

The Crystal Structure of L-Alanine

Harry James Simpson, Jr.

March 22, 1965

Abstract

The crystal structure of L-alanine, $C_3H_7O_2N$, has been determined and refined by analysis of complete three-dimensional diffraction data from copper X-radiation. The crystals are orthorhombic with space group $P2_1P2_1P2_1$; the unit cell dimensions are $a = 6.032$, $b = 12.343$, $c = 5.784$ Å. Least-squares refinement of the positional parameters of all 13 atoms and of the individual anisotropic temperature factors of the six heavy atoms, using 581 observed reflections of non-zero weight, yielded a final R factor of 0.049.

The structure bears a striking resemblance to that of DL-alanine, and involves the use of all available protons in hydrogen bonding, with lengths of 2.83, 2.85, and 2.81 Å.

Introduction

The determination of the crystal structure of l-alanine was undertaken as part of a program of research on the structures of compounds of biological importance. Particular interest in this compound arose because of an unusual similarity to dl-alanine, the structure of which has been previously determined by Levy & Corey (1941). In an earlier determination of cell dimensions by Bernal (1931), the values for l- and dl-alanine were found to be quite close. The cell dimensions determined for dl-alanine in the later and more exact investigation by Levy & Corey (1941) were in good agreement with Bernal's work. The unit cell of both compounds was determined to be orthorhombic, with dimensions as listed in Table 1.

In a preliminary investigation for the present work, it was found by examining intensity photographs of the 0-layers, taken about the c-axis for both compounds, that the structures of l- and dl-alanine are almost identical in projection.

From the differences in packing properties which must exist for these two species, it was felt to be surprising that their cell dimensions and structures along one projection could be so similar. The structure of l-alanine was determined to provide a comparison with that of dl-alanine.

Cell dimension determination

Crystals of L-alanine were obtained by several procedures, including evaporation of saturated methanol solutions, crystallization from slowly cooling, supersaturated water solutions, and slow evaporation of saturated water solutions. The latter was the most successful method, producing well-formed needle crystals, elongated along c, which are stable for periods of at least several months.

Straumanis-type rotation photographs about the c-axis were taken to determine dimensions of the a- and b-axes. A total of twenty-three reflections were indexed, and a least-squares analysis was made using ten high-angle reflections resulting from Cu K α_1 radiation ($\lambda = 1.54050 \text{ \AA}$).

A second crystal was mounted perpendicular to the needle axis and to one of the other parallel faces of the crystal. Alignment was made, with the aid of Laue photographs and a reciprocal lattice plot, along an axis perpendicular to the 110 plane. A total of thirty-one reflections were indexed on a Straumanis-type rotation photograph. Seven high-angle reflections, three of which were Cu K α_2 reflections ($\lambda = 1.54434 \text{ \AA}$), were used for a least-squares analysis to determine values for the c dimension and the quantity $(\frac{1}{a^2} + \frac{1}{b^2})$. This latter value, when compared with the previously determined a and b values indicated that either a or b or both were now significantly longer. Since the photographs around the two axes were taken under different temperature conditions, thermal expansion was postulated as an explanation for the discrepancy. To test this postulate, a new photograph of the L-alanine crystal mounted along the c-axis was made under the same temperature conditions as the $[110]$ photograph. Least-squares analysis yielded a significantly longer value for a than in the first determination and a b dimension identical

with the earlier determination. This indicated that nearly all the thermal expansion occurred along the a-axis. The second and third photographs were made under measured temperature conditions, and hence provide the basis for the best set of values for the determination. The calculations were performed by hand and yielded the results shown in Table 2.

Experimental

Complete intensity data from Cu K α radiation were collected by the multi-film, equi-inclination Weissenberg technique. The same crystal used for photographs 1 and 3 in the cell parameter determination was used to record layers 0 to 5 about the c-axis. A third crystal was mounted along the a-axis and shaped approximately to a cylinder by means of water soaked filter paper, since cleaving was found to destroy the crystal for data collection purposes. Layers 0 to 5 were then recorded about a.

In addition, a dl-alanine crystal was mounted and hk0 data collected about c for later use in a difference Fourier analysis to help solve the structure.

The intensities were estimated visually and corrected for Lorentz and polarization factors; no absorption correction seemed necessary.

