Coulson, C. A., Higgs, P. W. & March, N. H. (1951). *Nature, Lond.* **168**, 1039.
James, R. W. & Brindley, G. W. (1931). *Phil. Mag.* **12**, 81.
Klement, O. (1951). *Helv. Chim. Acta*, **34**, 2230.
March, N. H. (1952). *Acta Cryst.* **5**, 187.
Mathieson, A. McL., Robertson, J. M. & Sinclair, V. C. (1950). *Acta Cryst.* **3**, 251.
McWeeny, R. (1951). *Acta Cryst.* **4**, 513.
McWeeny, R. (1952). *Acta Cryst.* **5**, 463.
Slater, J. C. (1930). *Phys. Rev.* **36**, 57.
Torrance, C. (1934). *Phys. Rev.* **45**, 388.

Acta Cryst. (1953). **6**, 241

The Crystal Structure of DL-Serine*

By David P. Shoemaker†, Robert E. Barieau‡, Jerry Donohue and Chia-Si Lu§

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

(*Received* 28 *July* 1952)

The crystal structure of DL-serine has been determined by the use of complete data accessible with Cu $K\alpha$ radiation. A satisfactory trial structure was quickly obtained from a three-dimensional Patterson function by the application of a new method for interpreting these functions. This method, called the *Patterson superposition method*, is related to several other methods recently proposed, such as the vector convergence method of Beevers & Robertson, but differs from them in procedure and criteria employed.

The atomic parameters were refined by three-dimensional Fourier and least-squares procedures. Interatomic distances and interbond angles are close to those expected, and there is a satisfying system of hydrogen bonding throughout the crystal.

Introduction

Within the past decade and a half, determinations have been made in these Laboratories of the crystal structures of a number of amino acids and related substances by X-ray diffraction methods. The crystal structures of diketopiperazine (Corey, 1938), glycine (Albrecht & Corey, 1939), DL-alanine (Levy & Corey 1941), L_s-threonine (Shoemaker, Donohue, Schomaker & Corey, 1950), acetylglycine (Carpenter & Donohue, 1950), β-glycylglycine (Hughes & Moore, 1949) and hydroxy-L-proline (Donohue & Trueblood, 1952) have been determined, and the positional parameters for DL-alanine (Donohue, 1950) have been further refined. The bond distances and bond angles, modes of hydrogen bonding, and steric relationships of the molecules of these substances have been found very useful in the formation of conceptions and definite conclusions regarding the configurations of polypeptide chains in proteins (Corey & Donohue, 1950; Corey, 1940, 1948). These past investigations, together with the present one and others now in progress, are a part of a current program of research on the structure of protein molecules at this Institute. Because of its importance to the amino acid and protein program, DL-serine (α-amino-β-hydroxypropionic acid) received attention in these Laboratories first in 1942, when its unit cell and space group were determined (Lu, unpublished work). Continuation of that work was unsuccessful in revealing the structure, which did not appear to be amenable to treatment by the two-dimensional methods which were used.

In the investigation of the crystal structure of threonine (Shoemaker *et al.*, 1950) the use of the three-dimensional Patterson function enabled the determination of the approximate structure to be made in the face of a complexity that had defied two-dimensional methods and trial-and-error procedures. Consequently the decision was made to resume work on DL-serine with the intention of applying the three-dimensional Patterson method. This method led quickly to the approximate structure and was, moreover, also used to advantage in carrying out a preliminary refinement of parameters, thereby obviating altogether the use of Fourier projections and trial-and-error procedures for this purpose, since parameters of sufficient precision to justify immediate use of three-dimensional least-squares and Fourier refinement methods were obtained. These methods quickly converged to give the final atomic positional parameters.

Experimental

The DL-serine used in this work was prepared by Dr Roland N. Icke, and was made available to us by Prof. Carl Niemann of this Institute. Crystals for use

---

* Presented at the Second International Congress of Crystallography, Stockholm, June 1951. Contribution No. 1713 from Gates and Crellin Laboratories.

† Present address: Department of Chemistry, Massachusetts Institute of Technology, Cambridge 39, Massachusetts, U.S.A.

‡ Participation sponsored by California Research Corporation. Present address: California Research Corporation, Richmond, California, U.S.A.

§ Present address: University of Amoy, Amoy, China.

in intensity work, which were grown from water solution, were reduced to cylinders about 0·4 mm. in diameter by means of a fine camels-hair brush moistened with water. The crystals were found on goniometric examination to be monoclinic holohedral with

$$a_0:b_0:c_0 = 1{\cdot}176:1:0{\cdot}528 \text{ and } \beta = 106°\,31'\,.$$

The Laue symmetry was found to be C_{2h}. Rotation photographs were taken around each of the three principal axes with Cu $K\alpha$ radiation ($\lambda = 1{\cdot}5418$ Å); measurements of the layer line spacings gave approximate lattice constants which were used to index equatorial reflections on oscillation photographs. Measurements of these reflections gave a unit cell with the following values for the lattice constants:

$$a_0 = 10{\cdot}72 \pm 0{\cdot}02,\ b_0 = 9{\cdot}14 \pm 0{\cdot}02,\ c_0 = 4{\cdot}825 \pm 0{\cdot}006 \text{ Å};$$
$$\beta = 106°\,27' \pm 4'\,.$$

The quoted uncertainties are standard deviations. There are four molecules of serine ($HO-CH_2-CH(NH_3^+)-CO_2^-$) in this unit cell (3·99 calculated from the density of 1·537 g.cm.$^{-3}$ reported by Albrecht, Schnakenberg, Dunn & McCullough (1943)). The calculated axial ratios, 1·173:1:0·528, are in good agreement with the values obtained optically. The lattice parameters found by Albrecht *et al.* are 10·74, 9·15, 4·78 Å, and 106° 13′; their axial ratios are 1·174:1: 0·5224 (X-ray) and 1·194:1:0·522 (optical). Weissenberg photographs taken with nickel-filtered Cu $K\alpha$ radiation demonstrated the presence of ($0k0$) reflections only with k even and of ($h0l$) reflections only with h even. These observations lead uniquely to the centrosymmetric space group C_{2h}^5–$P2_1/a$, in agreement with the result of Albrecht *et al.*

Three-dimensional intensity data were obtained from equi-inclination Weissenberg photographs taken with nickel-filtered Cu $K\alpha$ radiation for all layer lines with equi-inclination angles less than 30°, with each of two crystals rotated respectively around their b and c axes. The estimation of intensities was facilitated by means of intensity strips and the multiple-film technique. The values of $|F_{hkl}|^2$, obtained from the intensities in the usual way, were placed on an absolute scale by the method of Wilson (1942). The coefficient B in the temperature factor, $\exp(-B\sin^2\theta/\lambda^2)$, was found to have the value 1·7 Å^2. With the use of this temperature factor, atom form factor curves (*Internationale Tabellen*, 1935) which include the temperature factor were prepared for use in the computation of calculated structure factors and of their derivatives later on.

Determination of the trial structure

In preparation for the calculation of the three-dimensional Patterson function the values of $|F_{hkl}|^2$ were sharpened (Patterson, 1935) and multiplied by a modification function (Waser, 1944; Schomaker, 1947; Shoemaker, 1947) of the form $(2\sin\theta/\lambda)^4 \exp\{-\alpha^2(2\sin\theta/\lambda)^2\}$ with $\alpha = 2{\cdot}2$ Å. This function has a zero value at $\theta = 0$ and a very small value at the limit of the data, and has a maximum value at about half the maximum value of $\sin\theta$. Its Fourier transform, which gives the peak shape to be expected for a resolved interaction, has a half-width at half-maximum of about 0·4 Å, and possesses a shallow minimum (of depth approximately 10% of the peak height) about 0·9 Å from the peak center. There is a very slight maximum (about 1% of the peak height) at about 1·6 Å, and no other significant maxima or minima. This peak shape is better, from the standpoint of a suitable compromise between resolving power and relative freedom from effects due to non-convergence, than the peak shape that would have been obtained either from the use of the unmodified F^2 or from the sharpened F^2/f^2 values.

The punched-card calculations for the three-dimensional Patterson function, as also for the three-dimensional Fourier syntheses later on, were carried out by a procedure developed by Prof. Verner Schomaker. All of these three-dimensional functions were calculated at intervals of $a_0/60$, $b_0/30$, and $c_0/15$. The corresponding spacings of the calculated points are

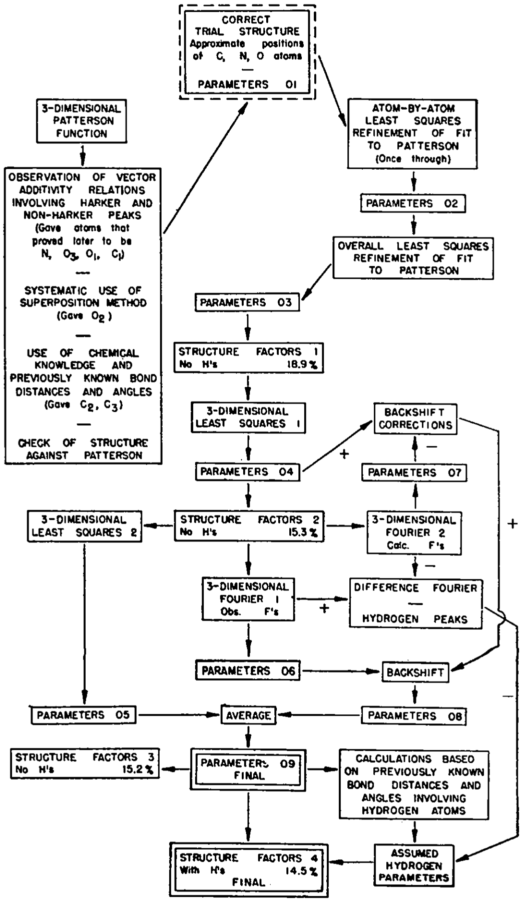


Fig. 1. Sequence of operations involved in the determination of the structure and the refinement of positional parameters.

