A Refinement of the Structure of N-Acetylglycine*

By Jerry Donohue

Department of Chemistry, University of Southern California, Los Angeles, California, U.S.A.

and Richard E. Marsh

Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California, U.S.A.
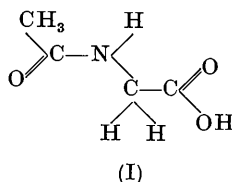
The crystal structure of N-acetylglycine, $C_4H_7NO_3$, which was originally determined by Carpenter & Donohue in 1950, has been refined by three-dimensional least-squares methods. Coordinates and individual anisotropic temperature parameters of the carbon, nitrogen, and oxygen atoms as well as the coordinates of four hydrogen atoms were included in the refinement. The remaining three hydrogen atoms of the methyl group are disordered, the methyl group apparently being free to assume all orientations about the C–C bond.

The bond distances are very close to the values reported earlier, but the estimated standard deviations in the distances have been reduced from about 0·015 to about 0·006 Å.

Introduction

The crystal structure of N-acetylglycine (I) was determined by

$$\begin{array}{c} CH_3 \qquad H \\ \diagdown C - N \diagup \\ O \diagup \qquad \diagdown C - C \diagup O \\ \diagup \diagdown \diagdown OH \\ H \quad H \end{array}$$

(I)

Carpenter & Donohue (C.D.) (1950). They obtained complete Cu $K\alpha$ intensity data from Weissenberg photographs about three different crystallographic axes, determined the approximate structure from Patterson projections and refined the positional parameters by Fourier methods. They included a single anisotropic temperature factor and the contributions of the four non-methyl hydrogen atoms in the calculated structure factors, but were unable to determine the positions of the hydrogen atoms of the methyl group. Their final R factor was 0·134 and they estimated a probable error in the bond distances of about 0·01 Å.

With the availability of high-speed digital computing facilities, we have thought it worth while to extend the refinement of the structure to take account of anisotropic temperature vibrations of the individual atoms and to determine, if possible, the locations of the methyl hydrogen atoms.

Refinement procedure

N-acetylglycine is monoclinic, space group $P2_1/c$, with the following lattice constants as determined by C.D.:

$$a_0 = 4·86, \quad b_0 = 11·54, \quad c_0 = 14·63 \text{ Å}$$
$$(\lambda(\text{Cu } K\alpha) = 1·542 \text{ Å}); \quad \beta = 138° 12';$$

the maximum error is given as 0·2%. The density measured by C.D. is 1·43 g.cm.$^{-3}$; the density calculated on the basis of four molecules in the unit cell is 1·422 g.cm.$^{-3}$.

Table 1. *Corrections to Table 3 of Carpenter & Donohue*

	F_o	
hkl	Reported by C.D.	Corrected value
169	—	4·3
254	< 3·9	3·9
$\bar{1}$,1,12	< 3·0	3·0
$\bar{2}$,11,9	3·1	< 3·1
$\bar{5}$04	< 0·7	0·7
$\bar{5}$,1,11	< 2·9	2·9
$\bar{5}$,1,12	< 2·9	2·9
$\bar{6}$,1,15	< 1·7	1·7

We have used as the basis of our refinement the observed structure factors of C.D. (their Table 3) with a few corrections based on examination of the original notebooks as noted here in Table 1. The starting parameters were those of their Table 2 with an over-all anisotropic temperature factor of $\exp - (0·20h + 0·045l)^2$ (C.D.) on top of an isotropic factor with $B = 2·36$ Å^2 (used but not reported by C.D.). The refinement consisted of 18 three-dimensional

least-squares cycles, the first eight being carried out on a Datatron 205 computer and the remaining on a Burroughs 220 computer; three three-dimensional difference maps were also calculated in attempts to locate the various hydrogen atoms.

