

The Crystal Structure of *l*-Lysine Monohydrochloride Dihydrate*

BY DONALD A. WRIGHT† AND RICHARD E. MARSH

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

(Received 27 February 1961)

The crystal structure of *l*-lysine monohydrochloride dihydrate has been determined by analysis of complete three-dimensional data from copper X-radiation. The approximate structure was obtained from a Patterson superposition map derived from the Cl-Cl vector. Refinement of the positional parameters for all 32 atoms and of the individual anisotropic temperature factors for the 13 heavier atoms was by the method of least squares. The final *R* factor for 1261 observed reflections of non-zero weight is 0.057, and the estimated limits of error are 0.015–0.018 Å in the bond distances involving the C, N, and O atoms and 33–40' in the bond angles.

The lysine molecule is a zwitterion, each nitrogen atom having an extra proton. The two C-NH₃⁺ distances are equal within experimental error, the average value being 1.482 Å. The side chain is in an extended, nearly planar configuration, and the average C-C distance is 1.524 Å. The apparent C-H, N-H, and O-H bond lengths are only slightly shorter than the standard internuclear separations.

Introduction

Lysine—NH₂(CH₂)₄CH(NH₂)COOH—is one of four essential amino acids having basic side chains. The detailed crystal structure analysis of lysine monohydrochloride dihydrate is one of a series of X-ray diffraction studies of protein derivatives that has been in progress in these Laboratories for many years. The results reported here reflect extensive three-dimensional refinement of the positions of all 32 atoms and of the anisotropic vibrational parameters of the 13 heavier atoms. The accuracy achieved allows structural features to be discussed with confidence.

An approximate solution of the structure was published during the course of this investigation by Raman (1959), and is in substantial agreement with the results found here. Raman did not, however, attempt complete refinement.

Experimental

l-Lysine monohydrochloride crystallizes readily as a dihydrate from saturated aqueous solutions allowed to evaporate at room temperature (20 °C). The colorless monoclinic crystals are stable for some months in dry air, particularly if the crystal faces are not damaged, before surface decomposition to a sticky white powder occurs, presumably due to loss of water. Though most crystals were twinned, a few single crystals were found in each crystallization, and were

used for the collection of the X-ray data. Serious radiation damage was noticed, particularly under hot humid conditions, but was greatly reduced by desiccation with silica gel during photography.

Complete intensity data for Cu *K*α radiation were collected by the multiple-film equi-inclination angle Weissenberg technique. As the linear absorption coefficient of the salt is 31.8 cm.⁻¹, considerable care was taken to choose crystals having regular shapes. They tended to be elongated along *c*, but if necessary they could be cut with a scalpel blade. No pronounced cleavage plane was observed. The five crystals used in this work were thick needles elongated about either *a* or *c*, and 0.2–0.4 mm. on the edge. All layer lines about the *a* and *c* axes with equi-inclination angles less than 45° (that is, layer lines 0 to 6 about *a* and 0 to 5 about *c*), as well as the zero layer about the *b* axis, were recorded with Cu *K*α radiation. The intensities were estimated visually and corrected as usual for Lorentz and polarization factors, no absorption corrections being applied. Of 1348 independent reflections within the effective copper sphere ($\sin \theta < 0.98$), 1284 were observed. Of these, 1024 reflections were measured about both the *a* and *c* axes and used to put all reflections on a single scale, as well as to give an estimate of the expected standard deviation, $\sigma(F)$, in an observed structure factor (Ibers, 1956). These standard deviations were averaged over regions of similar *F* values and fitted to the equation

$$\sigma(F) = a + b|F| + c|F|^2;$$

the resulting values for *a*, *b*, and *c* were 4.00, 2.34×10^{-3} , and 2.58×10^{-4} . The weighting schemes used in the least-squares refinement of the atomic parameters were based on this pattern of estimated uncertainties.

* Contribution No. 2681 from the Gates and Crellin Laboratories of Chemistry. This investigation was supported in part by a research grant No. H-2143 from the National Heart Institute, National Institutes of Health, Public Health Service.

† Present address: Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

Three-dimensional intensity data for molybdenum $K\alpha$ radiation were also collected (layer lines 0–8 about c , and 0–2 about a for correlation). Time did not allow further work with these data.

Table 1. *Unit-cell dimensions of l-lysine monohydrochloride dihydrate*
(Cu $K\alpha_1$, $\lambda=1.54051$)

	This investigation*	Shrinivasan (1956)†
a	$7.492 \pm 0.001 \text{ \AA}$	7.48
b	13.320 ± 0.004	13.31
c	5.879 ± 0.001	5.85
β	$97^\circ 47.4' \pm 0.7'$	$97^\circ 48'$

* Estimated errors set at about 3 times the standard deviations given by the least-squares analysis.

† Errors estimated at about 1%.

The unit-cell dimensions, given in Table 1, were determined from Straumanis-type single-crystal rotation photographs about b and c . Eighteen high-angle $h0l$ reflections ($\sin \theta > 0.76$) and six $hk0$ reflections ($\sin \theta > 0.84$) were indexed by comparison with the corresponding Weissenberg photographs, and their measured values of $\sin^2 \theta$ were used as input for a least-squares determination of the four cell parameters. The results are in good agreement with those of Srinivasan (1956), which are also listed in Table 1. The systematic absence of $0k0$ reflections with k odd confirms Srinivasan's choice of space group $P2_1$.

The density of newly precipitated crystals, measured by flotation in mixed solvents, is $1.250 \pm 0.003 \text{ g. cm.}^{-3}$, which corresponds to 1.91 molecules of water of hydration per lysine molecule. The calculated density for two molecules of the perfect dihydrate per unit cell, $1.259 \text{ g. cm.}^{-3}$, is just significantly higher. The apparent deficiency of 0.35 electrons for each water oxygen was not detected during the refinement, either in the least-squares calculations or in the difference Fourier syntheses. Any effect on the structure factors of such a slight deficiency would probably be masked by the temperature-factor parameters.