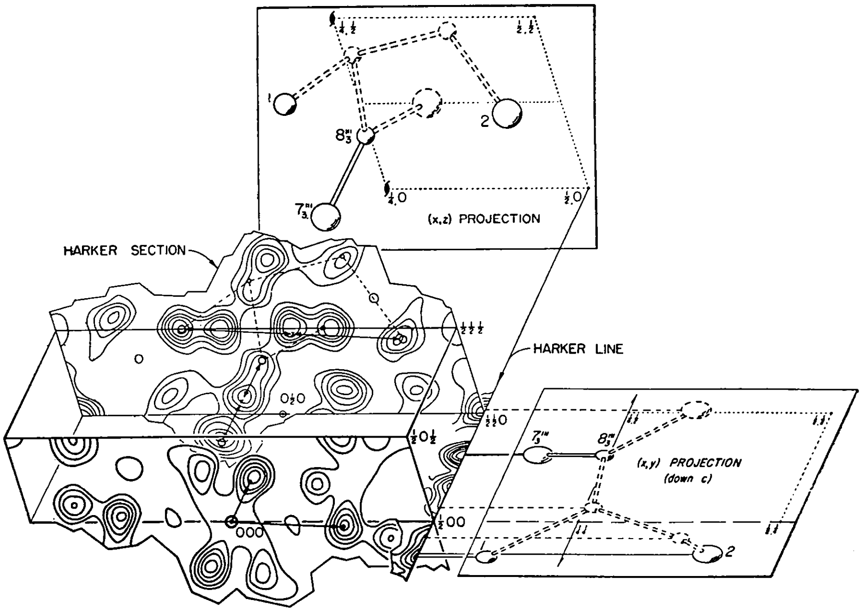


Fig. 2. Part of the three-dimensional Patterson function, showing how positions of atoms N, O_3, O_1 and C_1 were obtained.

0·18, 0·30 and 0·32 Å. The last two spacings are about as large as can be tolerated for precise work.

The Patterson function was, as expected, comparatively well resolved. The one-quarter cell within which the function was calculated contained about 50 resolved maxima (giving half weight to maxima on faces of the quarter-cell and quarter weight to those on edges). This is about half the expected number of interactions per quarter-cell, i.e. 98 excluding the origin peak, which was removed before the function was calculated (Patterson, 1935).

The procedure which was followed in obtaining the structure from the Patterson function is outlined briefly in Fig. 1, and is illustrated in Figs. 2–4.

No immediate attempt was made to construct the molecule out of interactions corresponding to covalently bonded distances, as this procedure did not prove fruitful in the case of threonine. A search was made for peaks at a distance of about 2·25 Å from the origin in the hope of determining the relative orientation of the two carboxyl oxygen atoms. Such a peak was found, but it did not appear to be completely resolved. The first progress toward an interpretation of the Patterson function was made by use of peaks on the Harker section $(u, \frac{1}{2}, w)$, the Harker line $(\frac{1}{2}, v, 0)$, and the zero section parallel to the Harker section $(u, 0, w)$, together with some other more general peaks.

The Harker section was not of much value by itself, because of the large number of non-Harker interactions on it. However, these non-Harker interactions are related to the general peaks on the zero section (Albrecht & Corey, 1939). Vector additivity relations were found involving one such peak on the zero section with three others on the Harker section, to give the x and z coordinates of two atoms, designated 1 and 2 (see Fig. 2). The common y coordinate of these two atoms was found by making trial fits to the Harker lines and looking for additional required general peaks. Indications of four more atoms (3, 4, 5 and 6) were found to be spurious through inability to fit required Patterson interactions to peaks. The positions of two more atoms (7 and 8) were found in the same manner as 1 and 2, but ambiguity remained as to which of four equivalent pairs of Harker peaks on the section and two equivalent Harker peaks on the Harker line correspond to the arbitrary choice of Harker peaks to represent atoms 1 and 2. Of the eight possible choices all but one, with atoms 7_3 and 8_3, were ruled out through inability to fit required Patterson interactions to peaks. The relationships between the four atoms 1, 2, $7_3''$ and $8_3''$ and the Harker and general Patterson peaks are shown in Fig. 2. The primes distinguish among symmetry-equivalent atoms.

Patterson superposition method

In order to find the positions of additional atoms, the above rather unsystematic trial-and-error procedure was abandoned in favor of a method suggested originally by Prof. Verner Schomaker in 1948. This method, which we call the 'Patterson superposition method', is substantially identical in principle with the method applied in two dimensions by Buerger (1951), but differs from it somewhat in procedure and in the criteria employed. It is also related to the 'vector convergence method' of Beevers & Robertson (1950), but is much more general; this last method, according to its authors, 'is not generally applicable, since it depends upon the presence of a heavy atom, and upon the multiplicity of the space group'.

The Patterson superposition method constitutes an operational procedure for carrying out a fundamental and straightforward interpretation of a Patterson function. In principle no *a priori* structural knowledge is needed except knowledge as to the numbers, kinds and 'shapes' of the atoms present (and perhaps a minimum limit for interatomic distances) and present indications are that this should usually prove true in practice. The most important condition for the applicability of the method is that at least a substantial number or fraction of the Patterson peaks in a unit cell must be resolved 'interaction peaks', or resolvable into them (by application of a theoretically predicted peak 'shape') to the extent of assigning coordinates to the interaction points. This condition is better fulfilled for small unit cells than for large ones, for typical organic structures than for structures of special types in which there is considerable systematic clustering of interactions, for centrosymmetric structures than for non-centrosymmetric ones, and for structures containing a few well-placed heavy atoms (when light atom–light atom interactions can be neglected) than for structures in which all atoms have comparable scattering powers, other things being constant in each instance. Especially is it better fulfilled with a three-dimensional Patterson function than with a projection; for moderately complicated organic structures without heavy atoms or very severe structural limitations the difference is likely to be decisive.

An obvious property of the Patterson function is that if the origin of the function is placed at any atom center, all atom centers must coincide with interaction points; moreover, if two (or more) identical Patterson functions are placed with their respective centers at any atom centers, all atom centers must coincide with interaction points of both (or all) Patterson functions. Let, therefore, the Patterson function or projection be represented by contour lines drawn on each of two (or more) identical maps, or in three dimensions two (or more) identical sets of maps, designated P and $P^*(\ldots P^N)$, prepared on transparent paper. Each map in a set represents one of a number of closely and equally spaced 'levels' in the Patterson function. The interaction points may be marked on these maps, having been located by application of a theoretical interaction peak 'shape' (Fourier transform of the modified square of the form factor) either by eye or by more quantitative means; or this may be done partially or entirely in conjunction with the interpretive steps. An initially blank transparent-paper map or set of maps S is also made ready. Now any observed resolved interaction point may be considered *by definition* as representing a vector from atom X_1 to atom X_2, the atoms being otherwise unspecified. Let now the atoms X_1 and X_2 be marked on S, in relative positions corresponding to the observed interaction, and let the origin of P be placed at X_1 and that of P^* at X_2 by superposing the maps (or in three dimensions interleaving the sets of maps) in the appropriate way. The map S, or in three dimensions one map of S at a time, is placed over an illuminated glass screen, with the required maps of P and P^* in correct position underneath it, and coincidences between interaction points on P and P^* are marked on S with a symbol of some kind (see Fig. 3). Coincidence or significant overlaps involving large peaks which cannot readily be resolved into component interactions are also marked, the lower significance of fit being indicated perhaps by making the symbol correspondingly smaller; here the judgments are somewhat intuitive, and involve consideration of tentative abstraction of possible interaction points from the unresolved peaks being considered. The attempt to match interaction points objectively abstracted from peaks in all possible cases, and to make somewhat subjective or intuitive matches otherwise, constitutes an additional distinction between the present method and others which make use of a sum or product function of P and P^*, or employ the minimum function of Buerger (1951). When all levels have been thus treated, S contains symbols representing possible positions of atoms relative to X_1 and X_2, half of which are due to a spurious center of symmetry half way between X_1 and X_2 which should disappear in the next step. There may also be other spurious symbols corresponding to accidental coincidences between unrelated interactions.

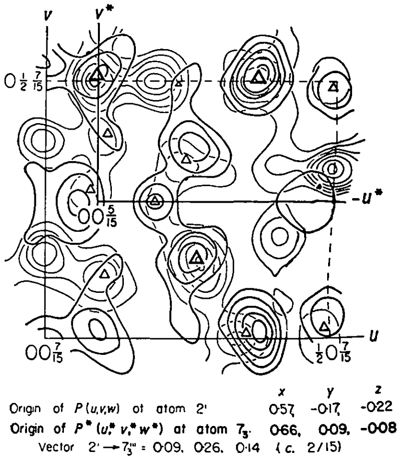


Fig. 3. Illustration of the Patterson superposition method of locating atoms with the three-dimensional Patterson function.

In order to carry out the next step the 'best' of the coincidences obtained in this way is labelled X_3 on S and the origin of P^* (say) is moved to it, the origin of P remaining at X_1. Again coincidences are marked, this time with a different symbol, and should approximately coincide with those obtained in the first step except where they are spurious. The process is

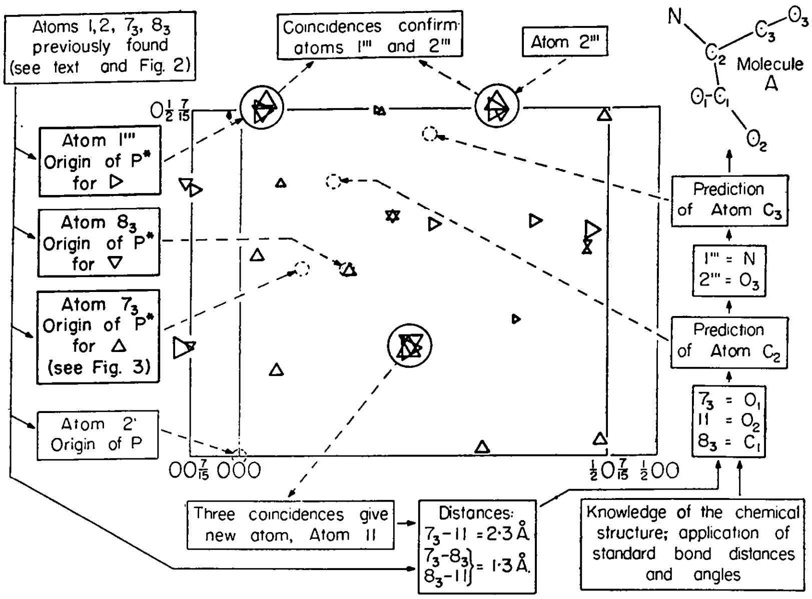


Fig. 4. Illustration of the Patterson superposition method of locating atoms with the three-dimensional Patterson function. Only superpositions shown are those on layer 7/15 of P. Origin of this figure is arbitrarily taken at that of P, namely at atom 2'.

continued in the same way, taking the origins of P and P^* at different pairs of established atoms, interpreting the best available multiple coincidences as additional established atoms X_n where necessary. Subject to practical limitations more than two sets of Patterson maps may be used at a time; the superposition of three or more Patterson functions does not in general have an unwanted center of symmetry. The operation is carried through a number of steps (perhaps half a dozen or so) until all of the atoms in the structure (except presumably hydrogens) are indicated by multiple coincidences in the form of clusters of superposition symbols on S, while any spurious symbols should be scattered more or less at random, or until enough atoms have been found to define the structure with the aid of any *a priori* structural information available. The atoms need yet to be identified as to kind, though in principle this could be done by quantitative consideration of amplitudes of interaction peaks. Symmetry relations among some or all of the established atoms will appear in due course (probably at an early stage) and permit the representation of symmetry elements on S and the assignment of *absolute* positions to the atoms; this may considerably facilitate the remaining steps. At any stage, after a few atoms have been found, a check on their validity may be obtained by calculating their mutual interaction point positions and testing the fit with the Patterson function, and the positions can often be improved by least squares as described later.