In order to establish an appropriate weighting system for the least-squares calculations we have made comparisons of the observed F values for all reflections measured about two or three different rotation axes. The results of these comparisons are summarized in Fig. 1. In this figure are plotted the average deviations $\overline{\Delta F}$ of the two or three different observations from the average value F, the individual deviations being averaged over small regions of F. We have fitted the points in two different ways: the dashed curve in Fig. 1 is the function $\overline{\Delta F} = 0.15 + 6 \times 10^{-5} F^3$ and the solid lines represent the relationship

$$F \leq 6: \quad \overline{\Delta F} = 0.158;$$
$$6 \leq F \leq 23: \quad \overline{\Delta F} = 0.0264F;$$
$$F \geq 23: \quad \overline{\Delta F} = 0.181F - 3.556 .$$

The weighting factor $\sqrt{w}$ was taken inversely proportional to this second relationship (with an additional factor denoting the number of observations) in the eight least-squares cycles carried out on the Datatron 205 computer; in the cycles carried out on the Burroughs 220 computer the quantity being minimized was $\Sigma w (F_o^2 - F_c^2)^2$ and the weighting function was chosen as $\sqrt{w} = 1/\sigma(F^2) = 0.6(0.3F + 1.2 \times 10^{-4} F^4)^{-1}$. (The average deviations in Fig. 1 are approximately 0.6 times as large as the estimated standard deviations calculated according to the prescription given by Ibers (1956)). Unobserved reflections were included
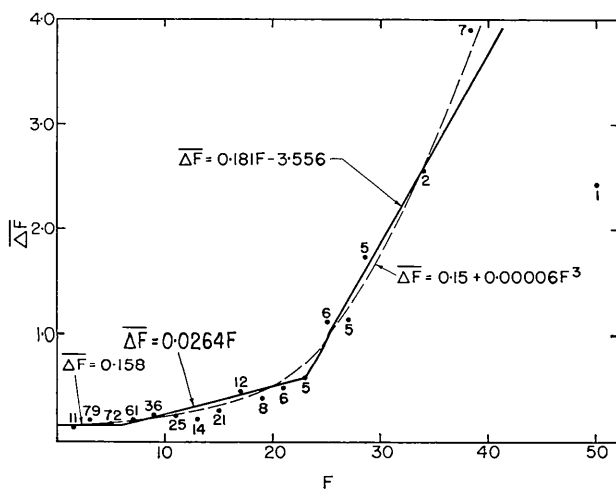

Fig. 1. The pattern of observational discrepancies. Here, ΔF is the deviation of a single observation from the average value of F for that reflection; $\overline{\Delta F}$ is the average over a number of reflections of approximately the same magnitude. The number included in each average is shown adjacent to the plotted points. The lines represent the functions on which the least-squares weights were based (see text).

in the least-squares totals only when F_c was greater than F_o, in which case ΔF was taken as the difference between the two. Atomic form factors for C, N, and O were taken as the average of the Berghuis et al. (1955) and the Hoerni & Ibers (1954) values; the hydrogen form factor was that of McWeeny (1951).

During the first two least-squares cycles only the positional parameters of the eight heavy atoms were adjusted; the R factor dropped from 0.14 to 0.12 and the sum of the squares of the weighted residuals dropped by 80%. Coordinates of the four non-methyl hydrogen atoms and corrections to the individual anisotropic temperature factors were obtained from a difference map and were included as parameters in the subsequent least-squares calculations, the hydrogen atoms being assigned a constant isotropic temperature factor with $B = 3.0$ Å². During the next twelve cycles the R factor dropped to 0.088 and the sum of the squares of the weighted residuals dropped by another 80%.

A disturbing feature of the refinement up to this point was the tendency of the carboxyl hydrogen atom (H_4) to move along the c axis away from the carboxyl atom O_1 (to which it is presumably bonded) towards the carbonyl oxygen atom O_3, until finally the O_1–H_4 distance was 1.4 Å and the O_3–H distance was only 1.1 Å. The suggested tautomeric structure, which would imply a zwitterion, seemed improbable. A survey of the observed and calculated structure factors turned up a number of fairly weak (and, hence, high-weight) reflections showing rather poor agreement and which are particularly sensitive to the z coordinate of H_4, prominent among them being $0k2$ and $\bar{1}k2$ reflections with k even. These planes are also highly sensitive to the inclusion of the hydrogen atoms of the methyl group, whatever their orientation about the C–C bond, and, hence, it seemed likely that the peculiar z coordinate of the H_4 atom was related to the omission of the methyl hydrogen atoms from the calculated structure factors.