Determination of the structure

The first attempts to derive an approximate structure made use of the three sets of two-dimensional data. The position of the chlorine atom was readily deduced from the Patterson projections and used to calculate approximate phases for the zonal structure factors. However, the subsequent electron density projections could not be interpreted. The use of a single phase-determining atom introduced a false mirror plane of symmetry in the short a and c axis projections, making them too confused to be solved, and the (010) projection had severe overlap (indicating that the carbon chain might lie approximately parallel to the b axis). After testing and rejecting several structures, two-dimensional work was abandoned.

Two three-dimensional syntheses—an electron density map based on the phase angles of the chlorine atom (and which also contained a false mirror plane), and a sharpened Patterson map—were now calculated. The near-origin peaks of the Patterson map confirmed that the carbon-atom side chain of the lysine molecule extended along the b axis. However, this was not enough of a clue to allow the Fourier map to be readily interpreted. Accordingly the Patterson superposition map was plotted, using the Cl–Cl vector in the Harker section at $v = \frac{1}{2}$ as the second origin; only superpositions of Patterson peaks greater than 10% of the Cl–Cl peak height were noted. The correct structure was now obvious. All atoms except one water oxygen were immediately identified, and there were only four spurious peaks. The missing oxygen was found as a well-defined peak in the electron density map by considering the possible hydrogen-bonding network. It was now apparent that the electron density map had defined the atoms of the polar groupings, but not those of the carbon chain.

Finally, all coordinates were adjusted to give the best agreement with the Patterson vectors. Estimated isotropic temperature factors were assigned to all atoms and structure factors were calculated for the reflections out to $\sin \theta = 0.6$. The initial R factor was 0.22.

Refinement of the parameters

The structure, now essentially determined, was refined by a series of least-squares analyses of the complete Cu $K\alpha$ data. Atomic form factors of Berghuis *et al.* (1955) were used for C, N, O, and Cl⁻, and McWeeny (1951) values were used for hydrogen. First calculations were carried out on a Datatron 205 computer with a program written by Pasternak (1956) which accepted individual isotropic or anisotropic temperature factors in its structure-factor routine but only adjusted positional parameters in the least-squares calculations. Hence, preliminary adjustments of the temperature factors were estimated from three-dimensional difference Fourier syntheses. A constant weighting factor applied to all reflections in the first least-squares calculations was quickly rejected as inadequate, probably because of large observational errors in the large structure factors. The function being minimized was $\sum w(\Delta F^2)^2$, and for each reflection the weighting factor \sqrt{w} should be taken proportional to the reciprocal of the expected uncertainty in ΔF^2 . As large errors due to incorrect parameters were still present, a compromise factor, $\sqrt{w} = 1/\sigma(F)$, was chosen (see Experimental). An auxiliary weight, allowing for the number of independent observations of a given reflection and also giving low or zero weight to observations considered of low reliability, was also used. As this weighting scheme led to a rapid convergence it was retained until the last stages of refinement.

Four structure-factor least-squares cycles, during which the scale factor and an average isotropic tem-

perature parameter were adjusted graphically, reduced the R factor for all observed reflections of non-zero weight to 0.16. It was now noticed that, with two minor exceptions, every large structure factor calculated appreciably larger than its observed value. This effect was presumably due to extinction, and an empirical correction was found from the relation discussed by Hamilton (1957),

$$\ln(F_o^2/F_T^2) = kT_o,$$

where k is a secondary extinction factor and F_T , the true value of the structure factor, was assumed equal to F_o . A reliable value of k was found only for the crystal mounted about the a axis, allowing 56 reflections to be corrected. The minimum correction was 6% in F .

Actually, the apparent extinction effect could, in substantial measure, arise from another source—a low film factor in the multiple-film visual technique of estimating intensities. However, the form of the correction would still be valid and its use is justified empirically. When refinement was almost complete the corrections were recalculated, but no adjustment seemed feasible. The corrected structure factors were given half weight in all succeeding least-squares calculations, and all uncorrected strong reflections were given zero weight.

A series of five least-squares cycles and three difference-Fourier syntheses completed refinement on the Datatron 205. The difference maps were interpreted by the method of Leung *et al.* (1957). Individual anisotropic temperature factors were assigned to the chlorine and water oxygen atoms, and corrections calculated for the isotropic temperature factors of all other atoms. After these calculations the R factor had fallen to 0.102.

At this time the Datatron 205 computer was replaced by a faster computer, the Burrough's 220. Also during this period a solution of the structure was reported by Raman (1959), who solved the phase problem by considering the anomalous dispersion effects of the chlorine atom and carried through partial refinement of the two acentric projections (a and c). The structure he derived is in substantial agreement with that published here, though it should be noted that Raman's coordinates refer to a monoclinic unit cell with an acute, rather than obtuse, angle β . The sign of the y coordinate of his atom C_1 is reversed, and his analysis did not allow the hydrogen bonding network to be fully elucidated. The observed structure factors he reports are in general agreement with those used in this analysis with the exception of $6k0$ data with $k=2n$, where serious discrepancies occur.

The monoclinic structure-factor least-squares program written for the Burrough's 220 computer accepts and optimises ten parameters for each atom: a scale (population) parameter, three positional parameters and six vibrational parameters. The matrix of the normal equations is semi-diagonalized by neglecting

Table 2. The final parameters and their standard deviations

The standard deviations, obtained from the least-squares analysis, are multiplied by 10^4 . The thermal parameters are of the form $T_i = \exp[-(\alpha_i k^2 + \beta_i h^2 + \gamma_i l^2 + \delta_i hk + \epsilon_i hl + \eta_i kl)]$