All the reflections in an effective copper sphere of $\sin \theta = 0.99$ were recorded. Of the 589 independent reflections, 434 were recorded about both a- and c-axes. Eight reflections were not scaled, seven because they were too strong to estimate, and one because of lying too near the edge of the sphere. The 224 and 320 reflections were given reduced weight because of strong evidence of double reflection occurring. Renninger spots resulting from double reflection were observed on both the 2kl and 3kl photographs, and the 224 and 320 reflections both calculated much weaker than the observed values of 224 and 320 about this axis. Consequently, only one measurement for each of these reflections, that about the c-axis, was used in the refinement. All reflections recorded about only one axis were given one-half weight.

The 434 reflections measured on two axes were used to put all the data on a single scale. These 434 reflections also served as a basis for an estimation of the quality of the observed data. The final weighting scheme was

determined by making a plot of $\sigma(F^2)$ against F . The points were fitted to the function $\sigma(F^2) = R + SF_0 + TF_0^2$, resulting in a best fit curve of:

$$\sigma(F^2) = .25 + 0.083 F_0^2$$

This relationship was the basis of the weighting scheme used in the final least-squares refinement cycles.

The systematic absence of the odd orders of $h00$, $0k0$, and $00l$ reflections determined the space group as $P2_1P2_1P2_1$ (D_2^4). Density considerations determined the number of molecules per unit cell to be four.

Determination of the structure

A trial structure was developed from considerations of the similarities to the structure of dl-alanine. Since the distribution of intensities in the hk0 projections of l- and dl-alanine are so similar, the phases of the $|F|$'s for l-alanine were initially assigned the same as for dl-alanine (Levy & Corey), except for a few questionable cases, where calculations were necessary to determine the phases. A Fourier projection down the c-axis was made, using Beevers-Lipsom strips, with these assigned phases. The locations of the dl-alanine atoms were superimposed on this plot to see if any systematic rotation of the molecule could bring the atomic positions into coincidence with the peaks of the Fourier projection. Due to extensive overlap of the atoms in this projection, no conclusive possible shifts were found. Then a Fourier projection using $(F_{01} - F_1)$ as coefficients was tried. The F_{01} values were derived from the hk0 data collected from the dl-alanine crystal mentioned earlier. This difference map was used to suggest parameter changes to solve the structure of l-alanine. The calculation of the difference Fourier was made on a Burroughs 220 computer, using a Fourier program developed by Dr. K. Hoogsteen.

After making the indicated changes in the parameters, a structure-factor least-squares program, developed by Dr. A. Hybl, was run. Using the hk0 data, the R factor remained at about 0.23 for three cycles. Isotropic temperature factors were used and the positions of the hydrogen atoms, as determined from the dl-alanine paper, were included in the structure factors, but not the least-squares. Again the overlap of the atoms in this projection prevented the convergence of the calculated structure.

Next structure-factor least-squares calculations were tried using hk1

and hk2 data and only the six heavy atoms. The R factor dropped from 0.24 to 0.14 in four cycles. The refinement was continued using complete scaled data collected around two axes, again including only the six heavy atoms. After the R factor reached 0.13 with the complete data, the four non-methyl hydrogens were added, but were not allowed to shift positions. Their positions were determined by calculating approximately what locations they would occupy, the amino hydrogens being placed in the best positions for hydrogen bonding. The R factor went from 0.12 to 0.096 in seven cycles. Then a difference Fourier in the plane of the methyl hydrogens was calculated to determine their positions. The methyl hydrogens were located on a circle of radius = 1 Å, centered on the extension of the α -carbon--methyl carbon bond. The final refined positions are shown in Figure 1, along with the calculated difference Fourier and the constructed circle. After locating these hydrogens, they were added to the structure factor calculations and three more cycles were run, dropping the R factor from 0.079 to 0.061. From this point on in the refinement, the hydrogen positions were allowed to shift.

Throughout the early part of refinement, a weighting function of $\frac{1}{1 + F}$ was used. By the time the methyl hydrogens were placed in the refinement, the weighting function was changed to $\frac{1}{1 + F^2}$. The final weighting function, $w = \frac{1}{0.40 + 0.083F_o^2}$, as determined by the uncertainty in the intensity data and a scaling factor introduced in the refinement, is discussed in the experimental section. This final weighting function was introduced in the last few cycles of refinement and resulted in significant improvement of the agreement. After the hydrogen positions had stabilized somewhat, the hydrogen isotropic temperature factors were also allowed to shift. After several cycles, in which the hydrogen temperature factors did not make sensible shifts, the hydrogen temperature factors were assigned values of 3.00 for the methyl

hydrogens and 1.00 for the remaining hydrogens. These values were placed in the parameters, but were not allowed to shift. After these adjustments in hydrogen temperature factors were made the least-squares refinement was continued until all shifts were less than 30% of their standard deviations, resulting in a final R factor of 0.049. The observed and calculated structure factors are listed in Table 3.