Accordingly, a three-dimensional difference map was computed, the contributions of all seven hydrogen atoms being omitted from the calculated structure factors. The four non-methyl hydrogen atoms showed up clearly, H_4 being approximately in its proper position (that is, close to O_1). On the circle defining the locus of possible positions for the methyl hydrogen atoms, however, the electron density was low and diffuse, varying between 0.1 and 0.3 e.Å^{-3} with no pronounced maxima. It seemed apparent, then, that the methyl group is not in a fixed orientation.

In order to approximate the scattering from the methyl group in a disordered and presumably random orientation, we introduced twelve atoms each having one-quarter the scattering power of a hydrogen atom. They were positioned equally spaced on a circle defining a tetrahedral configuration about the C_3–C_4 bond at a distance of 1.0 Å from C_4. All twelve 'atoms' were assigned isotropic temperature factors with

$B = 4\cdot0$ Å², and their parameters were held constant through the remaining calculations.

Table 2. The final positional parameters and their estimated standard deviations

All values have been multiplied by 10^4

	D.M.			C.D.		
Atom	$x(\sigma_x)$	$y(\sigma_y)$	$z(\sigma_z)$	x	y	z
C_1	0113(12)	7169(3)	5170(4)	0113	7167	5165
C_2	0039(12)	7295(4)	4122(4)	0031	7293	4120
C_3	−0178(13)	6084(4)	2730(4)	−0167	6087	2739
C_4	−0546(19)	4900(4)	2234(6)	−0493	4900	2242
N	−0178(10)	6171(3)	3637(3)	−0172	6165	3640
O_1	0089(9)	8184(3)	5565(3)	0099	8179	5570
O_2	0199(11)	6255(3)	5576(3)	0192	6262	5577
O_3	0052(11)	6968(3)	2306(3)	0053	6966	2304

Table 3. Coordinates of the hydrogen atoms

The position of the 12 one-quarter atoms of the methyl group were assumed and no standard deviations are given. All values have been multiplied by 10^3

Atom	$x(\sigma_x)$	$y(\sigma_y)$	$z(\sigma_z)$
$H_1(C_2)$	−240(13)	776(5)	340(4)
$H_2(C_2)$	262(14)	776(5)	458(5)
$H_3(N)$	−001(12)	549(5)	398(4)
$H_4(O_1)$	007(12)	810(4)	620(4)
$\frac{1}{4} H_5$	−053	496	155
$\frac{1}{4} H_6$	091	489	198
$\frac{1}{4} H_7$	195	474	245
$\frac{1}{4} H_8$	230	457	284
$\frac{1}{4} H_9$	187	442	303
$\frac{1}{4} H_{10}$	078	432	297
$\frac{1}{4} H_{11}$ (C₄)	−068	431	269
$\frac{1}{4} H_{12}$	−212	438	226
$\frac{1}{4} H_{13}$	−315	453	179
$\frac{1}{4} H_{14}$	−350	470	141
$\frac{1}{4} H_{15}$	−308	486	122
$\frac{1}{4} H_{16}$	−199	495	127

The inclusion of the methyl hydrogen atoms led to a marked improvement between calculated and observed structure factors, particularly for the $0k2$ and $\bar{1}k2$ reflections mentioned earlier. After four least-squares cycles the sum of squares of the weighted residuals had dropped by 20% and the R factor was 0·065; more encouraging, the hydrogen atom H_4 settled down at a reasonable position near O_1. And thus endeth the refinement. The final parameters and their estimated standard deviations as calculated from the residuals and the diagonal terms of the inverse

matrix of the normal equations* are given in Tables 2, 3, and 4.

