The Crystal Structure of Glycyl-L-Asparagine*

By R. A. Pasternak, Lewis Katz† and Robert B. Corey

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

Crystals of glycyl-L-asparagine are orthorhombic, space group $P2_12_12_1$, $a_0 = 4\cdot81$, $b_0 = 12\cdot85$, $c_0 = 13\cdot52$ Å. Its crystal structure has been determined by three-dimensional Patterson, Fourier, and least-squares methods. The dimensions of the planar peptide amide group agree with those found in other peptides. The succinamic acid part of the molecule is nearly planar and its plane is approximately perpendicular to the plane of the peptide group. The structure is held together by a three-dimensional network of hydrogen bonds.

Introduction

The geometric problem of protein structure is mainly a matter of the configuration of the polypeptide chain; exact information concerning the arrangement of the atoms associated with the peptide link is therefore of prime importance. Even at the conclusion of the present investigation the authors were aware of only four crystal-structure determinations from which precise dimensions of the peptide link may be derived: β-glycylglycine (Hughes & Moore, 1949); N-acetylglycine (Carpenter & Donohue, 1950); α-glycylglycine (Hughes & Biswas, 1954); N,N′-diglycylcystine (Yakel & Hughes, 1952).

The present determination of the structure of glycyl-L-asparagine‡ provides additional data concerning the dimensions of the peptide bond; it also gives information concerning the dimensions and configuration of asparagine, a constituent of many proteins. Recently a cyclic structure has been proposed for asparagine by Steward & Thompson (1952); it is not confirmed for glycyl-L-asparagine by this crystal analysis.

Experimental

Unit cell and space group

Glycyl-L-asparagine was obtained in the form of needle-like crystals from water–ethanol solutions. Laue photographs showed the crystals to be orthorhombic. The needle axis was taken to be the a axis; (011) faces were predominant. The axial lengths were determined from zero-layer Weissenberg photographs taken about all three crystallographic axes with Ni-filtered Cu radiation. The unit cell dimensions, with their limits of error, were found to be

$$a_0 = 4\cdot81\pm0\cdot02,\ b_0 = 12\cdot85\pm0\cdot05,\ c_0 = 13\cdot52\pm0\cdot05\ \text{Å},$$

and the density (by flotation) $1\cdot506\pm0\cdot008$ g.cm.$^{-3}$. The unit cell contains 4 molecules (calculated 4·01).

The only systematic absences observed were the $(h00)$, $(0k0)$, and $(00l)$ reflections when h, k or l (respectively) was odd, establishing the space group as D_2^4–$P2_12_12_1$. The range of observation was 0–6 in h, 0–16 in k, and 0–17 in l.

Intensity data

Photographs were taken around all three crystallographic axes. For those around the a axis a needle crystal with a cross section about 0·3 mm. on edge was used; for those around the b and c axes specimens having a maximum dimension of about 0·4 mm. were selected from fragments obtained by cleaving the needle crystals parallel to the (100) plane. Weissenberg equi-inclination photographs were taken for all layer lines which required inclination angles smaller than 30°. About 1000 reflections, practically all those accessible to Cu $K\alpha$ radiation, were recorded. The intensities were estimated visually from multiple films with the help of intensity strips prepared from the same crystals. Independent estimates were made by two observers; in general, they agreed within 10%. The intensities were corrected for Lorentz and polarization factors; no absorption corrections were applied. The corrected intensities from each layer were multiplied by an appropriate factor in order to bring all intensity data to the same relative scale. When independent values for the corrected intensities of individual reflections as estimated from photographs taken around different axes were compared, it was found that they declined with increasing order in a different manner. Since the photographs taken around the short a axis were of superior quality, and far more intensities could be estimated on a single layer line than for photographs around other axes, they were used as the source of most of the intensity data.

A temperature factor and a scale factor for converting the observed values of $|F_{\text{rel.}}|^2$ approximately

to the absolute scale were calculated by Wilson's (1942) method.