Atom	x	σ_x	y	σ_y	z	σ_z	$10^4\alpha$	σ_α	$10^4\beta$	σ_β	$10^4\gamma$	σ_γ	$10^4\delta$	σ_δ	$10^4\epsilon$	σ_ϵ	$10^4\eta$	σ_η
O_1	0.3568	(4)	0.0596	(3)	-0.3904	(5)	174	(7)	64	(2)	180	(8)	5	(5)	73	(9)	13	(6)
O_2	0.1038	(4)	0.1213	(3)	-0.2393	(5)	153	(6)	64	(2)	255	(10)	36	(5)	22	(11)	60	(7)
C_1	0.2503	(5)	0.0761	(3)	-0.2483	(6)	130	(7)	47	(2)	167	(11)	-20	(6)	44	(11)	7	(7)
C_2	0.3026	(5)	0.0339	(3)	-0.0066	(6)	113	(7)	38	(2)	182	(11)	6	(5)	35	(11)	1	(6)
N_1	0.1982	(4)	0.0855	(2)	0.1560	(5)	148	(7)	43	(2)	193	(10)	14	(5)	66	(11)	-1	(6)
C_3	0.2721	(6)	-0.0793	(3)	-0.0082	(7)	183	(9)	33	(2)	251	(14)	6	(6)	58	(14)	2	(7)
C_4	0.3126	(6)	-0.1293	(3)	0.2250	(7)	186	(9)	42	(2)	235	(13)	9	(6)	39	(14)	20	(8)
C_5	0.2920	(6)	-0.2430	(3)	0.2044	(8)	218	(11)	42	(2)	261	(14)	-6	(7)	78	(16)	20	(8)
C_6	0.3034	(6)	-0.2927	(3)	0.4389	(7)	183	(9)	47	(2)	202	(12)	-1	(6)	33	(13)	10	(7)
N_2	0.2800	(5)	-0.4027	(3)	0.4158	(6)	174	(8)	43	(2)	229	(11)	-14	(5)	46	(12)	34	(7)
$W_1, (H_2O)_1^*$	0.1283	(8)	0.3156	(3)	0.5342	(8)	584	(19)	68	(3)	402	(17)	-59	(12)	278	(27)	22	(11)
$W_2, (H_2O)_2^*$	0.2945	(7)	0.2897	(3)	0.1563	(7)	465	(16)	52	(2)	368	(15)	1	(8)	153	(21)	46	(9)
Cl	0.2077	(1)	0.5003	(1)	-0.1137	(2)	163	(2)	56	(1)	242	(3)	8	(2)	87	(3)	30	(2)

* The alternate symbols W_1, W_2 are used throughout the succeeding tables.

Table 3. *The hydrogen atom parameters*

- (a) Positional parameters, and standard deviations, from the final least-squares analysis. The thermal parameters, B_i , where $T_i = \exp(-B_i \sin^2 \theta / \lambda^2)$, were not refined
 (b) Positional parameters from the final difference synthesis

Atom	Bonded to	(a)						(b)			
		x	$10^3 \sigma_x$	y	$10^3 \sigma_y$	z	$10^3 \sigma_z$	B	x	y	z
H ₁	(N ₁)	0.197	(8)	0.153	(5)	0.140	(10)	4.0	0.198	0.155	0.140
H ₂		0.263	(8)	0.076	(5)	0.317	(11)	4.0	0.266	0.077	0.318
H ₃		0.086	(8)	0.066	(5)	0.111	(10)	4.0	0.083	0.070	0.106
H ₄	(N ₂)	0.172	(9)	-0.421	(6)	0.291	(12)	4.5	0.168	-0.420	0.276
H ₅		0.380	(9)	-0.427	(6)	0.411	(11)	4.5	0.394	-0.431	0.397
H ₆		0.263	(8)	-0.430	(6)	0.565	(11)	4.5	0.267	-0.429	0.566
H ₇	(W ₁)	0.122	(13)	0.358	(9)	0.650	(15)	8.0	0.110	0.356	0.664
H ₈		0.129	(13)	0.248	(9)	0.612	(14)	8.0	0.108	0.254	0.602
H ₉	(W ₂)	0.291	(10)	0.354	(10)	0.063	(13)	7.0	0.301	0.353	0.068
H ₁₀		0.247	(11)	0.314	(7)	0.302	(15)	7.0	0.252	0.315	0.312
H ₁₁	(C ₆)	0.425	(9)	-0.282	(6)	0.543	(13)	5.0	0.436	-0.282	0.537
H ₁₂		0.221	(9)	-0.258	(7)	0.543	(11)	5.0	0.213	-0.252	0.532
H ₁₃	(C ₅)	0.414	(9)	-0.277	(6)	0.121	(11)	5.0	0.418	-0.274	0.123
H ₁₄		0.169	(9)	-0.270	(7)	0.110	(11)	5.0	0.170	-0.264	0.125
H ₁₅	(C ₄)	0.227	(8)	-0.107	(6)	0.357	(11)	4.5	0.228	-0.104	0.352
H ₁₆		0.438	(9)	-0.114	(6)	0.307	(11)	4.5	0.438	-0.112	0.292
H ₁₇	(C ₃)	0.140	(8)	-0.091	(6)	-0.096	(10)	4.0	0.140	-0.090	-0.088
H ₁₈		0.367	(9)	-0.115	(5)	-0.109	(11)	4.0	0.366	-0.116	-0.106
H ₁₉	(C ₂)	0.420	(8)	0.041	(5)	0.058	(10)	3.0	0.424	0.039	0.061

all interatomic cross terms, though for an individual atom all cross terms between the temperature and scale parameters, as well as those between the x and z positional parameters, are collected.

The initial refinement was completed with four least-squares cycles during which the R factor fell to 0.078 and the maximum corrections in the parameters became approximately equal to their standard deviations. Positional parameters for all (19) hydrogen atoms were now calculated. Tetrahedral configurations were assumed for the carbon and nitrogen atoms, with C-H and N-H distances of 0.95 and 0.90 Å respectively. The orientations of the $-\text{NH}_3^+$ groups were indicated by the hydrogen-bond network. The hydrogen atoms of the water molecules were also positioned by considering the hydrogen bonding, and by assuming O-H distances of 0.90 Å and H-O-H angles of 104.5°. Once calculated, each hydrogen was identified as a peak in the latest difference map. Each hydrogen atom was assigned an isotropic temperature factor with a value of B approximately one unit higher than the average value for the atom to which it was bonded.