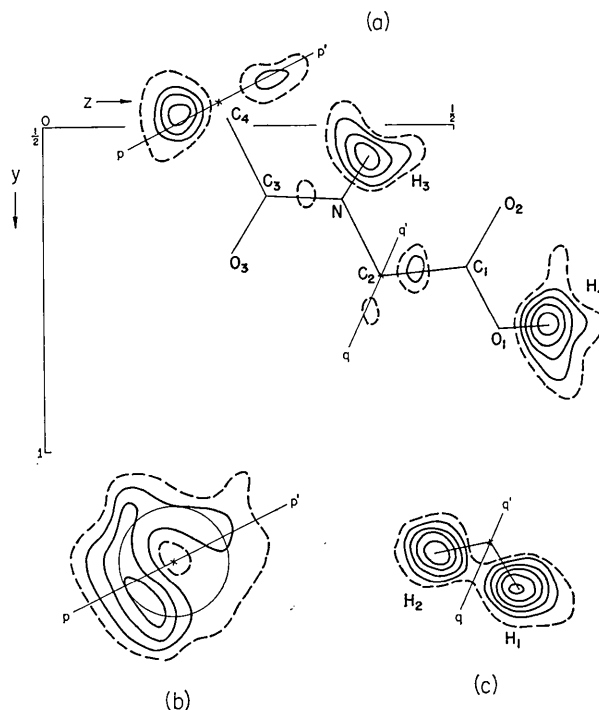At the conclusion of the refinement a three-dimensional difference map was calculated, again omitting

Fig. 2. Portions of a three-dimensional difference map, calculated near the end of the refinement, for which the contributions of the hydrogen atoms were omitted from the calculated structure factors. (a) The section in the plane of the molecule, approximately at $x = 0$. (b) A section through the plane of the methyl hydrogen atoms. The circle has a radius of 1 Å. (c) A section through the carbon and hydrogen atoms of the methylene group. All contours are at intervals of 0·1 e.Å⁻³, with the 0·1 e.Å⁻³ contour dashed.

the contributions of the hydrogen atoms from the F_c values. Portions of this map are shown in Fig. 2.

The final scale factor applied to the data of Table 3

Table 4. The final temperature factor parameters and their estimated standard deviations

The temperature factors are given in the form $T_i = \exp -(B_{11}h^2 + B_{22}k^2 + B_{33}l^2 + B_{12}hk + B_{13}hl + B_{23}kl)$

All values have been multiplied by 10^4

Atom	$B_{11}(\sigma)$	$B_{22}(\sigma)$	$B_{33}(\sigma)$	$B_{12}(\sigma)$	$B_{13}(\sigma)$	$B_{23}(\sigma)$
C_1	842(38)	54(3)	74(4)	1(20)	406(23)	1(6)
C_2	1043(51)	50(3)	87(5)	−35(22)	512(28)	−5(7)
C_3	889(49)	52(3)	78(5)	−7(22)	430(30)	2(7)
C_4	1691(99)	54(4)	131(10)	−15(34)	824(59)	−8(10)
N	1034(45)	48(3)	82(4)	−29(19)	498(25)	−2(6)
O_1	1292(40)	59(3)	102(4)	−6(16)	632(23)	−12(5)
O_2	1885(54)	58(3)	138(5)	13(18)	921(30)	19(6)
O_3	1405(48)	58(3)	110(5)	−30(19)	695(28)	8(6)

Table 5. *Bond distances and angles*

The values in parentheses have been corrected for libration effects implied by the out-of-plane temperature vibrations

	C.D. ($\sigma \cong 0.015$ Å)	D.M. ($\sigma \cong 0.006$ Å)		C.D. ($\sigma \cong 1.5°$)	D.M. ($\sigma \cong 0.5°$)
C_1-C_2	1·506 Å	1·514 Å	$O_1-C_1-C_2$	111·5°	111·0°
C_2-N	1·448	1·444	$O_1-C_1-O_2$	124·2	125·4
$N-C_3$	1·323	1·331	$O_2-C_1-C_2$	124·3	123·6
C_3-C_4	1·503	1·493 (1·507)	C_1-C_2-N	110·2	110·4
C_1-O_1	1·312	1·310 (1·319)	C_2-N-C_3	119·6	120·2
C_1-O_2	1·192	1·196 (1·219)	$N-C_3-C_4$	117·7	117·7
C_3-O_3	1·243	1·241 (1·252)	$N-C_3-O_3$	121·3	120·2
			$O_3-C_3-C_4$	121·0	122·1
$O_1 \cdots O_3'$	2·558	2·567	$C_1-O_1 \cdots O_3'$	113·3	112·7
$N \cdots O_2''$	3·030	3·030	$C_2-N \cdots O_2''$	131·8	131·6
			$C_3-N \cdots O_2''$	108·5	108·1

of C.D. by the least-squares process was 0·98, with an estimated standard deviation of 0·02.