Trial structure

In the molecule of glycylasparagine,

$$\overset{*}{NH_3^+}-CH_2-CO-NH-\overset{|}{CH}-CH_2-CO-NH_2,$$
$$\underset{COO^-}{|}$$
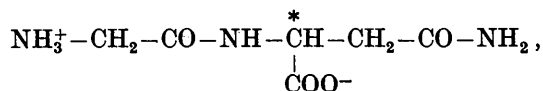
the carboxyl group and its α-carbon atom $(C-CO_2^-)$, the terminal amide group and its α-carbon atom $(C-CO-NH_2)$, and the peptide amide group with its adjacent carbon atoms $(C-CO-NH-C)$, may all be expected to be planar groups, either because of the presence of double bonds or because of resonance. However, free rotation around the single bonds not involved in resonance, especially those formed by the asymmetric carbon atom, probably would permit these planar groups to assume a great variety of spatial arrangements relative to one another, a consideration which made the choice of a probable configuration for the molecule extremely uncertain. Nevertheless, the outstanding (004) reflection, with a unitary structure factor of about 0·4, and the much weaker (002) reflection, together with considerations of molecular dimensions and packing, indicated that the molecule was elongated in the y direction and lay near $z = \frac{1}{4}$. Because of the presence of several other conspicuously strong ($0kl$) reflections, trial-and-error methods for deriving a more detailed picture of the molecule seemed to be promising. Prolonged efforts to obtain a plausible trial structure by these means were, however, unsuccessful.

The use of Patterson vector diagrams proved to be more fruitful. A three-dimensional Patterson function sharpened and modified as described elsewhere (Patterson, 1935; Shoemaker et al., 1950; Donohue & Trueblood, 1952a, b) was prepared*. This Patterson

function showed only two conspicuous non-equivalent peaks at distances short enough to correspond to vectors between bonded atoms. The heights of both these peaks were so great as to indicate that they contained several interactions. The positions of these peaks with respect to the axes of the vector diagram may be conveniently described by reference to Fig. 1. This figure represents a cube constructed about the origin as center and with its edges parallel to the axes of the vector diagram. One set of vectors, contained in peak A and its equivalent peak, points along the v direction; and the other set, contained in peak B and its seven equivalent peaks, falls approximately along the body diagonals of the cube. The coplanar vectors A, B, and B' approximate in both their lengths and their arrangement the bonds between the four atoms associated with the amide and the carboxyl groups. Additional peaks are also found in the vector diagram which can be attributed to the vectors relating the non-bonded atoms of this grouping.

Thus the orientation of at least one planar group of four atoms was derived with a fair degree of probability. It seemed that its absolute position might well be found from the three Harker sections, especially as the one vector parallel to v was expected to give rise to vector triples on the two sections at $u = \frac{1}{2}$ and $w = \frac{1}{2}$ (Albrecht & Corey, 1939). These Harker sections are shown in Figs. 2 and 3, respectively. As the modification of the Patterson function removes the

Fig. 2. An asymmetric unit of the Harker section $u = \frac{1}{2}$. The first contour is drawn at $P = 0$, others at arbitrary intervals. Black dots represent the positions of Harker interactions.
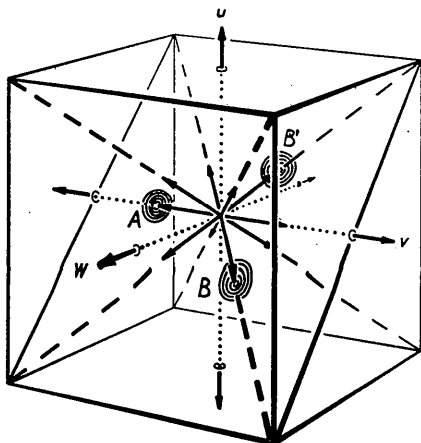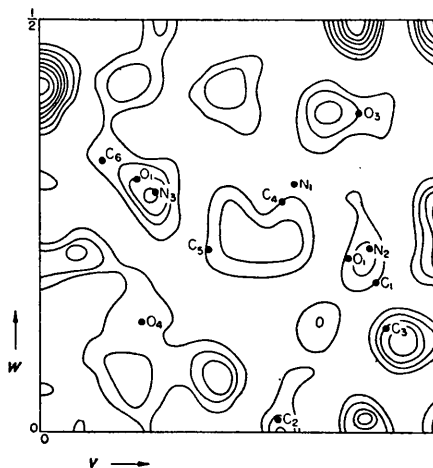

Fig. 1. A diagrammatic representation of the conspicuous vectors near the origin of the Patterson diagram.