It is always possible that a spurious coincidence will be chosen as an X_n early in the procedure, with chaotic consequences which will soon dictate a fresh start. If, however, the number of resolved or readily resolvable interactions, though perhaps small, is sufficient to enable the successful consummation of the first three or four steps, it is highly probable that a continuation of the procedure will converge to give the structure even when the general conditions of resolutions are so poor that much reliance must be placed on intuitive matching.

In the application of the method to serine the positions of four atoms (and of their symmetrical equivalents) were already known, though there appears to be no reason why these positions could not have been found perhaps more easily by straightforward application of the method from the start. Three steps of the method were carried out, with the origin of P at atom 2' and that of P^* successively at atom 1''', 8_3 and 7_3, as shown in Figs. 3 and 4. These sufficed to give a good fifth atom, designated 11, as well as two atoms (9 and 10) which proved spurious on a check of calculated interaction positions against the Patterson function. The process was stopped at this point because application of *a priori* structural knowledge permitted the structure to be derived (see Fig. 4) and the remaining two carbon atoms, which had not been clearly indicated as multiple coincidences, to be located and checked against the Patterson function. Presumably, continuation through a few more steps would have indicated these atoms clearly.

From the trial structure (Parameters 01) the positions of all Patterson interactions were calculated, and these were plotted on the Patterson layers. The fit was highly satisfactory, and the positions and shapes of nearly all peaks were fairly well explained. Only four interactions, all of which involved carbon atoms, failed altogether to lie on a region where the Patterson function was positive. From the fit it could be concluded that the trial structure was correct beyond reasonable doubt in all significant details except for small errors in atomic position, of the order of 0·1 or 0·2 Å.

Refinement of positional parameters

The steps and procedures that were followed in refining the positional parameters are shown schematically in Fig. 1, and the parameters at several stages of the work are listed in Table 1. The percentage numbers given in the structure-factor boxes in Fig. 1 are percentage discrepancies (reliability factors), R.

Refinement of fit to Patterson function

In previous determinations in this Laboratory it has been the usual practice to precede three-dimensional Fourier and least-squares treatments by a preliminary refinement of the trial structure by calculating structure factors for zones and making parameter shifts with the aid of structure-factor maps and Fourier projections. In the present case the degree of resolution of the three-dimensional Patterson function was good enough to suggest the application of methods proposed in 1948 by Prof. Verner Schomaker and tested by himself and one of the present authors by application to the three-dimensional Patterson function for L-threonine. These methods amount to least-squares refinements of the fit of the structure to the Patterson function. They are applicable only if the Patterson function is sufficiently well resolved to indicate unambiguously the direction of movement of a postulated interaction toward an observed Patterson peak.

Two different modes of approach were then tried. In the first, all but one of the atoms are fixed with the best values of the parameters available at the time; the position of the remaining atom is varied in such a way as to best improve the fit of the interactions involving that atom to the peaks on which these interactions lie. Observational equations are written for each interaction, and these are summed in the usual way to obtain normal equations which are solved to give the shifts (Δx_i, Δy_i, Δz_i) required of that atom. The new positions for that atom are substituted for the old in the recalculation of the positions of the interactions, and the procedure is repeated for another atom. The method may be applied over a number of cycles, until it converges. If the starting parameters are very crude, the judgment as to which way a given interaction should be moved in order to improve the fit may change in some cases during the early stages, particularly if the conditions of resolution are not good. In the present instance the method was carried through one complete cycle (in the order O_1, O_2, O_3, N, C_1, C_2, C_3) to give Parameters 02, and was not carried further because there seemed to be no case in which a change in judgment regarding the best position for each interaction was justified.

The second method of approach, which is the preferred one if good judgment as to the best positions of all or nearly all interactions can be made, is as follows: Observational equations are written for all or nearly all interactions, each equation corresponding to a desired shift in one coordinate of the position of an interaction if the peak on which the interaction lies is resolved, or to simultaneous shifts in two or more interaction coordinates if the peak is not resolved. The observational equations are rewritten in terms of the desired atom shifts, so that each equation contains at least two atomic parameter shifts. Weights are assigned to the observational equations and normal equations are obtained and solved in accordance with the least-squares procedure. This method has the advantage over the first described in that off-diagonal matrix elements connecting different atomic coordinates are not neglected. In the present work 135 observational equations were written and solved to obtain the 21 atomic parameter shifts.

It will be seen from Table 1 that the interatomic distances obtained from Parameters 03, which are the parameters following the two refinements of the fit to the Patterson function, are within a few hundredths of an Ångström unit of the values expected from previous structural work on amino acids; the mean difference between the distances obtained and those expected is about 0·04 Å (taking as the expected distances 1·24 Å for carboxyl C–O, 1·53 for C–C, 1·47 for C–N, 1·43 for hydroxyl C–O). This agreement appeared to be sufficiently good to justify the initiation of three-dimensional least-squares and Fourier refinement procedures. Therefore, the entire set of structure factors (hydrogen atoms omitted) were calculated from Parameters 03. The calculation of structure factors (and of structure factor derivatives for the least-squares operations) was done with punched cards (Donohue & Schomaker, 1949). The percentage discrepancy, R, was 18·9%, and appeared to be low enough to provide additional justification for the initiation of three-dimensional least-squares and Fourier refinement procedures.

Least-squares refinement

Two successive least-squares refinements were carried out in three dimensions by the procedures outlined by Hughes (1941).

In the case of threonine (Shoemaker *et al.*, 1950) the complete normal equations were solved, and the parameter shifts, which averaged 0·007 Å, were found to differ from those obtained with neglect of the off-diagonal matrix elements by an average of 0·002 Å, or about 30% of the mean shift. At the stage of refinement of the threonine parameters where the least-squares treatment was carried out the effect of the off-diagonal matrix elements was clearly negligible and

Table 1. *Atomic positional parameters for molecule M*

Parameters*		01	03	05	08	09 (Final)
C_1	x	0·2442	0·2523	0·2509	0·2519	0·2514
	y	0·4050	0·4023	0·4054	0·4040	0·4047
	z	0·1567	0·1753	0·1706	0·1703	0·1704
C_2	x	0·2600	0·2570	0·2548	0·2546	0·2547
	y	0·2813	0·2817	0·2788	0·2790	0·2789
	z	0·3907	0·3913	0·3806	0·3810	0·3808
C_3	x	0·3863	0·3867	0·3886	0·3884	0·3885
	y	0·2157	0·2160	0·2098	0·2101	0·2100
	z	0·4573	0·4613	0·4712	0·4718	0·4715
N	x	0·1577	0·1527	0·1530	0·1534	0·1532
	y	0·1700	0·1670	0·1684	0·1687	0·1686
	z	0·2487	0·2500	0·2455	0·2464	0·2460
O_1	x	0·1642	0·1633	0·1627	0·1635	0·1631
	y	0·4050	0·4020	0·4020	0·4022	0·4021
	z	−0·0800	−0·0700	−0·0693	−0·0675	−0·0684
O_2	x	0·3333	0·3337	0·3344	0·3335	0·3340
	y	0·5100	0·5063	0·5050	0·5058	0·5054
	z	0·2500	0·2420	0·2543	0·2523	0·2533
O_3	x	0·4292	0·4297	0·4319	0·4313	0·4316
	y	0·1760	0·1697	0·1680	0·1686	0·1683
	z	0·2160	0·2240	0·2306	0·2282	0·2294
Distances						
C_1–O_1		—	1·29 Å	—	—	1·268 Å
C_1–O_2		—	1·27	—	—	1·261
C_1–C_2		—	1·51	—	—	1·528
C_2–C_3		—	1·46	—	—	1·513
C_2–N		—	1·54	—	—	1·491
C_3–O_3		—	1·42	—	—	1·425

Hydrogen-atom parameters

	x	y	z
$H_1(C_2)$	0·2336	0·3222	0·5728
$H_2(C_3)$	0·3858	0·1147	0·6056
$H_3(C_3)$	0·4572	0·2889	0·5994
H_4(N)	0·1562	0·1182	0·0699
H_5(N)	0·1560	0·0988	0·4076
H_6(N)	0·0692	0·2232	0·2122
$H_7(O_3)$	0·3990	0·0708	0·1704

* See Fig. 1 for the stage of the refinement at which the parameters were obtained.

it was concluded that for three-dimensional work these elements can in general be safely neglected (except for those connecting different coordinates of the same atom when the lattice is not orthogonal) with a consequent material reduction in computational work. This conclusion seemed also to be justified on the theoretical grounds that since the terms which add to give the off-diagonal matrix elements are about as often negative as positive these elements will be smaller by an order of magnitude than the diagonal ones. Moreover, a least-squares treatment with neglect of off-diagonal terms is closely related, or perhaps equivalent, to one step or cycle in the 'method of steepest descents' of Booth, and, except for weighting factors, to a single step of Fourier refinement (Cochran, 1948); in the Fourier method, when carried out in three dimensions, the peak positions (corrected for non-convergence by the 'back-shift' method (Booth, 1946)) are expected to give atomic positions which differ from those ultimately obtainable by amounts that are small in proportion to the magnitudes of the shifts obtained in the one step in consequence of the fact that atoms are resolved in three dimensions.