Discussion of the results

The final R factor for 813 observed reflections of non-zero weight is 0·065, somewhat higher than the value of about 0·04 which we would have expected on the basis of the comparison between intensities measured about more than one axis. More significantly, the standard deviation of an observation of unit weight (Peterson & Levy, 1957) is calculated to be 2·0 as compared to the ideal value 1·0. This discrepancy implies either that some residual error exists in the model we have derived for the structure of N-acetyl-glycine or that the errors in the observed data are larger than the intensity comparisons indicated. We do not know which of the two alternatives is the more important, although the experience of many investigators using the visual method of estimating intensities might suggest that we have over-estimated the reliability of the observations. Purely subjectively, we feel that the estimated standard deviations in the parameters as given in Tables 2, 3, and 4 are reasonable, and we have used them in deriving the estimated standard deviations in the bond distances and angles given in Table 5.

Although our estimated standard deviations are less than half as large as those of C.D., our refinement has not produced any significant changes in the bond distances or angles (see Fig. 3 and Table 5)—which we find both disappointing and comforting. Indeed, the corrections which we have made to the bond lengths on the basis of the molecular librations implied by the temperature parameters (and surely on the basis of an incomplete understanding of these librations) are considerably greater than the differences between the two sets of distances. In making

Table 6. *Bond distances and angles involving the hydrogen atoms*

The estimated standard deviations in the positions of the hydrogen atoms are approximately 0·1 Å

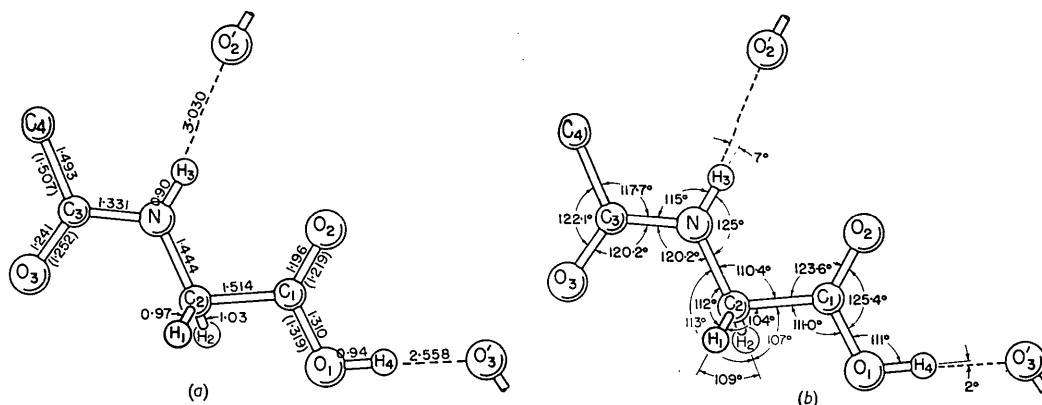|			
C_2-H_1	0·97 Å	$C_1-C_2-H_1$	104°
C_2-H_2	1·03	$C_1-C_2-H_2$	107
$N-H_3$	0·90	$N-C_2-H_1$	112
O_1-H_4	0·94	$N-C_2-H_2$	113
		$H_1-C_2-H_2$	109
		C_2-N-H_3	125
		C_3-N-H_3	115
		$C_1-O_1-H_4$	111
		$O_1-H_4 \cdots O_3'$	177
		$H_4-O_1 \cdots O_3'$	2
$H_4 \cdots O_3'$	1·63	$N-H_3 \cdots O_2''$	169
$H_3 \cdots O_2''$	2·14	$H_3-N \cdots O_2''$	7

Fig. 3. (*a*) The final bond distances. Values in parentheses have been corrected for librational effects. (*b*) The bond angles.