* The calculations for the Patterson diagram and all other major calculations were carried out by means of I.B.M. machines.
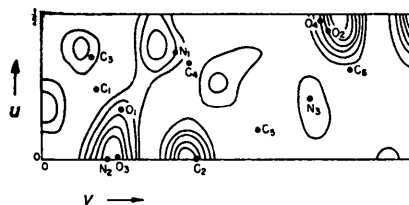

Fig. 3. An asymmetric unit of the Harker section $w = \frac{1}{2}$.

term $|F_{000}|^2$, the vector diagram contains negative and positive regions, the integral of the function over the unit cell being zero.

In our attempt to derive atomic positions from the Harker sections we applied the criterion which was found to be useful in a similar investigation (Donohue & Trueblood, 1952a, b), namely, the requirement that all interatomic vectors fall in or near positive regions of the diagram. Seventeen positions for the four-atom arrangement were found which satisfied this condition on all three Harker sections, giving seventeen sets of coordinates for a possible carboxyl or amide group. The probability of each of these sets of atomic positions was evaluated by calculating the corresponding Patterson interactions and comparing them with the experimental Patterson diagram. To our disappointment, no set fitted satisfactorily; however, three atoms of one set exhibited excellent agreement with our criterion. We therefore felt that at least these three positions might represent the positions of real atoms, although they need not necessarily belong to a planar four-atom arrangement. They were therefore selected as a starting point for finding the positions of other atoms by means of the vector convergence method (Beevers & Robertson, 1950; Shoemaker et al., 1953). Two of the three atoms and their equivalent positions were used for the superposition of the Patterson function, and about 20 prominent eightfold overlaps per asymmetric unit, representing 20 possible atomic positions, were obtained; 13 would be expected to arise from the 13 atoms (excluding hydrogen atoms) in the molecule. The significance of all these positions was tested by inserting the interactions calculated from them into the vector diagram. Eight of these positions, including the two originally chosen, gave interactions virtually all of which fell within positive regions of the vector diagram. These eight atoms formed a continuous, branched chain with reasonably good bond lengths and bond angles. This framework suggested almost unambiguously where to place the remaining five atoms of the molecule, and inspection showed that minor eightfold overlaps, or at least sevenfold overlaps, were actually present near these anticipated atomic positions. From these overlaps coordinates were derived which also gave acceptable Patterson interactions. A complete trial structure was thus obtained.

Our confidence in the correctness of the trial structure was strengthened by the calculation of intra- and intermolecular interatomic distances; all of them, in particular the lengths of the hydrogen bonds, were found to be of the right magnitude. The calculated and observed structure factors for the (0kl) reflections were found to be in fairly good agreement, except for the higher orders.

It is instructive that although the peaks near the origin (Fig. 1) were indentified correctly with vectors of the amide and carboxyl groups the two atoms C_2 and C_3 whose positions were used for the vector

convergence method were not atoms of these groups at all but were members of the carbon chain. The reason for this can be understood by reference to Figs. 2 and 3, in which the Harker interactions of all atoms, based on the final parameters, are plotted as black dots on the Harker sections $u = \frac{1}{2}$ and $w = \frac{1}{2}$. On the former, our criterion that all interactions must fall on or near positive regions is satisfied. On the latter, however, four interactions lie in negative regions; three of them (C_1, C_4, and C_5) belong unfortunately to the three different planar groups. It is therefore not surprising that the seventeen possible positions derived for the planar groups from the Harker sections did not include the correct ones.