However, since in the present case the least-squares method was applied to parameters (03) that appeared to be in error (*a priori*) by about 0·03 Å on the average, and since it seems clear that the expected errors due to neglect of off-diagonal matrix elements are proportional to the average magnitude of the shifts obtained, the effect of neglecting off-diagonal matrix elements in one least-squares step might again be expected to give parameters that are in error by an average of about 0·01 Å with respect to the ultimate parameters obtainable with the data. However, a repetition of the procedure should be expected to reduce by the same ratio the errors starting from neglect of such elements. In fact, a direct correspondence can be drawn between the procedure (I) of carrying out a succession of N least-squares refinements with neglect of off-diagonal matrix elements, and the procedure (II) of carrying out a single least-squares refinement without neglect of any matrix elements and solving the normal equations by iteration through N successive approximation steps where, in the nth iteration step, the equations, before being solved, are simplified by substituting the results of the $(n-1)$ step (or zeros, if $n = 1$) into the terms corresponding to all matrix elements that would have been neglected in the first procedure. The correspondence is exact if in the first procedure the structure-factor derivatives are not recalculated and if the new values of the calculated structure factors to be used in each step are calculated with the use of the previous shifts and the structure-factor derivatives (in other words, if the problem is treated as if the observational equations were linear in the parameters), provided further that the signs assigned to the observed structure factors are not changed while this procedure is being carried out. Under such conditions the first procedure is probably the most economical of computational effort if the number, N, of steps is limited to, say, two or three, for the time and labor required for the computation of all off-diagonal matrix elements is considerable.

In the present case the average of the shifts obtained in the first least-squares step is 0·018 Å, which may presumably be taken as a good measure of the average deviation of the starting parameters (03) from the ones ultimately obtainable at convergence of the procedure. The average of the shifts in the second least-squares step is 0·0056 Å, or 30% of the average of the first shifts; the value 0·0056 Å is presumably a good measure of the error introduced in the first step by the omission of off-diagonal matrix elements, and the percentage figure presumably applies to the second step. From this it can be inferred that omission of off-diagonal matrix elements resulted in an average error of about 0·0017 Å in Parameters 05. This average error is negligible in comparison with other errors which enter into these parameters as the result of errors in the experimental data.

Three-dimensional Fourier refinement

It is apparent from the discussion above that no significant further improvement in parameter accuracy would be likely to result from additional least-squares refinement. On the other hand it seemed advisable to carry out a stage of refinement by another method subject to computational errors somewhat different in character from those to which the least-squares method is subject. Such a method is the three-dimensional Fourier method. As has been stated above, this method, when used with the back-shift rule, is theoretically equivalent to the least-squares method to first order, except in the matter of weights, but the two methods differ in computational technique and comparison of the results obtained affords a control on the magnitudes of computational errors.

The starting point of the Fourier refinement was Parameters 04, which were obtained from the first least-squares operation. Two three-dimensional Fourier syntheses were prepared, one with the observed structure factors and the other with the calculated structure factors. In each, all planes were included for which a non-zero value was found for the observed intensity. The computations were carried out by means of punched cards in the way described in connection with the Patterson function.

The positions of the maxima in the observed and calculated electron-density functions were calculated by use of an analytical method involving least-squares (Carpenter & Donohue, 1950; Shoemaker *et al.*, 1950). From these, which were designated Parameters 06 and 07 respectively, new parameters, designated 08, were deduced by application of the back-shift rule whereby the shifts 04 to 07 were subtracted from the respective shifts 04 to 06 to obtain shifts corrected for effects due to peak overlap and termination of the series at a finite limit, and these shifts were added to the starting parameters, 04, to obtain the new ones, 08.

A composite Fourier synthesis representing sections through atomic peaks in three-dimensional Fourier Synthesis 1 (computed with observed F's) is shown in Fig. 5. The peaks shown represent the carbon, nitrogen and oxygen atoms in one molecule, which is the one designated M in the drawings of the structure given in Fig. 7.

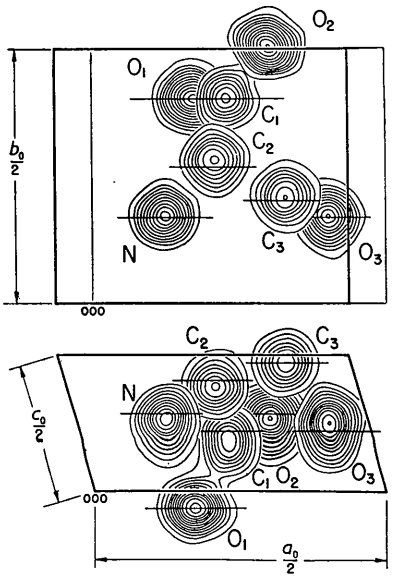


Fig. 5. Fourier synthesis (composite).

Parameters 05, obtained from the second least-squares refinement, and Parameters 08, obtained from the Fourier refinement, were found to be in very good agreement, the average difference being 0·006 Å and the maximum difference 0·013 Å.

Parameters 05 and Parameters 08 were assumed to be about equally precise and therefore were averaged with equal weight to give new parameters, which were designated 09. The standard deviation of the mean, as calculated from the root-mean-square difference between the two sets of Parameters 05 and 08, is 0·0036 Å.

Hydrogen atoms

Since visual intensity data and the assumed form factors are generally incapable of enabling precise values to be obtained for the hydrogen-atom parameters in a structure of this kind, hypothetical parameters were computed on the basis of previous knowledge of bond distances, bond angles and hydrogen-bond distances. Hydrogen atoms were placed in such a way as to give bond angles as close to tetrahedral as possible, and the distances were made to be 1·09 Å for C–H, 1·01 Å for N–H, and 0·97 Å for O–H. The molecule was assumed to have the *Zwitterion* structure, as in other amino acids the crystal structures of which have been determined, and therefore three hydrogen atoms were placed on the nitrogen atom. In view of the one rotational degree of freedom around the C–N bond, and the fact that the nitrogen atom has three oxygen neighbors (in other molecules) at distances (2·79, 2·81 and 2·87 Å) that are in essential agreement with previously found N–H · · · O bond distances and at angles that are not far from tetrahedral with respect to the C–N bond and to each other, the three hydrogen atoms were assumed to lie at positions that are the average of positions lying on the N · · · O lines and strictly tetrahedral positions adjusted with the one rotational degree of freedom so as to be as close to those lines as possible. The position assumed for the hydroxyl hydrogen atom was obtained in the same way, since the hydroxyl oxygen atom (O_3) has a carboxyl oxygen neighbor (in another molecule) at a distance of 2·67 Å, which is in essential agreement with values previously found for O–H · · · O hydrogen bonds, and at an angle that is not far from tetrahedral with respect to the hydroxyl C–O bond. The hydrogen-atom parameters obtained in this way are given in Table 1.

Although precise values for the hydrogen-atom parameters cannot be obtained or checked experimentally in work of this kind, it may be expected on the basis of experience that small peaks corresponding to the hydrogen atoms should be present in the electron-density function calculated with the observed structure factors. No attempt was made to find hydrogen peaks on the function obtained in Fourier Synthesis 1, because it was anticipated that these small peaks would be partially obscured by effects due to overlap of the larger peaks and termination of the series. In principle they should, however, appear on the difference Fourier obtained by subtracting the function obtained in Fourier Synthesis 2 (which, being computed with calculated structure factors containing no hydrogen contributions, should of course contain no hydrogen peaks at all but should have much the same effects due to peak overlap and series termination as the function obtained in Synthesis 1) from the function obtained in Fourier Synthesis 1, if the experimental data are sufficiently precise and if the assumed form factors for the heavy atoms represent the shapes of those atoms well enough. Such a difference Fourier was prepared and examined. It was found to contain many irregular peaks, most of which could not reasonably correspond to the position of a hydrogen atom. However, all of the hypothetical hydrogen positions obtained as described above coincide roughly with peaks on the difference Fourier. This may be taken as partial confirmation of the assumed positions and as good evidence that the assumed hydrogen bonding picture is essentially correct. Additional confirmation of the correctness of the assignment of hydrogen parameters was obtained later in the observation that inclusion of the hydrogen atoms with these parameters in the computation of calculated structure factors resulted in significant improvement in the agreement with observed structure factors.

In order to determine whether or not additional refinement of the heavy-atom parameters should be carried out with structure factors containing hydrogen-atom contributions, a study was made of the effect of the omission of hydrogen atoms in the Fourier refinement. Since structure-factor calculations based on Parameters 09 showed that the signs of only one or two very weak reflections were changed by the inclusion of hydrogen atoms, the effect of the omission of hydrogen atoms on Parameters 08 was assumed to lie entirely in the back-shift corrections obtained with Fourier Synthesis 2. The amount by which the addition of a hydrogen peak to Synthesis 2 shifts the maximum of a heavy-atom peak at a given distance can easily be calculated from the known slope of the hydrogen-atom peak at that distance and the known curvature of the heavy-atom peak. The shifts in all peaks were calculated in this manner; for each peak, only shifts due to hydrogen atoms bonded to the corresponding atom were taken into account. The parameter shifts obtained in this way were all 0·001 Å or less. Hence the effects of the omission of hydrogen atoms from the Fourier refinement process can be considered negligible, and it can be reasonably inferred that the same is true of the least-squares refinement.