Eight more cycles of least-squares calculations completed the refinement. The ten parameters for each of the heavier atoms and three positional parameters of each hydrogen atom were optimized. A shift factor of 0.5 was applied to all parameter changes. After two cycles the R factor had fallen to 0.062, showing the discrepancies to be little larger than the observational errors, and accordingly the weighting scheme was changed to reflect the expected uncertainties in F_o^2 —that is, \sqrt{w} was changed from

$1/\sigma(F)$ to $1/\sigma(F^2)$. The calculated standard deviations in the atomic parameters decreased after this change, particularly for the chlorine atom, and small but significant corrections to the parameters followed. The final R factor was 0.057, again including only those reflections of non-zero weight.

The final positional and temperature-factor parameters for the heavier atoms are given in Table 2 and for the hydrogen atoms in Table 3. The observed and final calculated structure factors are given in Table 4.

Finally, two three-dimensional Fourier summations were calculated: a difference map, to confirm the hydrogen positions, and an electron density map. For the difference map, the amplitudes, ΔF , were taken as the differences between the absolute values of the observed structure factors and those calculated from the heavier atoms alone; the phase angles, however, included the calculated contributions of the hydrogen atoms. A composite drawing of this synthesis is given in Fig. 1, while for comparison the final three-dimensional electron density map is shown in Fig. 2. Every hydrogen atom peak in the difference map is greater than 0.35 e.Å⁻³, and with the exception of one peak of 0.3 e.Å⁻³ no other positive region exceeds 0.2 e.Å⁻³. There are some negative regions (as large as -0.5 e.Å⁻³), but they are usually associated with the chlorine atom. The two oxygen atoms of the water molecules are in regions of 0.0 and -0.25 e.Å⁻³ respectively, giving no conclusive evidence of a defect structure.

Parameters for the hydrogen atoms were obtained from the difference map using a 27-point Gaussian

Table 4. Observed and calculated structure factors

The five columns within each group contain the values of k, 10F_o, 10|F_c|, 10A_o and 10B_o. Reflections indicated by an asterisk (*) were given zero weight in the least-squares calculations. The values listed for 10F_o, 10A_o, and 10B_o for the 117 reflection are in error; the correct values are 36, 27, and -24

Table with multiple columns of numerical data, organized into groups labeled with reflection indices (e.g., 0k0, 0k1, 0k2, 0k3, 0k4, 0k5, 1k6, 1k7, 1k8, 1k9, 2k0, 2k1, 2k2, 2k3, 2k4, 2k5, 2k6, 2k7, 2k8, 2k9, 3k0, 3k1, 3k2, 3k3, 3k4, 3k5, 3k6, 3k7, 3k8, 3k9, 4k0, 4k1, 4k2, 4k3, 4k4, 4k5, 4k6, 4k7, 4k8, 4k9, 5k0, 5k1, 5k2, 5k3, 5k4, 5k5, 5k6, 5k7, 5k8, 5k9). Each group contains five columns of values.

from the parameters given by the final least-squares calculations. Of 57 parameters, 7 differ by more than one standard deviation (as obtained from the least-squares residuals) from the least-squares values, the maximum difference being 1.6σ . The bond lengths and angles listed in Table 6 are derived from the least-squares coordinates.

During the final eight refinement cycles in which the hydrogen atom contributions were included, the average shift in the parameters of the carbon, nitrogen, and oxygen atoms was 0.019 \AA , the maximum shift being 0.042 \AA for the water oxygen W_1 . The chlorine atom moved by 0.004 \AA . Since the heavy atom parameters seemed to have converged before the hydrogen atoms were introduced, the effect of the hydrogen atoms was to alter the apparent positions of the heavier atoms by several standard deviations. The hydrogen atoms moved an average of 0.18 \AA from their original calculated positions, the maximum shift being 0.34 \AA for H_6 .

The Patterson analysis had placed all the carbon, nitrogen, and oxygen atoms quite accurately. During the entire refinement their average shift was 0.15 \AA , with the maximum shift, again for W_1 , being 0.39 \AA .

Accuracy of the results

It seems that the structure has been refined to the limit of accuracy feasible with the present experimental observations. The final R factor for 1261 observed reflections with non-zero weight stands at 0.057 , close to the value predicted by comparing the values of intensities measured about both a and c axial settings. Moreover, the standard deviation of an observation of unit weight (Peterson & Levy, 1957) is calculated to be 1.16 , close to the ideal value of 1.00 . This agreement also implies that the weighting scheme adopted during the final stages of refinement reflects the true observational errors.

The standard deviations of the parameters given in Tables 2 and 3 were calculated from the sum of the weighted residuals and the diagonal terms of the inverse matrices of the least-squares normal equations. The limits of error (chosen as three times the standard deviations) are 0.015 – 0.018 \AA in the bond lengths and 33 – $40'$ in the bond angles between the heavier atoms. The internal consistency of bond lengths and angles expected to be equivalent shows these uncertainties to be realistic. The estimated limits of error for the hydrogen parameters are 0.2 – 0.3 \AA .

The values for the uncertainties in the temperature parameters are more difficult to assess. The standard deviations in their numerical values, calculated as above from the least-squares coefficients (see Table 2), give no clue as to the importance of such effects as extinction, absorption, scaling, and spot shape. However, all these effects have much the same dependence upon the scattering angle as an isotropic vibration, and a single isotropic temperature parameter should

be sufficient to allow for these systematic experimental errors, particularly as the crystals were regular in shape and without pronounced cleavage. The uncertainties reported for the temperature parameters, then, can be extended to apply to the relative values of the vibrational amplitudes, and there seems no reason to doubt the pattern of anisotropic thermal motions discussed below.

Discussion of the results

(i) Geometry of the molecule

The lysine molecule is dominated by two planar configurations, the carboxylate grouping $O_1O_2C_1C_2$ and the aliphatic side chain terminating in the ϵ nitrogen atom N_2 . As expected, the molecule is a zwitterion, with both nitrogen atoms accepting a proton (from the two acid groupings) and assuming the tetrahedral configuration $C-NH_3^+$. The bond lengths and angles found for the heavier atoms are given in Table 5 and shown in Fig. 3, while those involving the hydrogen atoms are given in Table 6.