We believe that the difficulties that we encountered in the interpretation of the Patterson diagram, and especially of the Harker sections, are not peculiar to the structure under investigation, but are of a general occurrence with complicated structures containing no heavy atoms. If N is the number of atoms in the unit cell, then on a Harker section the number of Harker peaks per unit area, and therefore their contribution to the density of the Patterson function, is approximately proportional to $N^{\frac{1}{2}}$. The contribution of the non-Harker peaks to the density of the Patterson function increases much faster; namely, about proportional to N. Thus, with increasing complexity of a Harker section its principal features are more and more determined by the non-Harker peaks and the identification and localization of specific Harker interactions become more and more difficult.

Refinement of parameters

In the trial structure as derived from the three-dimensional Patterson diagram the atoms were distributed so that little overlapping was to be anticipated in a Fourier projection along the short a axis of the crystal. A first projection, which included about one-third of the observed (0kl) structure factors, was therefore made on (100); 12 of the 13 atoms were indeed resolved and the position of the 13th was clearly indicated. Further refinement of the y and z parameters proceeded rapidly by means of successive Fourier projections. In the third projection virtually all of the 210 (0kl) reflections were included, and no structure factor of importance changed sign, even when back-shift corrections (Booth, 1946) were applied to the atomic parameters; the discrepancy factor $R = \Sigma||F_o| - |F_c|| \div \Sigma|F_o|$ calculated without the inclusion of the hydrogen atoms was 0·19.

Preliminary x parameters had already been obtained from the Patterson function; they were modified in accordance with the refined y and z parameters and bond lengths found in similar compounds. The value of R calculated from them for the (h0l) and (hk0) reflections showed that they were still only very approximate. Unfortunately the centrosymmetric Fourier projections along the b and c axes were useless for their

refinement because of lack of resolution of the atoms. Refinement of the x parameters alone was therefore carried out by the method of least squares; since the y and z parameters were so much better than the x parameters, they were taken to be constant. In the first two refinements the complete three-dimensional data were used but only the diagonal terms in the normal equations were evaluated. In the three subsequent refinements about 300 reflections were selected and all terms of the normal equations were calculated. At the end of these five stages of refinement the values of R had dropped to 0·25 for the ($h0l$) and ($hk0$) zones, and to 0·20 for the general reflections. These values, indicated that the x parameters were now nearly comparable in accuracy with the y and z parameters, and that further refinement should be carried out simultaneously on all parameters. To this end a three-dimensional Fourier synthesis was prepared. Fig. 4
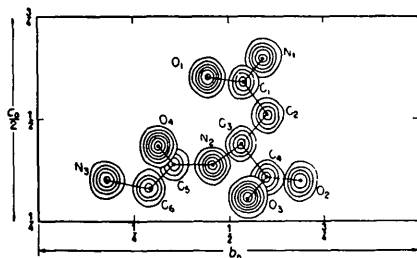

Fig. 4. A representation of the results of the three-dimensional Fourier synthesis. Sections passing approximately through the centers of the atoms are projected on (100). Contour intervals of about 1·5 e.Å⁻³; first contour at 1·5 e.Å⁻³.

is a two-dimensional representation of the results of this synthesis, in which sections passing approximately through the centers of the atoms are projected on (100). The peaks are slightly elongated in the direction of the c axis, especially the one representing oxygen atom O_2; this peak is also distinguished by a surprisingly low electron density when compared with the peaks representing the other oxygen atoms. In general, the background is low and uniform with variations less than $\pm\frac{1}{2}$e.Å⁻³. Some additional peaks of about $\frac{1}{2}$–1 e.Å⁻³ are situated near positions which hydrogen atoms might be expected to occupy.

Parameters for the carbon, nitrogen, and oxygen atoms were derived from the Fourier peaks by the analytical method described by Shoemaker et al. (1950). They were corrected for termination-of-series errors in the usual way (Booth, 1946). Finally, the n-shift rule, with $n = 1·7$ (Shoemaker et al., 1950), was applied to the parameter shifts, a procedure which corrects for the incomplete convergence of a non-centrosymmetric Fourier synthesis. The average parameter changes given by the three-dimensional Fourier synthesis were about 0·015 Å in each of the three axial directions; a few atoms, however, moved as much as 0·05 Å.