Final structure factors

Two sets of structure factors were calculated from Parameters 09: Set 3 without hydrogens, and Set 4

Table 2. *Observed and calculated (final, set No. 4) structure factors*

h	F_o	F_c
	*h*00	
0	—	320·0
2	38·4	−49·5
4	25·8	−28·3
6	0	5·0
8	1·9	0·9
10	18·4	−20·5
12	12·2	16·9
	*h*10	
1	8·4	− 7·9
2	2·8	1·3
3	17·7	−16·2
4	36·5	40·4
5	12·4	12·2
6	6·6	6·4
7	9·1	9·1
8	18·4	−19·2
9	2·8	− 1·6
10	2·7	3·3
11	0	− 0·4
12	2·2	1·7
	*h*20	
0	18·2	−18·3
1	7·4	5·9
2	15·1	−16·0
3	37·3	−45·5
4	4·0	− 5·4
5	6·7	2·2
6	24·7	26·0
7	3·8	− 2·2
8	8·1	− 8·4
9	8·6	8·5
10	0	− 1·2
11	2·3	− 3·6
12	0	3·2
	*h*30	
1	27·7	28·4
2	29·0	−27·3
3	8·3	− 7·4
4	15·3	14·0
5	7·3	− 6·1
6	3·4	4·3
7	10·9	− 8·2
8	4·0	− 4·7
9	4·7	− 6·9
10	1·2	1·6
11	5·2	5·1
	*h*40	
0	7·6	− 8·2
1	21·7	27·2
2	3·5	− 2·4
3	25·6	22·3
4	2·4	1·4
5	9·0	− 6·8
6	0	0·9
7	1·8	2·4
8	2·6	3·5
9	0	0·6
10	5·0	4·6
11	5·7	− 5·9
	*h*50	
1	7·4	6·7
2	1·3	− 0·2
3	29·5	−28·0
4	12·6	−12·4
5	20·5	17·6
6	8·3	− 7·6
7	12·5	14·6
8	4·3	3·0
9	12·2	−13·6
10	5·3	5·2
11	0	1·0
	*h*60	
0	12·6	11·4
1	7·5	− 8·4
2	12·4	11·9
3	9·6	− 9·6
4	15·4	−15·1
5	10·7	10·3
6	8·0	7·1
7	1·6	− 1·4
8	10·8	−11·5
9	3·8	1·1
10	1·3	− 0·5
	*h*70	
1	8·7	9·0
2	6·5	6·8
3	19·2	−18·0
4	8·0	7·1
5	0	1·0
6	3·3	3·4
7	10·5	10·0
8	2·8	− 1·9
9	2·0	− 2·9
	*h*80	
0	0	− 0·2
1	23·9	−22·2
2	6·0	− 7·2
3	9·1	6·7
4	1·9	1·8
5	4·2	− 3·7
6	7·2	6·1
7	2·8	− 2·2
8	0	− 0·7
	*h*90	
1	0	0·6
2	0	− 0·8
3	6·2	6·0
4	3·3	− 3·0
5	0	1·4
6	1·3	1·1
7	4·3	− 8·3
	h,10,0	
0	13·2	13·4
1	4·1	4·3
2	9·4	− 8·9
3	1·4	0·2
4	1·4	− 1·4
5	1·8	4·3
	*h*01	
−12	5·2	7·1
−10	4·6	3·7
− 8	8·6	−10·7
− 6	4·8	6·2
− 4	9·2	− 9·0
− 2	31·5	−38·3
0	2·1	1·5
2	22·6	22·4
4	0	0·2
6	25·6	23·6
8	11·6	− 8·9
10	10·4	−13·4
	*h*11	
−13	3·6	− 9·4
−12	4·8	− 4·2
−11	0	1·1
−10	12·9	12·7
− 9	2·7	3·0
− 8	2·0	0·0
− 7	0	− 3·1
− 6	10·1	8·2
− 5	8·5	− 4·8
− 4	6·9	− 5·9
− 3	38·0	40·4
− 2	34·6	32·6
− 1	21·6	−20·5
0	51·8	−57·6
1	5·2	2·3
2	19·6	−19·0
3	12·6	− 9·2
4	6·0	5·6
5	6·0	− 6·7
6	18·2	19·4
7	6·2	− 5·0
8	3·1	− 2·2
9	8·4	10·1
10	7·2	6·8
11	0	− 2·6
	*h*21	
−12	0	2·3
−11	3·9	3·2
−10	2·6	2·9
− 9	0	− 1·8
− 8	6·2	− 7·3
− 7	20·7	−20·6
− 6	1·9	0·4
− 5	3·6	2·7
− 4	29·0	27·7
− 3	23·8	24·4
− 2	30·0	−35·7
− 1	17·9	−16·8
0	33·9	31·8
1	24·6	24·6
2	7·4	8·2
3	23·7	24·8
4	25·0	−25·3
5	22·1	−21·3
6	10·8	−10·0
7	13·4	−12·5
8	18·4	20·5
9	3·6	3·4
10	4·9	− 5·3
11	0	− 0·3
	*h*31	
−12	1·3	− 1·4
−11	9·9	10·3
−10	5·0	− 7·4
− 9	1·7	− 2·1
− 8	6·4	7·6
− 7	18·4	−18·7
− 6	8·7	8·8
− 5	23·6	22·3
− 4	23·0	−21·8
− 3	24·4	−25·6
− 2	5·7	5·9
− 1	32·3	−38·4
0	10·2	8·2
1	29·2	32·9
2	8·3	− 6·2
3	26·8	24·8
4	6·7	4·3
5	9·9	− 9·2
6	3·8	4·2
7	13·6	13·0
8	11·1	−11·2
9	6·9	− 6·5
10	5·5	5·5
11	6·6	− 9·4
	*h*41	
−12	3·2	− 5·1
−11	6·8	7·3
−10	4·4	5·6
− 9	2·8	− 2·9
− 8	10·4	−12·2
− 7	8·8	7·4
− 6	6·2	4·9
− 5	3·7	3·9
− 4	20·6	20·9
− 3	1·2	− 0·6
− 2	3·0	1·4
− 1	26·2	−21·5
0	28·2	−26·1
1	7·8	4·6
2	24·4	20·5
3	20·8	−19·8
4	15·5	−14·0
5	10·3	8·7
6	4·2	− 3·2
7	13·8	13·8
8	1·7	2·0
9	4·3	4·2
10	4·1	3·5
11	2·8	− 9·7
	*h*51	
−11	0	2·7
−10	3·3	− 3·0
− 9	0	− 0·7
− 8	10·6	− 9·7
− 7	3·8	− 3·1
− 6	5·0	4·2
− 5	12·4	14·0
− 4	9·9	−10·0
− 3	8·3	− 8·4
− 2	11·4	10·4
− 1	6·3	5·1
0	18·6	17·1
1	14·9	15·1
2	14·6	3·2
3	20·3	−19·1
4	10·0	−11·2
5	13·7	−15·3
6	4·4	4·1
7	13·9	14·1
8	6·0	− 5·5
9	0	− 1·6
10	2·9	− 3·7
	*h*61	
−10	7·5	8·7
− 9	1·8	2·0
− 8	4·4	3·6
− 7	3·3	4·7

Table 2 (*cont.*)

h	F_o	F_c
−6	12·4	−13·4
−5	4·3	3·0
−4	7·5	−6·2
−3	14·1	−14·2
−2	0	−0·4
−1	12·3	10·5
0	8·9	−7·7
1	4·2	−2·5
2	11·0	11·2
3	5·6	5·5
4	9·7	11·2
5	5·8	−3·2
6	1·7	0·4
7	0	−1·3
8	3·6	−2·1
9	3·9	−4·9
	h71	
−10	3·4	4·8
−9	0	−0·6
−8	3·6	3·4
−7	7·0	−7·9
−6	6·8	−5·4
−5	18·5	19·7
−4	16·0	16·4
−3	6·0	−3·5
−2	15·6	−14·4
−1	0	−2·1
0	10·5	−10·4
1	1·6	−3·9
2	2·2	−1·8
3	3·6	2·4
4	7·1	6·3
5	5·8	−6·0
6	3·3	−4·0
7	6·7	7·1
8	7·8	9·9
9	2·2	−7·1
	h81	
−8	6·4	−5·8
−7	6·6	−5·1
−6	3·9	3·2
−5	5·7	−6·8
−4	10·0	8·8
−3	6·8	7·5
−2	6·8	−6·6
−1	11·4	12·4
0	1·7	2·5
1	3·9	−2·9
2	5·9	5·7
3	3·2	−4·5
4	12·6	−12·6
5	7·9	6·1
6	0	−2·0
7	7·6	−5·9
8	1·6	7·2
	h91	
−7	0	−2·0
−6	0	−0·3
−5	6·8	−7·0
−4	4·4	3·9
−3	2·5	2·1
−2	9·0	−8·8
−1	14·3	−13·5
0	6·4	4·8
1	14·2	15·2
2	0	2·5
3	2·8	2·5
4	0	0·5
5	2·2	−2·4
6	5·0	−5·0
	h,10,1	
−5	5·3	5·3
−4	8·8	9·1
−3	10·9	−11·8
−2	9·6	−9·5
−1	2·2	1·6
0	2·8	1·7
1	3·2	−2·4
2	2·8	2·4
3	0	−0·7
4	4·2	−5·3
	h02	
−12	6·7	−7·4
−10	18·8	21·1
−8	5·2	3·9
−6	2·5	−2·5
−4	4·6	6·2
−2	10·1	−12·5
0	29·5	−33·0
2	14·3	12·2
4	3·6	−1·9
6	2·1	−0·3
8	11·5	12·4
10	4·0	5·2
	h12	
−13	4·0	−3·8
−12	2·0	−1·4
−11	3·4	−4·4
−10	10·5	−10·5
−9	4·2	7·6
−8	2·6	2·9
−7	8·5	−8·8
−6	19·3	18·2
−5	26·3	−24·5
−4	12·2	9·7
−3	41·2	45·1
−2	21·4	−15·4
−1	6·9	−5·3
0	10·0	7·6
1	18·5	−16·3
2	6·2	−6·4
3	21·8	21·1
4	8·3	−9·7
5	0	−0·5
6	9·8	−9·4
7	16·4	−19·1
8	4·8	5·0
9	6·5	7·8
10	3·3	4·2
	h22	
−13	1·3	1·6
−12	0	−0·5
−11	9·2	9·7
−10	6·6	−5·3
−9	7·4	5·8
−8	4·4	−5·2
−7	21·0	−22·6
−6	2·0	−3·1
−5	4·2	3·8
−4	10·6	9·1
−3	7·0	−4·8
−2	5·7	5·3
−1	23·1	−21·4
0	10·4	−7·9
1	14·1	13·0
2	15·9	14·8
3	21·1	23·3
4	2·1	−0·5
5	12·0	−10·3
6	3·7	−4·7
7	7·2	6·0
8	4·4	−4·4
9	2·0	0·6
10	1·7	−0·8
	h32	
−12	2·2	−1·3
−11	1·3	1·0
−10	4·2	3·8
−9	9·8	9·1
−8	12·3	−10·4
−7	13·2	11·5
−6	15·2	15·0
−5	6·0	−4·4
−4	24·9	−20·0
−3	7·4	−5·5
−2	29·2	23·9
−1	15·4	−13·8
0	7·0	−5·3
1	0	−2·9
2	3·4	−4·4
3	6·9	−4·5
4	10·5	−9·2
5	0	0·9
6	15·8	18·6
7	8·9	9·5
8	10·3	−9·8
9	5·5	7·0
10	1·6	2·6
	h42	
−12	1·9	−5·0
−11	9·5	−8·8
−10	1·5	0·0
−9	5·4	5·2
−8	7·4	8·6
−7	3·9	2·8
−6	18·5	−19·7
−5	6·8	6·8
−4	2·0	1·2
−3	9·3	8·4
−2	21·7	19·9
−1	4·1	3·0
0	1·5	−3·2
1	8·1	−8·1
2	0	−1·5
3	2·4	0·6
4	8·8	8·0
5	10·5	−10·5
6	11·2	−13·9
7	0	−0·9
8	2·9	3·5
9	0	1·8
	h52	
−12	0	1·4
−11	5·4	−5·7
−10	0	0·2
−9	4·3	3·6
−8	9·5	7·5
−7	20·3	−19·6
−6	11·1	−9·0
−5	16·8	17·5
−4	11·6	−10·3
−3	0	1·1
−2	6·6	−6·1
−1	0	−1·4
0	8·8	7·2
1	9·3	−6·6
2	1·3	0·8
3	19·8	20·8
4	8·9	10·9
5	14·1	−13·4
6	2·8	0·4
7	2·6	1·6
8	3·0	−3·4
9	3·2	−6·2
	h62	
−11	1·8	3·2
−10	6·0	5·9
−9	2·3	−3·1
−8	14·2	14·3
−7	2·6	−2·4
−6	14·0	−14·3
−5	0	−0·7
−4	10·0	9·2
−3	2·6	3·1
−2	6·8	4·8
−1	0	2·8
0	26·1	−28·1
1	0	−1·6
2	4·0	−3·0
3	9·1	−7·9
4	16·1	18·9
5	6·5	6·6
6	8·1	−8·0
7	1·8	2·1
8	1·8	3·9
	h72	
−10	0	0·5
−9	3·1	3·9
−8	4·0	−4·0
−7	5·4	−5·3
−6	0	0·8
−5	9·2	−9·4
−4	18·2	19·4
−3	10·4	11·3
−2	8·8	−7·5
−1	0	−0·4
0	4·6	−4·8
1	4·4	5·2
2	6·5	5·3
3	0	0·7
4	6·9	−7·1
5	3·6	−5·6
6	8·3	−8·9
7	3·4	−3·8
	h82	
−9	1·6	−4·0
−8	1·7	−1·9
−7	4·7	6·1
−6	1·5	−2·0
−5	3·7	−3·1
−4	1·7	2·4
−3	10·0	−10·8
−2	2·8	2·8
−1	4·1	−3·8
0	0	−0·7
1	10·7	12·2
2	3·5	3·8