Table 5. Bond lengths and angles

Bond		Angle	
C_1-O_1	1.250 \AA	$O_1-C_1-O_2$	125.5°
C_1-O_2	1.246	$C_2-C_1-O_1$	116.8
		$C_2-C_1-O_2$	117.7
C_2-N_1	1.484	$C_1-C_2-N_1$	109.7
C_6-N_2	1.480	$C_3-C_2-N_1$	111.8
		$C_5-C_6-N_2$	110.9
C_1-C_2	1.529	$C_1-C_2-C_3$	109.8
C_2-C_3	1.524	$C_2-C_3-C_6$	114.6
C_3-C_4	1.518	$C_3-C_4-C_5$	111.0
C_4-C_5	1.526	$C_4-C_5-C_6$	111.5
C_5-C_6	1.521		

The dimensions of the amino acid group are quite similar to those reported for glycine (Marsh, 1958), and are in general agreement with the values reported for other amino acids (summarized recently in a paper by Hahn, 1957). The two C–O bond lengths are equal within experimental error, with their average value of $1.248 \pm 0.004 \text{ \AA}$ (the uncertainty is the e.s.d. assuming the individual values to be equivalent) in good agreement with the average value of $1.254 \pm 0.003 \text{ \AA}$ for glycine.* The two C–NH₃⁺ bond lengths are also equal within experimental error, and again the average value of $1.482 \pm 0.004 \text{ \AA}$ is close to the (single) value found in glycine, $1.474 \pm 0.003 \text{ \AA}$. Somewhat higher values for this bond length have been reported else-

* The magnitudes of the principal thermal vibrations of atoms O_1 and O_2 (Table 8) indicate that a correction of approximately 0.012 \AA should be applied to each of the C–O bonds (Cruickshank, 1956). The resulting value is 1.260 \AA , close to the similarly corrected value 1.263 \AA in glycine and in agreement with the value deduced by Hahn for 50% double-bond character.

Table 6. Bond distances and angles involving the hydrogen atoms, calculated from the least-squares parameters of Table 3

	$d(X-H)$		$\angle \left(\begin{smallmatrix} C \\ N \end{smallmatrix} - X - H \right)$		$\angle (H-X-H)$
N_1-H_1	0.91 Å	$C_2-N_1-H_1$	113°	$H_1-N_1-H_2$	103°
N_1-H_2	1.01	$C_2-N_1-H_2$	108	$H_2-N_1-H_3$	124
N_1-H_3	0.88	$C_2-N_1-H_3$	103	$H_3-N_1-H_1$	106
N_2-H_4	1.04	$C_6-N_2-H_4$	112	$H_4-N_2-H_5$	120
N_2-H_5	0.82	$C_6-N_2-H_5$	107	$H_5-N_2-H_6$	97
N_2-H_6	0.97	$C_6-N_2-H_6$	108	$H_6-N_2-H_4$	111
W_1-H_7	0.89			$H_7-W_1-H_8$	103
W_1-H_8	1.00				
W_2-H_9	1.01			$H_9-W_2-H_{10}$	102
W_2-H_{10}	1.02				
C_5-H_{11}	1.04	$C_5-C_6-H_{11}$	114	$H_{11}-C_6-H_{12}$	98
C_6-H_{12}	1.03	$C_5-C_6-H_{12}$	112		
		$N_2-C_6-H_{11}$	106		
		$N_2-C_6-H_{12}$	115		
C_5-H_{13}	1.18	$C_4-C_5-H_{13}$	110	$H_{13}-C_5-H_{14}$	108
C_5-H_{14}	1.07	$C_4-C_5-H_{14}$	116		
		$C_6-C_5-H_{13}$	105		
		$C_6-C_5-H_{14}$	105		
C_4-H_{15}	1.11	$C_3-C_4-H_{15}$	117	$H_{15}-C_4-H_{16}$	101
C_4-H_{16}	1.02	$C_3-C_4-H_{16}$	113		
		$C_5-C_4-H_{15}$	105		
		$C_5-C_4-H_{16}$	108		
C_3-H_{17}	1.07	$C_2-C_3-H_{17}$	106	$H_{17}-C_2-H_{18}$	108
C_3-H_{18}	1.09	$C_2-C_3-H_{18}$	109		
		$C_4-C_3-H_{17}$	116		
		$C_4-C_3-H_{18}$	104		
C_2-H_{19}	0.91	$C_1-C_2-H_{19}$	118		
		$C_3-C_2-H_{19}$	104		
		$N_1-C_2-H_{19}$	104		

where, leading Hahn to deduce an average value 1.503 Å. This is possibly a little high.

The equation for the least-squares plane through atoms $O_1O_2C_1C_2$ is

$$0.4702X' + 0.8673Y + 0.2094Z' = 1.3732,$$

where the direction cosines refer to the orthogonal axes a^*bc . The deviations of the individual atoms from the plane are:

$$O_1 - 0.004, \quad O_2 - 0.004, \quad C_1 0.010, \quad C_2 - 0.003 \text{ Å}.$$

The nitrogen atom N_1 lies 0.446 Å out of the plane; the value found in glycine is 0.436 Å.

An interesting result of this investigation is the determination of accurate bond lengths and angles in a medium length (5 atom) aliphatic carbon chain. As Bonham & Bartell (1959) point out, there is no reliable value for the normal carbon-carbon single-bond length, for the very good reason that there is a dearth of accurate structural information. Bartell (1959) further argues that the standard most commonly accepted, the bond length in diamond, is unsuitable for this purpose because of interactions peculiar to the diamond structure. He proposes,

instead, the average value 1.533 ± 0.003 Å found in electron diffraction studies of the normal hydrocarbons butane through heptane. The average value of the C-C single bonds in lysine is $1.524 + 0.003$ Å, significantly shorter than the diamond value and slightly shorter than the value proposed by Bartell. The average value of the four C-C-C angles is 111.7° , significantly greater than the tetrahedral value but close to the value of 112.4° found by Bonham & Bartell in *n*-butane. However, the angle $C_2-C_3-C_4$, 114.6° , is appreciably larger than the other three. Similar results are found in other amino acids and peptides, and in the case of leucylprolylglycine (Leung & Marsh, 1958) it was suggested that the widening of this angle might be due to the steric effect of an atom hydrogen bonded to the NH_3^+ group. In lysine an oxygen atom O_1 , which is hydrogen bonded to N_1 , is only 3.37 Å from C_4 and 2.77 Å from H_{15} .