Further refinement of the parameters was carried out by means of a least-squares treatment in which

all data were used, and only the diagonal terms of the normal equations were evaluated. Some significant shifts were again obtained, mostly in the x parameters. In general, they were in the same directions as those indicated by the preceding Fourier synthesis (see Table 1). In the subsequent calculation of structure factors the hydrogen atoms were included for reflections with sin $\theta < 0·55$; some prism reflections were thus significantly improved. Parameters were assigned to the hydrogen atoms (Table 2) by assuming the bond lengths C–H = 1·1 Å and N–H = 1·0 Å, and bond angles in accord with the probable trigonal or tetrahedral arrangement of the bonds. At this stage a redetermination of the scale factor and the temperature factor constant was made, which gave a value $B = 3·3 \times 10^{-16}$ Å² for the latter.

Although the refinement of the parameters had apparently converged and the contribution of the hydrogen atoms had been taken into account, the agreement between the observed and calculated structure factors and the value of R (0·175) still were not as good as might be expected in view of the quality of the experimental data. Closer scrutiny of the observed and calculated F's revealed discrepancies which varied systematically with the order of the reflections. The elongation of the Fourier peaks (Fig. 4) in the direction of the c axis, together with these systematic discrepancies, indicated anisotropic temperature effects within the crystal; and the extension of the molecule of glycylasparagine in directions about parallel to the b and to the c axes of the crystal made it reasonable to choose the axes of the crystal as axes of temperature anisotropy. The temperature factor therefore takes the form

$$T = \exp(-\alpha h^2/4a^2 - \beta k^2/4b^2 - \gamma l^2/4c^2) :$$

The constants β and γ were determined independently for each layer line of constant h by an analytical method similar to the one used for finding isotropic temperature factor constants, and average values $\beta = 2·3 \times 10^{-16}$ Å² and $\gamma = 4·0 \times 10^{-16}$ Å² were obtained. For the short a axis, the constant α was not derived from the data but was taken arbitrarily to be equal to γ, and individual scale factor corrections for each layer line were introduced.

By means of the application of these anisotropic temperature factors the value of R for the ($0kl$) zone was reduced to 0·146 and that for all reflections to 0·163, and all systematic discrepancies appeared to have been eliminated.

The possibility had to be considered that the introduction of the anisotropic temperature factors might cause significant changes in the parameters of the atoms. In order to investigate this matter a three-dimensional difference Fourier synthesis was carried out. The difference Fourier did indeed indicate small changes in the parameters, the largest being 0·025 Å, including the n-shift. Only the shifts of 0·01 Å or more, which occurred in the case of five atoms, were consi-

Table 1. *Atomic coordinates for one molecule*

Atom	(1) Fourier parameters			(2) Least-squares parameters			(3) Difference Fourier parameters		
	x	y	z	x	y	z	x	y	z
C_1	0·124	0·5374	0·5892	0·119	0·5377	0·5908	0·121$_5$	0·5377	0·5899
C_2	0·252$_5$	0·6014	0·5088	0·251	0·6022	0·5091	0·251	0·6013	0·5084
C_3	0·423$_5$	0·5323	0·4388	0·426	0·5327	0·4383	0·426	0·5327	0·4383
C_4	0·583$_5$	0·5981	0·3580	0·588	0·5989	0·3604	0·588	0·5989	0·3604
C_5	0·297$_5$	0·3562	0·3903	0·297$_5$	0·3560	0·3891	0·297$_5$	0·3560	0·3891
C_6	0·090$_5$	0·2917	0·3329	0·096	0·2891	0·3330	0·096	0·2891	0·3330
N_1	−0·059$_5$	0·5894	0·6495	−0·068$_5$	0·5896	0·6481	−0·072	0·5896	0·6481
N_2	0·248	0·4576	0·3881	0·251	0·4573	0·3876	0·251	0·4573	0·3876
N_3	0·144$_5$	0·1788	0·3526	0·144$_5$	0·1776	0·3528	0·144$_5$	0·1787	0·3528
O_1	0·167	0·4465	0·6026	0·166	0·4456	0·6044	0·166	0·4456	0·6044
O_2	0·524	0·6878	0·3484	0·533	0·6895	0·3491	0·533	0·6895	0·3491
O_3	0·750$_5$	0·5496	0·3069	0·756$_5$	0·5499	0·3063	0·756$_5$	0·5499	0·3063
O_4	0·486$_5$	0·3158	0·4350	0·488	0·3156	0·4332	0·489$_5$	0·3151	0·4347

dered to be significant and were applied to the parameters. These shifts had, in general, little effect upon the interatomic distances and bond angles of the molecule.