Accuracy of the results

The final value of the quantity $\left(\frac{\sum w(F^2)^2}{m - s} \right)$ --- the standard deviation of an observation of unit weight --- is 1.36. The standard deviations in the positional parameters (Table 4) were calculated from the sum of the residuals and the diagonal terms of the least squares normal equations. The heavy atom positional parameters show an average standard deviation of about 0.002 Å, leading to a standard deviation in interatomic distance of about 0.003 Å, and a limit of error (chosen as three times the standard deviation) of 0.009 Å. The limit of error in a bond angle is about 0.5°. For bonds involving hydrogen atoms, the limits of error are about 0.09 Å in bond length and 5° in bond angle.

The standard deviations of the temperature factor parameters of the heavy atoms were calculated from the diagonal coefficients of the inverse matrices of the normal equations and include interactions among the six temperature parameters of the same atom.

Discussion of the structure

(i) The intermolecular environment

The structure of l-alanine as viewed down the b-axis is shown in Figure 2a, with the structure of dl-alanine, as determined by Levy & Corey, viewed down the a-axis, in Figure 2b. (Note the difference in axial labeling -- the present paper returns to the original labeling of Bernal). The structures show great similarities in these projections.

Hydrogen bonds are formed by all three available protons in both cases-- two from the amino nitrogen atom to O(1) atoms in adjacent molecules, and one from the amino nitrogen atom to an O(2) atom. The hydrogen bonding is almost exactly the same in both structures, with the only difference being a slight change in the angle of attack of one of the hydrogen bonds to O(1). The hydrogen bond distances are the same within the limits of error in the more approximate dl-alanine structure determination.

When viewed as in the two projections shown, the structure of l-alanine can be derived from the dl-alanine structure by taking the column of d-alanine molecules running along the c-axis in the dl-alanine structure, changing to l-alanine molecules, inverting lengthwise and translating along the c-axis until the hydrogen bonds coincide with those of the parent dl-alanine structure.

When viewed down the c-axis, the structures are essentially identical, giving rise to the identical hk0 intensity distributions which originally prompted this investigation.

(ii) The temperature factors

The magnitudes and direction cosines of the principal axes of thermal motion of the heavy atoms, derived from the anisotropic temperature parameters of Table 4, are listed in Table 5.

The largest thermal motions, as one would expect, are associated with the oxygen atoms and the methyl carbon. The motions of the other two carbon atoms and the nitrogen atom are more nearly isotropic and are of a smaller amplitude than the maximum movements of the first three mentioned.

The maximum thermal motion for both oxygen atoms is an out-of-plane vibration, which lies very close to perpendicular to the respective C--O bonds. The minimum thermal motion lies at an appreciable angle to the direction of the bond.

The maximum thermal motion for the methyl group is again perpendicular to the C--C bond. The minimum motion of C(3) also makes a sizable angle with the bond direction.

Since the maximum thermal motions of these atoms nearly perpendicular to their respective bonds (O(1), O(2), and C(3) form angles of 79° , 82° , and 80° , respectively, with their bonding directions), it is felt to be legitimate to correct the bond distances involved for implied rocking effects.

As mentioned in the discussion of the structure determination, the hydrogens were assigned isotropic temperature factors near the end of the refinement, in keeping with the relative amounts of motion expected for each hydrogen.

One other noticeable trend is the small magnitude of the thermal motions of all the heavy atoms in the direction of c. Since the long needle axis lies in this direction, there is the possibility of greater order existing in this crystallographic direction, thus reducing the noise level in the structure,

which would be reflected in smaller temperature factors in the direction of greater ordering. Strong hydrogen bond chains also exist in this direction and provide a possible explanation for reduced thermal motion along c.

(iii) Geometry of the molecule

The carboxyl group and adjacent carbon atom are very nearly planar, with the α -carbon and the two oxygens lying $0.002 \text{ \AA}$ below the best fit plane and the carboxyl-carbon lying $0.006 \text{ \AA}$ above it, the standard deviation in the atomic positions being $0.002 \text{ \AA}$. A view down this best fit plane is shown in Figure 3 to indicate the conformation of the methyl and amino groups with respect to the plane of the carboxyl group. The deviations from the least-squares plane are shown in Table 6.