The aliphatic side chain has a fully extended configuration, the five carbon atoms and the terminal nitrogen being approximately coplanar. The least-squares plane of these atoms is

$$-0.9729X' + 0.119Y + 0.2024Z' = 2.1535,$$

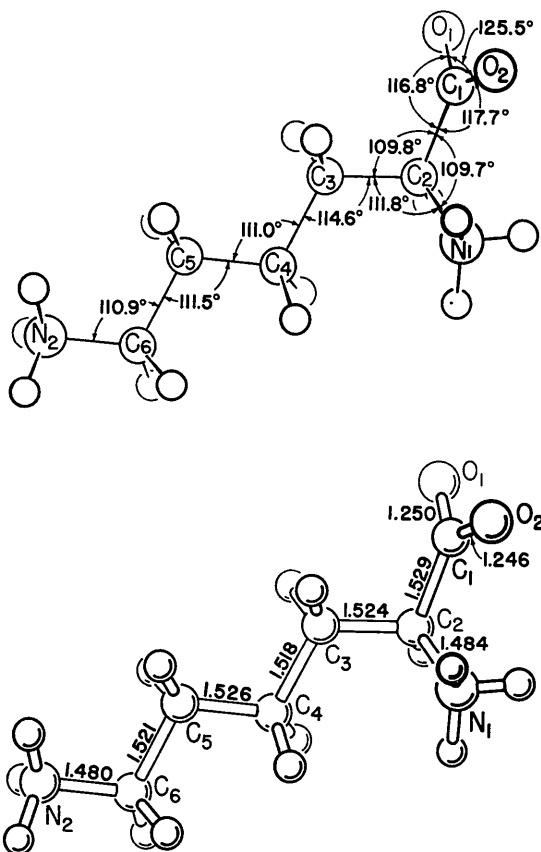


Fig. 3. Bond distances and angles.

and the small, though significant, deviations from it are,

$$\begin{array}{lll} C_2 & -0.003, & C_3 & -0.058, & C_4 & 0.040 \text{ \AA}, \\ C_5 & 0.081, & C_6 & -0.039, & N_2 & -0.022 \text{ \AA}. \end{array}$$

The dihedral angle between this plane and that of the carboxyl group is 71.4° .

The average C-H bond length is $1.06 \pm 0.04 \text{ \AA}$, quite close to the value of 1.09 \AA usually taken as the standard internuclear separation. Thus this investigation, unlike some other recent investigations carried out in these and other Laboratories, has given little support to the theories of an apparent shortening in the C-H bond length when it is measured by X-ray methods. In addition, the average O-H bond length, 0.98 \AA , is close to the internuclear separation, 1.01 \AA , found by Peterson & Levy (1957) in a neutron diffraction study of heavy ice. However, the average N-H bond length of 0.94 \AA is considerably shorter than the expected internuclear separation, though it is longer than the value of 0.87 \AA found in glycine. Thus it is still not clear just where the effective X-ray scattering center of a bonded hydrogen atom lies.

(ii) Packing of the molecules

Drawings of the structure viewed along [001] and [100] are shown in Figs. 4 and 5, respectively. The molecules of lysine are in an extended configuration parallel to the b axis, and the packing is dominated by a three-dimensional network of hydrogen bonds. All ten protons of the nitrogen atoms and the water molecules participate in this network, though one of the protons of nitrogen atom N_2 forms a bifurcated bond.

The hydrogen bond lengths and angles are given in Table 7. The nitrogen N_1 of the amino acid grouping forms three such bonds—to a chloride ion, a water oxygen, and an oxygen of the carboxylate group. The terminal (ϵ) nitrogen N_2 is at hydrogen bonding distances from four atoms—two chloride ions and two oxygens of different carboxylate groups. It forms two normal hydrogen bonds, one to a chloride ion and

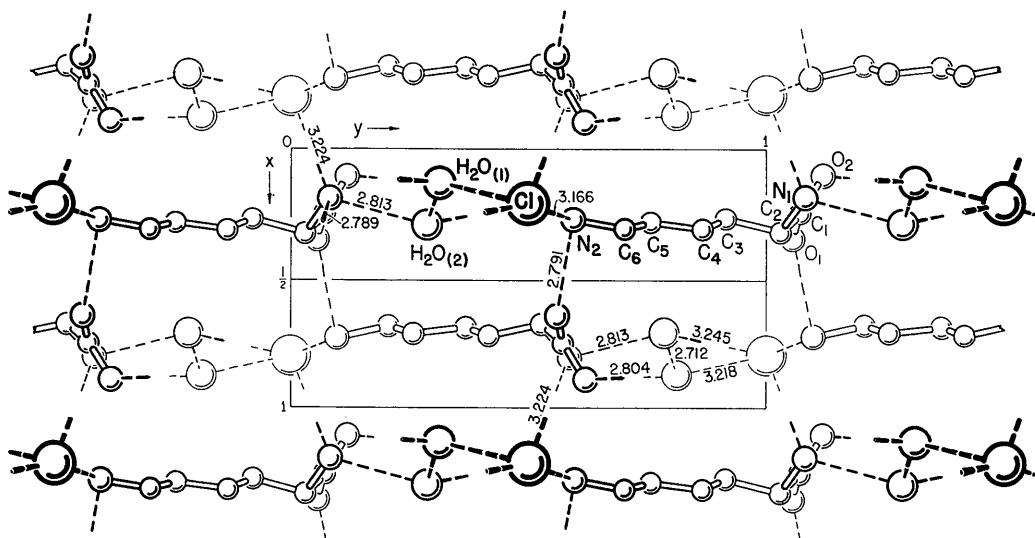


Fig. 4. A drawing of the structure viewed along [001].