Table 1 lists the parameters derived from (1) the three-dimensional Fourier synthesis, (2) the least-squares refinement, and (3) the difference Fourier synthesis. The changes in the y and z parameters are small, the average difference between parameters (1) and (3) being $\Delta y = 0·008$ Å and $\Delta z = 0·011$ Å. The x parameters, however, show a corresponding average shift of $\Delta x = 0·018$ Å and atoms N_1 and O_2 move more than 0·06 Å and 0·04 Å, respectively, in the x direction. This difference in the behavior of the y and z parameters on the one hand and the x parameters on the other is to be ascribed to the fact that the former had been nearly completely refined by the Fourier projections prior to the three-dimensional Fourier synthesis, whereas the latter had not yet converged to the same degree even at the completion of the Fourier synthesis.

A critical analysis of the changes in the parameters shows clearly that they have converged. Of the five atoms which are significantly shifted by the difference Fourier synthesis all except N_1 move back to positions near those found from the three-dimensional Fourier synthesis. We believe that the parameters derived from the difference Fourier are the most reliable, and for this reason our subsequent discussion of the structure is based on them.

In the difference Fourier a variation of density is found in the neighborhood of most atoms. This is not surprising because the same anisotropic temperature factors were applied to all atoms. Oxygen atom O_2 in particular seems to have a much stronger temperature anisotropy than the other atoms. The difference Fourier indicates that the direction of its maximum thermal vibration is approximately perpendicular to the plane of the carboxyl group of which it is a part, and also perpendicular to its one hydrogen bond. An analogous temperature anisotropy seems to exist for

a similarly bonded oxygen atom in L-glutamine (Cochran & Penfold, 1952). Undoubtedly the discrepancy factor R could be reduced by an assignment of individual anisotropic temperature factors to particular atoms. Since no further refinement of the parameters was intended, a procedure of this sort did not appear to be worthwhile, and the structure factors were not recalculated after the difference Fourier synthesis. The standard errors, calculated in the usual way (Shoemaker *et al.*, 1950), are therefore based on the least-squares parameters. The average standard error is 0·014 Å for an interatomic distance and $\frac{3}{4}°$ for bond angles near 110° between bonds of a length of about 1·5 Å.

The results of the difference Fourier synthesis also provided some experimental information about the positions of the hydrogen atoms. In the calculation of the difference Fourier the contribution of the hydrogen atoms to the values of F_c had purposely been omitted, so that their positions would be indicated by peaks in the difference Fourier. In a non-centrosymmetric structure hydrogen peaks of a height of about $\frac{1}{2}$ e.Å^{-3} are to be expected. Such peaks were indeed found. The peaks that represented hydrogen atoms attached to carbon atoms were fairly sharp and near the positions where the atoms had been assumed to be. No peak was found for the hydrogen atom attached to the peptide nitrogen atom, N_2. Peaks near the terminal amide nitrogen atom, N_1, were too poorly defined to give any definite information about the positions of the two hydrogen atoms. Three fairly well defined peaks appeared around the amino nitrogen atom, N_3. They were displaced from the positions originally assigned to the hydrogen atoms in a manner which seemed to be significant. A displacement of this sort was not surprising because of the possibility of rotation of the amino group NH_3^+ around the C–N single bond. The arrangement of the hydrogen atoms associated with this group will be discussed later in more detail. The assumed parameters of the hydrogen atoms used for the structure-factor calculations and those