Table 7. *Magnitudes B_i and direction cosines q_i^j relative to a^*, b and c of the principal axes of the temperature factor ellipsoids*

Atom	B_1	q_1^j	B_2	q_2^j	B_3	q_3^j
C_1	3·54	0·997	2·88	−0·007	2·17	0·078
		0·010		0·999		−0·040
		−0·078		0·041		0·996
C_2	4·42	0·989	2·64	0·151	2·03	−0·004
		−0·149		0·971		−0·189
		−0·025		0·187		0·982
C_3	3·72	0·995	2·78	0·079	2·23	0·057
		−0·063		0·968		−0·242
		−0·074		0·238		0·969
C_4	7·11	0·999	2·94	0·011	2·57	0·036
		−0·026		0·896		0·444
		−0·027		−0·445		0·895
N	4·38	0·990	2·59	0·138	1·86	0·034
		−0·126		0·963		−0·238
		−0·066		0·232		0·971
O_1	5·43	1·000	3·24	0·006	2·02	0·026
		−0·014		0·948		0·318
		−0·023		−0·318		0·948
O_2	7·92	1·000	3·32	−0·008	1·98	0·026
		0·018		0·903		−0·429
		−0·020		0·429		0·903
O_3	5·92	0·996	3·29	0·082	1·81	−0·018
		−0·083		0·924		−0·374
		−0·014		0·374		0·927

the corrections for libration we have taken into account only the out-of-plane motions of the atoms and have assumed that the differences in mean square amplitudes between the four peripheral atoms and the atom to which they are bonded (C_1 or C_3) are the libration amplitudes, C_1 and C_3 being assumed the centers of rocking.

The magnitudes and direction cosines relative to a^*, b, and c of the principal axes of thermal motion are given in Table 7. For all eight heavy atoms the direction of principal motion is very close to normal to the bc plane—that is, to the plane of the molecule. This is not surprising, since the strongest inter-molecular forces—the hydrogen bonds—all lie in this plane. The out-of-plane motions cannot be explained on the basis of rigid-body libration, as the smallest amplitudes are associated not with the central atoms C_2 and N but with their neighbors C_1 and C_3. The in-plane anisotropies can be qualitatively explained by a small rigid-body rocking motion about an axis passing through the mid-point of the C_2–N bond and a larger translational motion along b (see Fig. 4).

As was pointed out by C.D., the heavy atoms are significantly non-coplanar, deviations of individual atoms from the best molecular plane being as much as 0·08 Å. The twist leading to the non-planarity is about the C_1–C_2 and the C_2–N single bonds, in approximately equal amounts. The two sets of terminal atoms C_1, C_2, O_1, O_2, and C_2, N, C_3, C_4, O_3 are each coplanar within experimental error, the maximum deviation being 0·020 Å for C_4. The dihedral angle between the two planes is 5·2°, the same value as reported by C.D.
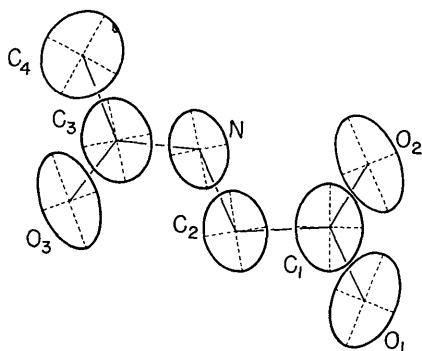

Fig. 4. The temperature-factor ellipsoids in the plane of the molecule. The lengths of the axes are proportional to the mean square amplitudes of vibration. For each of the eight atoms the third (major) axis is within 10° of perpendicular to the plane of the molecule.

References

BERGHUIS, J., HAANAPPEL, IJ. M., POTTERS, M., LOOP-STRA, B. O., MACGILLAVRY, C. H. & VEENENDAAL, A.L. (1955). *Acta Cryst.* 8, 478.

CARPENTER, G. B. & DONOHUE, J. (1950). *J. Amer. Chem. Soc.* 72, 2315.

HOERNI, J. A. & IBERS, J. A. (1954). *Acta Cryst.* 7, 744.

IBERS, J. A. (1956). *Acta Cryst.* 9, 652.

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

PETERSON, S. W. & LEVY, H. A. (1957). *Acta Cryst.* 10, 70.