Table 7. *Hydrogen-bond lengths and angles*

The positional parameters of donor atoms (D) are those of Table 2; those of acceptor atoms (A) are derived from the table values by the relations given

Donor	Acceptor, at	$d(D \cdots A)$	$(H-D \cdots A)$	$(D-H \cdots A)$	$(C-N \cdots A)$	$(A \cdots D \cdots A)$
W_1	$O_2, (x, y, z+1)$	2.804 Å		$W_1-H_8 \cdots O_2$ 170°		$O_2 \cdots W_1 \cdots Cl$ 118.9°
W_1	$Cl, (x, y, z+1)$	3.218		$W_1-H_7 \cdots Cl$ 156		
W_2	$Cl, (x, y, z)$	3.245		$W_2-H_9 \cdots Cl$ 164		
W_2	$W_1, (x, y, z)$	2.712		$W_2-H_{10} \cdots W_1$ 160		$Cl \cdots W_2 \cdots W_1$ 101.9
N_1	$Cl, (\bar{x}, y-\frac{1}{2}, \bar{z})$	3.224	13°	$N_1-H_3 \cdots Cl$ 162	$C_2-N_1 \cdots Cl$ 111.1°	$Cl \cdots N_1 \cdots O_1$ 107.8
N_1	$O_1, (x, y, z+1)$	2.789	3	$N_1-H_2 \cdots O_1$ 175	$C_2-N_1 \cdots O_1$ 111.4	$O_1 \cdots N_1 \cdots W_2$ 92.5
N_1	$W_2, (x, y, z)$	2.813	16	$N_1-H_1 \cdots W_2$ 157	$C_2-N_1 \cdots W_2$ 106.9	$W_2 \cdots N_1 \cdots Cl$ 125.4
N_2	$Cl, (x, y-1, z+1)$	3.166	3	$H_6-N_2 \cdots Cl$ 175	$C_6-N_2 \cdots Cl$ 110.9	$O_1 \cdots N_2 \cdots Cl$ 104.9
N_2	$O_1, (\bar{x}+1, y-\frac{1}{2}, \bar{z})$	2.791	13	$N_2-H_5 \cdots O_1$ 162	$C_6-N_2 \cdots O_1$ 94.5	$O_2 \cdots N_2 \cdots Cl$ 89.3
N_2	$Cl^*, (x, y-1, z)$	3.346	41	$H_4-N_2 \cdots Cl^*$ 124	$C_6-N_2 \cdots Cl^*$ 118.0	$O_2 \cdots N_2 \cdots O_1$ 171.8
N_2	$O_2, (\bar{x}, y-\frac{1}{2}, \bar{z})$	2.889	37	$H_4-N_2 \cdots O_2$ 126	$C_6-N_2 \cdots O_2$ 90.9	$Cl^* \cdots N_2 \cdots Cl$ 129.0
						$Cl^* \cdots N_2 \cdots O_1$ 85.2

one to a carboxylate oxygen, which orient the third proton in a direction approximately half way between the other two atoms, though at distances appreciably further than normal from both. The $H_4 \cdots Cl^*$ and $H_4 \cdots O_2$ distances are 2.64 and 2.14 Å, respectively, while the four other $H \cdots Cl$ distances in the structure range from 2.19 to 2.38 Å and the five other $H \cdots O$ distances from 1.73 to 2.00 Å. This configuration is usually called a 'bifurcated hydrogen bond', and though objection may be made to the use of the word 'bond' some added stabilization should result from the interaction between the hydrogen atom and its negatively charged neighbors. No attempt was made to judge if this stabilization is appreciably greater than that due to normal van der Waals interactions.

The water molecules are an integral part of the hydrogen bond system, each being the donor of two

Table 8. *Magnitudes and direction cosines of the principal axes of the vibration ellipsoids, referred to orthogonal axes a^*bc*

Atom	Axis i	B_i	g_i^1	g_i^2	g_i^3
O_1	1	4.601	0.139	0.986	0.094
	2	3.842	0.985	-0.147	0.086
	3	2.340	-0.099	-0.080	0.992
O_2	1	5.180	0.287	0.870	0.402
	2	3.710	0.715	0.085	-0.694
	3	2.567	-0.637	0.486	-0.598
C_1	1	3.571	-0.488	0.866	0.110
	2	2.664	0.867	0.466	0.174
	3	2.221	-0.100	-0.180	0.979
C_2	1	2.762	0.388	0.919	-0.068
	2	2.486	-0.422	0.243	0.873
	3	2.421	0.819	-0.310	0.482
N_1	1	3.462	0.829	0.551	0.091
	2	2.920	-0.504	0.809	-0.301
	3	2.530	-0.240	0.204	0.949
C_3	1	4.054	0.995	0.069	-0.068
	2	3.398	0.066	0.024	0.998
	3	2.321	-0.070	0.997	-0.019
C_4	1	4.153	0.974	0.098	-0.204
	2	3.406	0.110	0.583	0.805
	3	2.744	-0.198	0.807	-0.557
C_5	1	4.819	0.998	-0.062	0.003
	2	3.651	0.023	0.412	0.911
	3	2.792	0.058	0.909	-0.413
C_6	1	4.086	0.976	-0.068	-0.209
	2	3.356	0.114	0.970	0.215
	3	2.705	0.188	-0.234	0.954
N_2	1	4.057	0.806	-0.452	-0.382
	2	3.468	0.585	0.508	0.632
	3	2.522	0.092	0.733	-0.674
W_1	1	13.089	-0.988	0.135	-0.074
	2	5.510	0.018	0.577	0.817
	3	4.248	0.153	0.806	-0.572
W_2	1	10.252	-1.000	-0.018	0.010
	2	5.234	0.001	0.416	-0.909
	3	3.379	-0.021	0.909	-0.416
Cl	1	4.260	0.362	0.820	0.443
	2	3.541	-0.886	0.451	-0.111
	3	2.941	0.291	0.352	-0.889