Table 2. *Coordinates of the hydrogen atoms in one molecule*

	(1) Assumed parameters			(2) Difference Fourier parameters		
	x	y	z	x	y	z
H_1 (C_2)	0·387	0·660	0·543	0·39	0·66	0·54
H_2 (C_2)	0·087	0·640	0·467	0·08	0·64	0·46
H_3 (C_3)	0·590	0·495	0·481	0·58	0·48	0·48
H_4 (C_6)	0·122	0·303	0·254	0·13	0·31	0·27
H_5 (C_6)	−0·115	0·309	0·356	−0·13	0·30	0·36
H_6 (N_1)	−0·170	0·552	0·703		Diffuse	
H_7 (N_1)	−0·105	0·666	0·636		Diffuse	
H_8 (N_2)	0·080	0·486	0·353		No peak	
H_9 (N_3)	0·181	0·139	0·289	0·10	0·13	0·30
$H_{10}(N_3)$	−0·024	0·147	0·386	0·08	0·14	0·41
$H_{11}(N_3)$	0·311	0·169	0·398	0·30	0·17	0·38

estimated from the peaks in the difference Fourier plot are given in Table 2.

In Table 3 the observed structure factors and those calculated from the least-squares parameters are listed for the (0kl) reflections; data for the other reflections are omitted for the sake of economy.* The calculated

---

* A list of the observed and calculated structure factors for reflections other than (0kl), comprising 17 pages of typescript, has been deposited as Document No. 4105 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., U.S.A. A copy may be secured by citing the Document number and by remitting \$2·50 for photoprints, or \$1·75 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

structure factors contain the contribution of hydrogen atoms for reflections with sin $\theta < 0.55$.

Data for the individual layer lines may be summarized as follows:

	No. of reflections	$\bar{F}_o$	R
0kl	187	9·5	0·146
1kl	215	9·9	0·123
2kl	209	7·8	0·177
3kl	177	6·1	0·181
4kl	121	4·4	0·237
5kl	61	3·4	0·282
6kl	12	3·1	0·243
hkl	982	7·5	0·163

With increasing values of h the average value of F_o decreases, as might be expected, and the discrepancy factor R increases correspondingly. The pronounced rise in R at $h = 4$ might be attributed to the procedure by which the intensities were collected. Whereas most of the data for $h = 0$ to $h = 3$ were obtained from the same crystal rotated around a, the data for the higher layer lines had to be collected from two crystals rotated around b and c, respectively; they are therefore less consistent internally.

The observed values of the structure factors for (004) and (112), the two strongest reflections, are conspicuously smaller than the calculated values, the difference probably being due to extinction. These planes were therefore omitted in the calculation of R,