The bond distances and angles calculated from the parameters of Table 4 are shown in Figure 4 and listed in Table 7, along with the comparable angles and distances of the accurately determined structure of glycine (Marsh, 1958).

The values shown in parentheses are corrected for the thermal rocking motion perpendicular to the bond. No corrections were made other than for the maximum thermal motion of the three atoms showing the largest amplitude of motion.

The work described here was carried out as an undergraduate research project.

Table 1

Cell dimensions for l- and dl-alanine from previous determinations.

$\lambda = 1.539 \text{ \AA}$ for Cu K α doublet. (Note variation in axial labeling).

<u>Bernal (1931)</u>		<u>Levy & Corey (1941)</u>
<u>l-alanine</u>	<u>dl-alanine</u>	<u>dl-alanine</u>
a = 6.0 $\text{\AA}$	a = 6.0 $\text{\AA}$	a = 12.04 $\pm$ 0.01 $\text{\AA}$
b = 12.1	b = 12.0	b = 6.04 $\pm$ 0.01
c = 5.75	c = 5.8	c = 5.81 $\pm$ 0.01

Table 2

Summary of the cell dimension determination and a list of the best set of parameters for the experimental temperature conditions.

<u>Photo.</u> <u>number</u>	<u># of</u> <u>reflect.</u>	<u>axial</u> <u>align.</u>	<u>a</u> (Å)	<u>b</u> (Å)	<u>c</u> (Å)
1	10	<u>c</u>	$6.0302 \pm .0002$	$12.3431 \pm .0005$	
2	7	110	6.032 or	12.347	$5.7838 \pm .0008$
3	10	<u>c</u>	$6.0315 \pm .0003$	$12.3427 \pm .0007$	

Best set of values:

Temperature = $25 \pm 1^{\circ}\text{C}$

$a = 6.032 \pm 0.001 \text{ Å}$

$b = 12.343 \pm 0.001 \text{ Å}$

$c = 5.784 \pm 0.001 \text{ Å}$

Table 4

The final atomic parameters and their standard deviations. The heavy atom values have been multiplied by 10^4 . The temperature factors are 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 hydrogen parameters have been multiplied by 10^3 . The methyl hydrogens [H(1,2,3)] were assigned an isotropic B value of 3.0 and the remaining four values of 1.0.

Atom	x (σ_x)	y (σ_y)	z (σ_z)	$\alpha(\sigma_\alpha)$	$\beta(\sigma_\beta)$	$\gamma(\sigma_\gamma)$	$\delta(\sigma_\delta)$	$\epsilon(\sigma_\epsilon)$	$\eta(\sigma_\eta)$
O(1)	7287(3)	843(1)	6283(3)	206(6)	43(1)	150(6)	59(4)	-47(10)	16(4)
O(2)	4501(3)	1850(1)	7609(3)	227(7)	52(1)	91(5)	46(4)	18(10)	-10(3)
C(1)	5606(4)	1418(1)	6016(4)	149(6)	25(1)	119(7)	-6(4)	-18(10)	-8(4)
N	6560(3)	1382(1)	1856(3)	150(6)	36(1)	88(6)	5(4)	12(8)	2(4)
C(2)	4769(4)	1612(1)	3563(4)	138(6)	28(1)	98(7)	15(4)	5(10)	-4(4)
C(3)	2746(5)	915(2)	3025(5)	177(8)	58(2)	153(9)	-38(5)	-61(12)	-15(5)
H(1)	227(7)	108(3)	144(8)						
H(2)	153(7)	103(3)	410(9)						
H(3)	315(6)	14(3)	306(8)						
H(4)	442(5)	239(2)	346(6)						
H(5)	700(5)	64(3)	194(6)						
H(6)	769(6)	179(3)	223(6)						
H(7)	599(5)	150(2)	43(6)						

Table 5

The magnitudes b , root mean square displacements μ , and direction cosines q relative to the crystallographic axes, of the principal axes of the thermal vibration ellipsoids.