Table 2 (*cont.*)

h	F_o	F_c
3	0	2·1
4	0	0·1
5	0	1·0
6	2·1	−3·4
	*h*92	
−7	3·7	4·3
−6	6·9	−6·1
−5	1·3	1·4
−4	7·8	7·4
−3	1·5	−1·0
−2	1·8	−2·0
−1	3·1	−4·6
0	0	−0·9
1	10·3	−10·2
2	0	−0·7
3	5·4	6·1
4	1·1	1·4
	h,10,2	
−4	1·0	−2·0
−3	5·6	6·0
−2	1·1	2·2
−1	3·2	2·9
0	2·0	2·8
1	4·4	−4·9
2	3·6	4·6
	*h*03	
−12	0	0·7
−10	8·6	−8·2
−8	9·9	−9·7
−6	24·9	23·3
−4	18·2	17·2
−2	8·4	−6·0
0	9·8	−10·3
2	2·5	−0·3
4	7·5	−9·1
6	2·9	−0·2
8	0	0·3
	*h*13	
−12	4·7	5·6
−11	1·3	−2·6
−10	8·4	−7·0
−9	4·9	2·3
−8	13·9	−14·6
−7	8·0	8·3
−6	5·2	5·6
−5	23·1	−22·2
−4	8·6	−7·5
−3	4·4	−2·8
−2	18·8	15·8
−1	16·8	16·0
0	16·9	16·4
1	10·2	−10·1
2	3·9	−3·4
3	5·3	4·8
4	12·5	−12·6
5	9·0	9·4
6	11·5	9·7
7	7·2	−6·9
8	4·9	−5·9
	*h*23	
−11	4·7	−3·8
−10	7·3	6·0
−9	11·3	11·6
−8	0	2·4
−7	5·5	3·2
−6	1·6	−2·1
−5	6·8	4·6
−4	23·6	−27·2
−3	12·1	−9·4
−2	21·2	18·7
−1	11·4	−11·1
0	3·7	3·0
1	7·4	−7·1
2	2·5	−1·7
3	11·4	9·4
4	1·3	0·1
5	7·1	−4·9
6	5·9	6·3
7	6·8	7·0
8	4·5	−5·5
	*h*33	
−12	0	1·6
−11	13·0	−13·0
−10	5·4	−6·1
−9	0	−1·0
−8	7·2	6·2
−7	12·4	12·4
−6	4·5	5·6
−5	1·1	2·8
−4	8·1	−6·4
−3	9·1	6·5
−2	10·4	−9·1
−1	12·2	12·9
0	15·3	14·1
1	20·7	−22·4
2	7·3	−6·2
3	4·8	−5·7
4	2·2	1·9
5	7·3	6·0
6	2·0	−2·3
7	2·8	−3·0
8	0	0·8
	*h*43	
−11	3·4	−2·6
−10	12·4	−12·4
−9	7·1	5·9
−8	15·5	17·6
−7	21·4	−24·2
−6	11·3	−10·9
−5	5·1	3·7
−4	8·7	−9·6
−3	16·9	17·0
−2	4·0	6·9
−1	7·3	7·7
0	5·4	4·9
1	13·8	−11·4
2	17·3	−16·4
3	14·1	13·1
4	19·7	21·7
5	9·4	−8·2
6	2·6	−3·1
7	1·1	−2·2
8	4·5	−6·7
	*h*53	
−11	6·2	7·8
−10	7·3	7·3
−9	1·8	−1·5
−8	1·9	1·9
−7	1·4	−0·4
−6	1·4	−0·4
−5	5·5	−7·5
−4	0	1·4
−3	3·8	−4·0
−2	3·9	0·5
−1	0	−0·1
0	17·3	−16·0
1	14·4	12·2
2	0	0·6
3	5·5	−5·0
4	3·4	4·2
5	1·3	0·5
6	2·8	3·5
7	0	0·0
	*h*63	
−10	6·9	−8·4
−9	4·7	−3·6
−8	6·0	6·5
−7	12·8	11·2
−6	2·8	2·3
−5	8·3	−6·2
−4	4·4	3·6
−3	4·4	−2·5
−2	7·4	7·2
−1	3·9	−4·0
0	0	−0·9
1	13·6	9·4
2	13·2	−12·9
3	4·3	−3·9
4	5·7	3·4
5	3·5	1·7
6	6·3	−5·2
	*h*73	
−9	1·2	−2·2
−8	3·7	−2·7
−7	0	0·0
−6	4·8	−5·4
−5	4·5	4·2
−4	4·0	4·2
−3	8·0	−6·5
−2	3·7	4·5
−1	0	0·4
0	3·3	3·1
1	6·6	−3·5
2	2·4	2·6
3	8·4	8·3
4	2·2	−2·2
5	1·0	1·2
	*h*83	
−8	1·8	4·7
−7	8·2	6·0
−6	5·4	−3·8
−5	0	0·8
−4	9·3	−8·8
−3	4·6	−4·1
−2	6·8	5·1
−1	0	−0·1
0	4·5	3·2
1	0	−1·7
2	4·8	−3·5
3	9·7	−8·4
4	3·9	3·9
	*h*93	
−6	1·5	−2·6
−5	11·9	−10·6
−4	1·1	1·2
−3	6·2	7·2
−2	1·1	−0·7
−1	9·8	9·1
0	3·3	−3·4
1	7·9	−8·0
2	4·5	4·4
	*h*04	
−10	2·8	3·7
−8	15·6	−19·8
−6	10·9	11·9
−4	8·8	−8·4
−2	4·9	5·8
0	12·2	16·2
2	7·6	7·6
4	17·0	−23·2
6	7·2	7·4
	*h*14	
−11	2·2	−3·6
−10	0	−1·7
−9	0	−2·0
−8	7·3	8·0
−7	9·4	9·6
−6	4·2	−4·1
−5	5·4	4·9
−4	13·9	−14·8
−3	15·3	−19·4
−2	5·3	−6·6
−1	10·5	10·0
0	13·6	15·3
1	5·0	−4·8
2	0	0·1
3	4·1	−3·8
4	3·6	5·6
5	4·7	4·6
6	1·4	−2·0
	*h*24	
−11	7·8	−9·5
−10	2·0	−3·2
−9	7·0	8·1
−8	1·3	0·7
−7	2·0	−2·7
−6	5·1	6·2
−5	2·9	2·2
−4	7·7	6·1
−3	10·0	11·5
−2	16·0	−16·9
−1	5·2	5·6
0	2·4	1·2
1	15·2	−19·1
2	2·2	−2·0
3	3·6	1·4
4	5·9	5·9
5	0	−0·1
	*h*34	
−10	6·0	−6·3
−9	3·5	−4·4
−8	0	−1·2
−7	10·4	−8·4
−6	0	1·0
−5	0	−0·3
−4	6·2	4·1
−3	5·4	5·8
−2	7·6	−5·4
−1	7·1	7·8
0	10·7	7·9
1	4·6	7·1
2	8·6	−8·4
3	4·4	−5·9