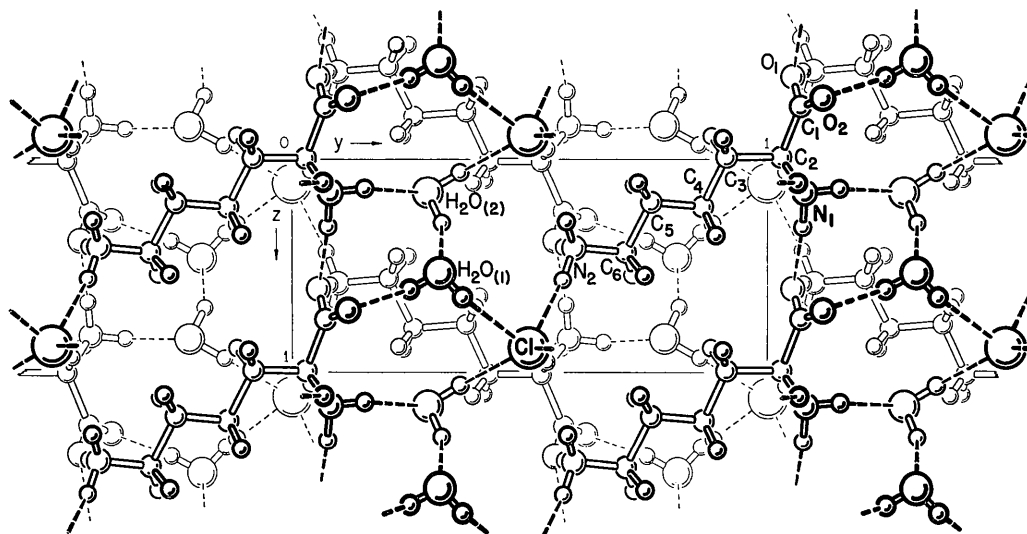


Fig. 5. A drawing of the structure viewed along [100].

bonds and the acceptor for one. Any substantial loss of water would presumably lead to the destruction of the crystal, though the density measurements seem to indicate that a slight deficiency (5%) can be tolerated. By way of contrast some recent investigations (Leung & Marsh, 1958; Sutor, 1958; Gerdil, 1961; Hughes, Yakel & Freeman, 1961) have shown structures stable with only 80% or less of the water sites occupied. In these cases the role of the water molecules appears to be principally one of space filling.

(iii) *The temperature factors*

The magnitudes and direction cosines of the principal axes of the ellipsoids of thermal motion, derived from the anisotropic temperature factor coefficients of Table 2, are given in Table 8.

By far the greatest thermal motion is associated with the two water molecules, as would be expected, for they are held in place solely by hydrogen bonds. In each case the direction of maximum vibration is within 12° of normal to the plane defined by the three neighboring hydrogen bonded atom, and, as it happens, nearly parallel to the a axis. The smallest vibration is associated with the central carbon atom C_2 , and is nearly isotropic. The major axes of the carboxylate oxygen atoms O_1 and O_2 are approximately perpendicular to the plane of that group, and the major axes of the carbon atoms of the side chain (C_3 to C_6) are approximately perpendicular to its plane.

The temperature factors cannot be interpreted in terms of rigid body vibrations. The greatest vibration of a side chain atom is associated not with the terminal atom N_2 but with the penultimate carbon atom C_5 . Also, the direction of maximum vibration of N_2 differs by more than 30° from that of the other side chain atoms. Apparently the hydrogen bonds formed by N_2

inhibit the thermal vibrations associated both with it and its immediate neighbors.

We should like to thank Mr David Barker who measured the density and assisted in the final calculations. D. A. Wright acknowledges gratefully a Fulbright grant for travel between Auckland, New Zealand, and Pasadena.

References

- ALBRECHT, G. & COREY, R. B. (1959). *J. Amer. Chem. Soc.* **61**, 1087.
 BARTELL, L. S. (1959). *J. Amer. Chem. Soc.* **81**, 3497.
 BERGHUIS, J., HAANAPPEL, IJ. M., POTTERS, M., LOOPSTRA, B. O., MACGILLAVRY, C. H. & VEENENDAAL, A. L. (1955). *Acta Cryst.* **8**, 478.
 BONHAM, R. A. & BARTELL, L. S. (1959). *J. Amer. Chem. Soc.* **81**, 3491.
 CRUICKSHANK, D. W. J. (1956). *Acta Cryst.* **9**, 757.
 GERDIL, R. (1961). *Acta Cryst.* **14**, 333.
 HAHN, T. (1957). *Z. Kristallogr.* **109**, 438.
 HAMILTON, W. C. (1957). *Acta Cryst.* **10**, 629.
 HUGHES, E. W., YAKEL, H. L. & FREEMAN, H. C. (1961). *Acta Cryst.* **14**, 345.
 IBERS, J. A. (1956). *Acta Cryst.* **9**, 652.
 LEUNG, Y., MARSH, R. E. & SCHOMAKER, V. (1957). *Acta Cryst.* **10**, 650.
 LEUNG, Y. & MARSH, R. E. (1958). *Acta Cryst.* **11**, 17.
 MARSH, R. E. (1958). *Acta Cryst.* **11**, 654.
 MCWEENY, R. (1951). *Acta Cryst.* **4**, 513.
 PASTERNAK, R. A. (1956). *Acta Cryst.* **9**, 341.
 PETERSON, S. W. & LEVY, H. A. (1957). *Acta Cryst.* **10**, 70.
 RAMAN, S. (1959). *Z. Kristallogr.* **111**, 301.
 SHOEMAKER, D. P., DONOHUE, J., SCHOMAKER, V. & COREY, R. B. (1950). *J. Amer. Chem. Soc.* **72**, 2328.
 SRINIVASAN, R. (1956). *Acta Cryst.* **9**, 1039.
 SUTOR, D. T. (1958). *Acta Cryst.* **11**, 453.