Atom	Axis i	$B_i (\text{\AA}^2)$	q_i^1	q_i^2	q_i^3	$\mu_i (\text{\AA})$
O(1)	1	3.74	0.778	0.624	-0.066	0.218
	2	2.37	-0.385	0.558	0.735	0.173
	3	1.59	-0.496	0.546	-0.675	0.142
O(2)	1	3.95	0.730	0.683	0.000	0.224
	2	2.59	-0.677	0.723	-0.135	0.183
	3	1.19	-0.093	0.099	0.991	0.123
C(1)	1	2.22	0.976	-0.107	-0.188	0.167
	2	1.68	-0.074	0.653	-0.754	0.146
	3	1.42	0.203	0.750	0.630	0.134
N	1	2.27	0.726	0.684	0.069	0.169
	2	2.10	-0.684	0.729	-0.039	0.163
	3	1.17	-0.077	-0.019	0.997	0.122
C(2)	1	2.14	0.872	0.490	0.005	0.164
	2	1.62	-0.479	0.854	-0.204	0.143
	3	1.30	-0.104	0.176	0.979	0.128
C(3)	1	3.80	-0.417	0.909	-0.009	0.219
	2	2.68	0.726	0.326	-0.606	0.184
	3	1.68	0.547	0.259	0.796	0.146

Table 6

Deviations from the least-squares plane through atoms O(1), O(2), C(1), and C(2).
 $0.5755X - 0.8154Y - 0.0625Z = 3.151$

<u>Atom</u>	<u>Deviation from plane (Å)</u>
O(1)	-0.002
O(2)	-0.002
C(1)	0.006
N	0.452
C(2)	-0.002
C(3)	-1.386
H(1)	-1.33
H(2)	-1.78
H(3)	-2.02
H(4)	0.66
H(5)	-0.14
H(6)	1.25
H(7)	0.43

Table 7a

Bond distances and angles. The average estimated standard deviations for bonds involving only heavy atoms are 0.003 Å and 0.2°, and for bonds involving hydrogen atoms the standard deviations are 0.03 Å and 2.0°.

C(1)-O(1)	1.245 Å (1.253 Å)	O(1)-C(1)-O(2)	125.6°
C(1)-O(2)	1.258 (1.267 Å)	O(2)-C(1)-C(2)	116.1
C(1)-C(2)	1.523	O(1)-C(1)-C(2)	118.3
C(2)-C(3)	1.526 (1.533 Å)	C(1)-C(2)-C(3)	111.6
C(2)-N	1.493	C(1)-C(2)-N	110.3
N-H(5)	0.95	N-C(2)-C(3)	109.8
N-H(6)	0.90	N-C(2)-H(4)	109
N-H(7)	0.88	C(3)-C(2)-H(4)	111
C(2)-H(4)	1.00	C(1)-C(2)-H(4)	105
C(3)-H(1)	0.98	H(5)-N-H(7)	112
C(3)-H(2)	1.00	H(5)-N-H(6)	107
C(3)-H(3)	0.99	H(6)-N-H(7)	117
Intermolecular hydrogen bond lengths		H(1)-C(3)-H(2)	108
N-O(1)	2.813 Å	H(1)-C(3)-H(3)	107
N-O(1)	2.827	H(2)-C(3)-H(3)	108
N-O(2)	2.853	C(2)-N-H(5)	108
H(5)-O(2)	1.94	C(2)-N-H(6)	106
H(6)-O(1)	1.99	C(2)-N-H(7)	107
H(7)-O(1)	1.94	C(2)-C(3)-H(1)	108
Intermolecular hydrogen bond angles		C(2)-C(3)-H(2)	115
O(2)-H(5)-N	161°	C(2)-C(3)-H(3)	109
O(1)-H(7)-N	174°		
O(1)-H(6)-N	154°		

Table 7b

Comparison of bond distances and angles in l-alanine and glycine

Bond Distance	l-alanine	glycine
C(1)-O(1)	1.245 Å (1.253 Å)	1.252 Å (1.261 Å)
C(2)-O(2)	1.258 (1.267 Å)	1.255 Å (1.265 Å)
C(1)-C(2)	1.523	1.523
C(2)-N	1.493	1.474

Bond Angle	l-alanine	glycine
O(1)-C(1)-O(2)	125.6°	125.5°
O(1)-C(1)-C(2)	118.3	117.4
O(2)-C(1)-C(2)	116.1	117.1
C(1)-C(2)-N	111.6	111.8

Legends

Figure 1: Difference Fourier in the plane of the methyl hydrogens. The contours are made at .2 .3, and .4 elect./Å². The constructed circle is of radius = 1 Å, to serve as a guide in the initial location of the methyl hydrogens. The circle is centered on the extension of the C(2)-C(3) bond. The final refined hydrogen positions are also indicated.

Figure 2a: The structure of l-alanine viewed down the b-axis. The dashed lines represent hydrogen bonds.

Figure 2b: The structure of dl-alanine viewed down the a-axis.

Figure 3: End view of the least-squares plane to show the conformation of the methyl and amino groups with respect to the carboxyl group.