Table 2 (*cont.*)

h	F_o	F_c
4	1·8	0·9
5	2·6	−3·0
6	3·0	3·8
	$h44$	
−10	2·7	1·9
−9	0	2·1
−8	0	1·6
−7	2·4	−1·5
−6	5·3	4·5
−5	4·5	−6·8
−4	9·6	−10·0
−3	2·4	−3·0
−2	4·5	2·8
−1	6·6	−6·2
0	2·8	1·8
1	9·0	9·9
2	7·3	−5·3
3	7·3	7·9
4	0	−0·9
5	3·9	−3·1
	$h54$	
−9	9·8	−9·7
−8	0	−0·6
−7	9·6	7·7
−6	5·9	5·6
−5	7·3	8·4
−4	5·6	5·8
−3	8·7	−7·2
−2	2·6	−3·2
−1	11·8	−12·4
0	0	−0·4
1	11·0	9·4
2	0	−1·2
3	4·0	−4·2
4	1·8	−3·4
	$h64$	
−9	0	−1·8
−8	0	0·3
−7	3·7	−3·6
−6	3·2	2·9
−5	5·6	4·6
−4	2·9	−3·4
−3	3·2	−1·2
−2	2·8	−0·3
−1	4·7	4·9
0	14·0	16·7
1	6·7	−6·2
2	8·1	−8·2
3	0	−1·1
4	0	−2·2
	$h05$	
−8	2·2	0·9
−6	10·0	−9·7
−4	9·4	8·9
−2	11·3	−13·3
0	2·9	3·5
2	4·2	5·5
	$h15$	
−9	3·6	3·5
−8	3·3	2·2
−7	2·0	−1·6
−6	7·2	−4·7
−5	5·1	4·6
−4	10·4	8·8
−3	3·5	−2·8
−2	9·1	−7·7
−1	6·5	5·1
0	0	−3·6
1	4·2	−4·4
2	5·8	−5·5
3	1·4	−2·2
	$h25$	
−9	0	1·1
−8	6·4	−7·1
−7	1·3	−1·5
−6	0	0·4
−5	8·0	−8·3
−4	5·2	6·9
−3	12·0	12·4
−2	2·1	−1·0
−1	0	−0·4
0	3·1	−3·0
1	0	1·1
2	5·1	4·9
	$h35$	
−8	0	−0·2
−7	5·0	−6·3
−6	3·6	4·0
−5	3·2	4·1
−4	1·8	−2·2
−3	0	−1·2
−2	5·9	−6·6
−1	10·2	−12·6
0	6·3	7·5
1	8·3	12·1
	$h45$	
−8	1·8	−2·0
−7	3·0	2·7
−6	4·6	5·0
−5	0	0·6
−4	4·6	4·8
−3	2·1	−2·3
−2	4·5	−5·4
−1	0	−1·8
0	0	−0·6
1	3·8	−4·9
	$h55$	
−7	1·5	−5·8
−6	0	0·4
−5	0	1·7
−4	1·7	2·1
−3	6·2	6·3
−2	1·5	0
−1	5·5	−4·1
0	4·8	7·2
	$h65$	
−4	4·6	−8·7
−3	1·8	−3·7
−2	3·0	4·1

with hydrogen-atom contributions out to about $\sin \theta = 0{\cdot}6$. The latter set is compared with the set of observed structure factors in Table 2.

The average discrepancy between observed and calculated structure factors (with hydrogens) is 14·5%. Inclusion of the hydrogen atoms reduced the percentage discrepancy for all planes from 15·2 to 14·5%. Within the range $\sin \theta \leqslant 0{\cdot}6$, in which hydrogen atom contributions were included, the percentage discrepancies without and with hydrogen atoms are respectively 13·6 and 12·1%.

Final parameters and precision of their determination

Since it appears that the refinement procedures have converged, and since the omission of hydrogen atoms from the refinement appears to have no significant effect on the results obtained, Parameters 09 have been taken as the final positional parameters of the carbon, nitrogen and oxygen atoms, and as the basis for the calculation of all interatomic distances and bond angles discussed later on.

Standard errors were calculated by the methods used for threonine (Shoemaker *et al.*, 1950), the parameter n being set equal to unity because of the presence of a center of symmetry. The mean standard errors in the least-squares parameters (05) due to experimental errors and form-factor errors alone were found to be 0·0063, 0·0056 and 0·0046 Å, for carbon, nitrogen and oxygen respectively, and those in the Fourier parameters (08) were found to be 0·0052, 0·0050 and 0·0042 Å respectively. These figures represent averages over the values for the standard errors in x, y and z parameters, and are in good agreement with one another. For reasons discussed in the threonine paper (Shoemaker *et al.*, 1950) the mean of the least-squares and Fourier standard errors for each parameter were combined with the standard error of the mean parameters (09), 0·0036 Å, in the sense of the root of the sum of the squares, to give the final standard errors which were found to be 0·0068, 0·0064 and 0·0057 Å, for carbon, nitrogen and oxygen respectively. These correspond to standard errors of 0·012, 0·011 and 0·010 Å in the positions of the respective atoms. The mean of the carbon, nitrogen and oxygen x, y, and z standard parameter errors, 0·0063 Å, corresponds to a standard error of 0·0089 Å in an interatomic distance. The corresponding probable error in a distance is 0·0060 Å, and the corresponding limit of error (taken as three times the probable error or twice the standard error) is 0·018 Å.

Discussion of the structure

From the final parameters given in Table 1 (column 09) all important interatomic distances were calculated. Intramolecular interatomic distances and bond angles are given in Table 3 and intermolecular interatomic

Table 3. *Intramolecular interatomic distances and bond angles*

	Distance (Å)		Distance (Å)
C_1–C_2	1·528	$H_1(C_2)\ldots H_5(N)$	2·26
C_2–C_3	1·513	$H_1(C_2)\ldots H_6(N)$	2·28
C_2–N	1·491	$H_1(C_2)\ldots H_2(C_3)$	2·48
C_1–O_1	1·268	$H_1(C_2)\ldots H_3(C_3)$	2·38
C_1–O_2	1·261	$H_2(C_3)\ldots H_4(N)$	2·97
C_3–O_3	1·425	$H_2(C_3)\ldots H_5(N)$	2·38
		$H_2(C_3)\ldots H_7(O_3)$	2·18
		$H_3(C_3)\ldots H_7(O_3)$	2·81
$O_1\ldots O_2$	2·25		
$N\ldots O_1$	2·64		Angle (°)
$N\ldots O_3$	3·01	O_1–C_1–O_2	125·3
$C_1\ldots O_3$	2·86	O_1–C_1–C_2	117·4
$O_2\ldots C_3$	2·88	O_2–C_1–C_2	117·2
$O_1\ldots H_1(C_2)$	3·06	C_1–C_2–N	110·0
$O_2\ldots H_1(C_2)$	2·70	C_3–C_2–N	111·1
$O_1\ldots H_4(N)$	2·68	C_1–C_2–C_3	110·3
$O_1\ldots H_6(N)$	2·51	C_2–C_3–O_3	112·0
$C_3\ldots H_4(N)$	2·75		
$C_3\ldots H_5(N)$	2·63		
$O_2\ldots H_3(C_3)$	2·68		

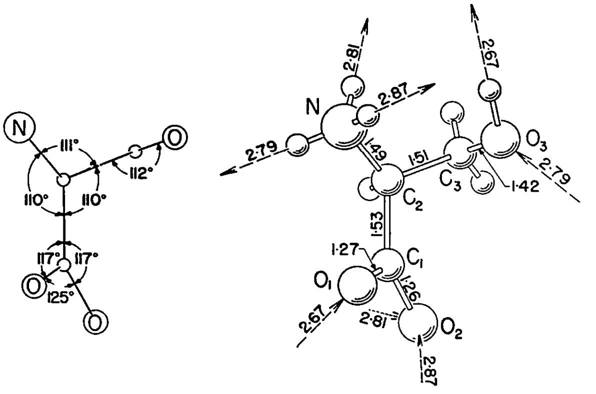


Fig. 6. The serine molecule, showing bond distances and angles.

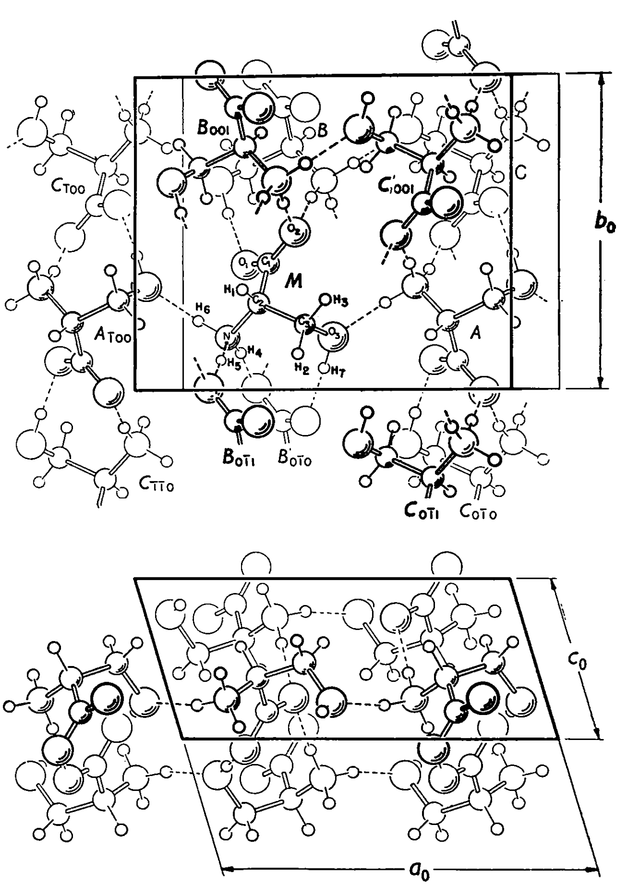


Fig. 7. Two views of the structure. (*a*) View normal to the *X*, *Y* plane. (*b*) View along the *b* axis, in the direction of increasing *y*.

distances (including hydrogen-bond distances) in Table 4. The molecule, with interatomic distances and angles, is shown in Fig. 6, and one unit cell of the structure (with a few additional molecules in surrounding unit cells) in Fig. 7.

The intramolecular interatomic distances and bond angles contain no surprises; all distances are in essential agreement with those found earlier for other amino acids and related molecules. In particular the α-carbon–nitrogen distances of 1·49 Å is in essential agreement with the distances 1·49 Å found for threonine, 1·50 Å for alanine (after recent three-dimensional refinement), 1·48 and 1·50 Å for hydroxyproline, and 1·48 and 1·51 Å for β-glycylglycine. The agreement with the value 1·45 Å reported for N-acetylglycine is not as good, and it is possible that there is a real difference.

Configurations around the C_2–C_3 and C_2–N bonds are ethane-like (staggered). The hydroxyl oxygen atom, unlike that in threonine, lies in what might be considered the least favorable of the three possible positions, being close to both the amino group and the carboxyl group. The carboxyl group is almost coplanar with the C_2–N bond. Van der Waals contacts are, however, essentially normal; the distances from O_1 of the carboxyl group to the two close hydrogens of the amino group are 2·68 and 2·51 Å, while the distance expected from conventional van der Waals radii is 2·6 Å, though the O_1–N distance of 2·61 Å is somewhat shorter than the sum of the van der Waals radii, 2·9 Å. Contacts of O_2 of the carboxyl group with H_1 on the α-carbon and H_3 on the β-carbon are respectively 2·70 and 2·68 Å, both of which exceed the sum of the van der Waals radii. Hydrogen–hydrogen contact distances in the molecule are about equal to the expected distance, 2·4 Å, except that the position assumed for H_7 of the hydroxyl group is only 2·18 Å from H_2 of the β-carbon; however, the actual distance may be longer as this short distance was not taken into consideration in assigning the position of this hydrogen atom, the primary consideration being the hydrogen bond in which this hydrogen is involved. The ethane-like distances H_1–H_5 and H_1–H_6 are understandably somewhat shorter than 2·4 Å.