Figure 4: Bond distances and angles. Bonds corrected for thermal vibration have corrected values in parentheses.

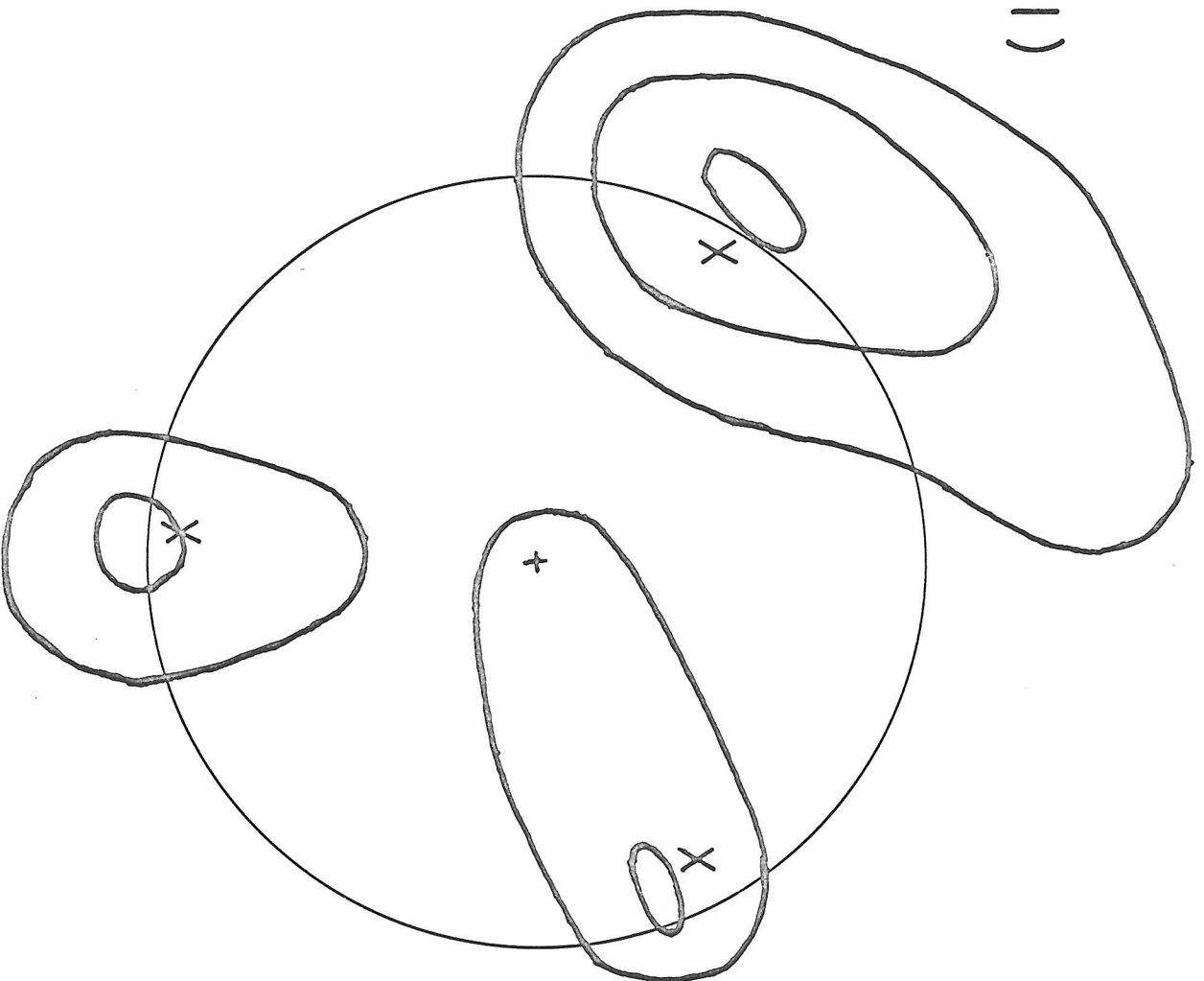
Table 3: Observed and calculated structure factors

Within each group the columns contain the values of h , $10F_o$, $|10F_c|$, $10A_c$ and $10B$. Reflections designated by an asterisk (*) were omitted from the least-squares and R factor calculations.

References

- Bernal, J. D., (1931), Z. Krist., 78, 363.
- Levy, Henri A. & Robert B. Corey, (1941), JACS, 63, 2095.
- Marsh, Richard E., (1958), Acta Cryst., 11, 654.

$H(1)$



$H(2)$

$H(3)$