The environment of the molecule is shown in Fig. 7. It is apparent that the amino nitrogen atom forms three hydrogen bonds; three oxygen atoms in three

Table 4. *Intermolecular interatomic distances**

From atom (X) on molecule M	To		Equivalent contact to molecule† (2)	Distance $X\ldots Y$ (Å)	Angle $C-X\ldots Y$ (°)
	atom (Y)	on molecule (1)			
		(A) Hydrogen-bond distance			
N	O_2	$B_{0\bar{1}0}$	B	2·87	121
N	O_2	$B_{0\bar{1}1}$	B_{001}	2·81	98
N	O_3	$A_{\bar{1}00}$	A	2·79	99
O_3	O_1	$B_{0\bar{1}0}$	B	2·67	107
		(B) Hydrogen-bond contacts			
H_4(N)	O_2	$B_{0\bar{1}0}$	B	1·87	
H_5(N)	O_2	$B_{0\bar{1}1}$	B_{001}	1·82	
H_6(N)	O_3	$A_{\bar{1}00}$	A	1·80	
$H_7(O_3)$	O_1	$B_{0\bar{1}0}$	B	1·70	
		(C) Other contacts‡			
O_2	H_7	B	$B_{0\bar{1}0}$	2·81	
O_1	H_4	B	$B_{0\bar{1}0}$	2·70	
C_1	H_7	B	$B_{0\bar{1}0}$	2·47	
C_1	H_4	B	$B_{0\bar{1}0}$	2·56	
O_2	H_2	B_{001}	$B_{0\bar{1}1}$	2·82	
C_1	H_2	B_{001}	$B_{0\bar{1}1}$	2·81	
C_1	H_5	B_{001}	$B_{0\bar{1}1}$	2·67	
H_1	H_2	B_{001}	$B_{0\bar{1}1}$	2·98	
H_1	H_5	B_{001}	$B_{0\bar{1}1}$	2·78	
H_3	O_2	C_{001}	C_{001}	2·86	
H_2	H_7	$C_{0\bar{1}1}$	$C_{0\bar{1}1}$	2·82	
H_3	H_5	A	$A_{\bar{1}00}$	2·75	
C_3	H_6	A	$A_{\bar{1}00}$	2·66	
H_3	H_6	A	$A_{\bar{1}00}$	2·49	
H_7	H_6	A	$A_{\bar{1}00}$	2·59	

* See Fig. 7.

† In every case where the contact is from X on M to Y on molecule (1), the equivalent contact is from Y on M to X on molecule (2).

‡ All distances of 3·00 Å or less, excepting those listed in parts (A) and (B), are given here.

different adjacent molecules are approximately at three vertices of a regular tetrahedron centered around the nitrogen atom, with the α-carbon atom directed toward the fourth vertex, and the nitrogen–oxygen distances are 2·87, 2·81 and 2·79 Å. It was on the assumption of hydrogen bonds that the positions of the three hydrogen atoms, H_4, H_5 and H_6 of the $-NH_3^+$ group were assigned. The first two hydrogen bonds are to carboxyl oxygen atoms O_2 in molecules $B_{0\bar{1}0}$ and $B_{0\bar{1}1}$ respectively, and the third and shortest is to a hydroxyl oxygen atom O_3 in molecule $A_{\bar{1}00}$. The hydroxyl group is apparently involved in another hydrogen bond to a carboxyl oxygen atom O_1 in molecule $B_{0\bar{1}0}$; the distance to that atom from O_3 is 2·67 Å, and the angle C_3–$O_3 \cdots O_1$ ($B_{0\bar{1}0}$) is approximately tetrahedral. As has been mentioned earlier, the position assumed for H_7 is based on the assumption of a hydrogen bond.

Atom O_2 of the carboxyl group accepts two hydrogen bonds (2·87 and 2·81 Å respectively) from amino groups of molecules B and B_{001}, and atom O_1 of the carboxyl group accepts one hydrogen bond (2·67 Å) from the hydroxyl group of molecule B. Atom O_3 accepts one hydrogen bond (2·79 Å) from the amino group of molecule A.

The hydrogen bonds mentioned above, excepting the one from the amino group to the hydroxyl group, tie the molecule into infinite sheets parallel to the (010) face, the molecules in a given sheet being tied together somewhat in the sense of a distorted graphite-layer arrangement. The sheets are tied together in the third dimension by the amino to hydroxyl hydrogen bonds, which are presumably weaker than the other bonds (notwithstanding the apparently shorter distance) because the hydroxyl oxygen atom has no formal charge. This description appears to be compatible with the observation of cleavage on the (100) face.

Expected van der Waals contact distances are as follows, for comparison with the distances given in part (C) of Table 4: O ··· H, 2·6 Å; C ··· H, 2·7 or 2·8 Å; H ··· H, 2·4 Å. All of the O ··· H and H ··· H distances listed in part (C) of Table 4 exceed these expected van der Waals contact distances. Of the five C ··· H distances listed, two are definitely shorter than the van der Waals sum; however, both involve atom C_1 and hydrogens which are hydrogen bonded to the carboxyl group of which C_1 is a part, and, moreover, some adjustment in the positions of these hydrogen atoms is permissable in view of the way in which the positions were assigned.

Evidently packing in the structure is good and all contacts are substantially normal.

We are indebted to Prof. Verner Schomaker for assistance with the punched-card calculations for the Patterson function and for many useful suggestions and for stimulating and helpful discussions, and to Prof. Linus Pauling for his interest in the work and for encouragement and helpful discussions. Miss Lillian Casler and Mrs Nan Arp assisted with the computational work and the preparation of the figures. This work was aided by a grant from the National Foundation for Infantile Paralysis.

References

Albrecht, G. & Corey, R. B. (1939). *J. Amer. Chem. Soc.* **61**, 1087.

Albrecht, G., Schnakenberg, G. W., Dunn, M. S. & McCullough, J. D. (1943). *J. Phys. Chem.* **47**, 24.

Beevers, C. A. & Robertson, J. H. (1950). *Acta Cryst.* **3**, 164.

Booth, A. D. (1946). *Proc. Roy. Soc.* A, **188**, 77.

Buerger, M. J. (1951). *Acta Cryst.* **4**, 531.

Carpenter, G. B. & Donohue, J. (1950). *J. Amer. Chem. Soc.* **72**, 2315.

Cochran, W. (1948). *Acta Cryst.* **1**, 138.

Corey, R. B. (1938). *J. Amer. Chem. Soc.* **60**, 1598.

Corey, R. B. (1940). *Chem. Rev.* **26**, 227.

Corey, R. B. (1948). *Advances in Protein Chemistry*, **4**, 385.

Corey, R. B. & Donohue, J. (1950). *J. Amer. Chem. Soc.* **72**, 2899.

Donohue, J. (1950). *J. Amer. Chem. Soc.* **72**, 949.

Donohue, J. & Schomaker, V. (1949). *Acta Cryst.* **2**, 344.

Donohue, J. & Trueblood, K. T. (1952). *Acta Cryst.* **5**, 419.

Hughes, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.

Hughes, E. W. & Moore, W. J. (1949). *J. Amer. Chem. Soc.* **71**, 2618.

Internationale Tabellen zur Bestimmung von Kristallstrukturen (1935). Berlin: Borntraeger.

Levy, H. A. & Corey, R. B. (1941). *J. Amer. Chem. Soc.* **63**, 2095.

Patterson, A. L. (1935). *Z. Krystallogr.* **90**, 517.

Schomaker, V. (1947). Private communication.

Shoemaker, D. P. (1947). PhD. Thesis, California Institute of Technology.

Shoemaker, D. P., Donohue, J., Schomaker, V. & Corey, R. B. (1950). *J. Amer. Chem. Soc.* **72**, 2328.

Waser, J. (1944). PhD. Thesis, California Institute of Technology.

Wilson, A. J. C. (1942). *Nature, Lond.* **150**, 151.

Acta Cryst. (1953). **6**, 256

The Crystal Structure of 4-*iso*Propylidene–Aminophenol

By D. R. Holmes[*] and H. M. Powell

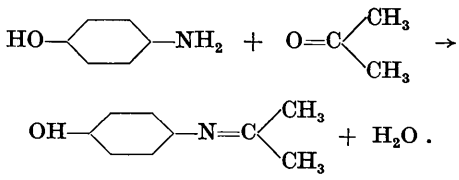
Department of Chemical Crystallography, University Museum, Oxford, England

(*Received* 25 *September* 1952)

The crystals obtained from a solution of *p*-aminophenol in acetone are shown to be a condensation compound, 4-*iso*propylidene–aminophenol (space group $P22_12_1$, $a = 5{\cdot}74$, $b = 12{\cdot}01$, $c = 12{\cdot}16$ Å). A detailed crystal-structure analysis by means of Fourier syntheses of the electron density projected perpendicular to the *a*, *b* and *c* axes is given. Bond lengths and bond angles in the molecule have been calculated and lie close to the accepted values (the standard deviation of a bond length is 0·04 Å). The molecules of 4-*iso*propylidene–aminophenol are linked into infinite chains running parallel to the *b* axis by hydrogen bonds of length 2·66 Å between the nitrogen and oxygen atoms of adjacent molecules.

Introduction

Caspari (1927) stated that *p*-aminophenol existed in two forms, an α-form crystallized from ethyl alcohol and a β-form from acetone. Analysis of samples of the two forms prepared in the way described showed that only the α-form had the appropriate empirical formula C_6H_7ON; the empirical formula of the β-form was almost exactly $C_9H_{11}ON$, indicating that the acetone had reacted with the *p*-aminophenol. The reaction is described in *Beilsteins Handbuch* (1929); condensation occurs and results in the formation of 4-*iso*propylidene–aminophenol:



The structure analysis of the condensation product is described in the present paper; a three-dimensional structure analysis of the parent substance has been given by Brown (1951).

---

[*] Now at I.C.I. Plastics Division, Welwyn Garden City, Herts., England.