Table 3. *Observed and calculated structure factors for the (0kl) reflections*

k	l	F_o	F_c	k	l	F_o	F_c	k	l	F_o	F_c
0	2	17·9	16·9	2	10	20·7	23·9	4	11	—	− 3·5
0	4	88·1	−114·7	2	11	2·3	1·9	4	12	—	1·4
0	6	15·4	14·8	2	12	4·7	− 5·8	4	13	2·8	1·4
0	8	8·0	7·9	2	13	7·4	5·7	4	14	4·5	− 4·6
0	10	3·9	− 2·9	2	14	3·2	− 4·4	4	15	9·0	8·1
0	12	2·9	5·1	2	15	—	− 1·1	5	1	17·4	15·6
0	14	4·8	5·3	2	16	2·7	2·0	5	2	38·5	36·6
0	16	1·9	− 0·3	2	17	2·8	− 0·8	5	3	9·8	−10·7
1	1	17·7	−18·7	3	1	41·1	39·8	5	4	4·7	2·8
1	2	24·8	26·2	3	2	27·1	−26·4	5	5	10·2	14·2
1	3	10·5	7·3	3	3	15·9	−14·5	5	6	12·2	8·3
1	4	10·4	−11·0	3	4	19·5	−20·1	5	7	16·4	−17·0
1	5	30·6	32·3	3	5	39·0	38·8	5	8	8·6	11·9
1	6	18·7	18·9	3	6	10·6	11·0	5	9	10·0	−12·4
1	7	16·1	19·0	3	7	6·2	3·4	5	10	7·7	− 8·5
1	8	1·5	2·7	3	8	7·7	4·5	5	11	—	− 0·4
1	9	1·4	− 0·6	3	9	19·8	−18·9	5	12	2·1	− 4·1
1	10	3·0	− 3·9	3	10	6·3	6·5	5	13	2·5	2·5
1	11	—	− 3·3	3	11	—	− 0·6	5	14	—	3·1
1	12	1·3	− 3·1	3	12	1·3	0·9	5	15	6·6	5·5
1	13	2·1	3·6	3	13	11·8	9·6	5	16	1·6	− 0·4
1	14	13·6	−11·2	3	14	5·5	− 5·1	6	0	12·4	15·0
1	15	4·2	3·3	3	15	8·3	6·8	6	1	12·5	−14·0
1	16	3·3	− 1·5	3	16	4·5	2·0	6	2	5·8	8·0
1	17	3·1	− 3·6	4	0	9·3	−10·0	6	3	2·1	3·4
2	0	38·5	44·1	4	1	6·1	− 6·8	6	4	8·3	− 8·7
2	1	12·8	−14·2	4	2	7·8	− 5·9	6	5	7·3	9·1
2	2	17·3	16·7	4	3	4·2	− 5·1	6	6	12·6	−11·2
2	3	21·7	20·9	4	4	1·7	3·2	6	7	5·8	− 7·0
2	4	7·5	− 6·2	4	5	39·0	−38·8	6	8	8·4	− 9·6
2	5	21·5	24·1	4	6	18·1	−14·9	6	9	—	− 2·0
2	6	21·1	−21·6	4	7	9·6	− 5·8	6	10	5·4	− 7·6
2	7	7·9	− 6·4	4	8	17·2	−17·5	6	11	4·6	− 4·8
2	8	2·3	3·4	4	9	21·6	26·7	6	12	—	0·6
2	9	2·5	− 3·4	4	10	3·0	4·7	6	13	3·0	2·3

Table 3 (cont.)

k	l	F_o	F_c		k	l	F_o	F_c		k	l	F_o	F_c
7	1	27·4	−28·4		9	6	11·8	−14·7		12	1	5·4	− 4·0
7	2	2·9	− 2·0		9	7	21·6	−24·1		12	2	—	− 0·7
7	3	6·6	2·8		9	8	2·9	1·9		12	3	—	− 0·5
7	4	17·1	17·8		9	9	15·1	−10·8		12	4	6·2	4·6
7	5	15·0	18·2		9	10	—	0		12	5	14·1	12·9
7	6	8·7	−11·1		9	11	3·4	4·4		12	6	9·5	7·7
7	7	20·6	−18·6		9	12	4·0	− 2·3		12	7	—	0·8
7	8	9·1	− 9·1		9	13	3·7	1·6		12	8	5·7	5·2
7	9	7·2	− 8·2		9	14	4·6	4·1		12	9	8·6	− 8·6
7	10	5·7	7·0		10	0	—	− 1·2		12	10	9·2	− 8·5
7	11	8·0	4·7		10	1	10·2	11·3		12	11	2·3	3·4
7	12	4·2	3·4		10	2	0·9	0·5		13	1	13·6	12·5
7	13	4·6	− 2·7		10	3	2·9	1·7		13	2	11·3	− 9·2
7	14	2·3	0·3		10	4	5·3	4·9		13	3	4·9	− 4·8
7	15	1·8	− 1·1		10	5	20·9	−18·0		13	4	—	2·5
8	0	51·5	−56·8		10	6	25·3	22·8		13	5	4·1	− 6·1
8	1	8·7	−10·2		10	7	7·5	− 5·3		13	6	4·0	− 2·8
8	2	5·5	7·1		10	8	3·1	3·0		13	7	8·5	9·3
8	3	10·2	9·6		10	9	3·5	4·1		13	8	4·5	− 4·9
8	4	24·7	27·6		10	10	20·4	−17·7		13	9	4·3	4·2
8	5	—	2·6		10	11	4·2	3·6		13	10	2·4	2·3
8	6	—	− 0·9		10	12	6·2	5·6		14	0	16·7	−18·8
8	7	18·9	−21·4		10	13	6·5	5·0		14	1	5·6	6·3
8	8	7·0	7·1		11	1	10·4	13·4		14	2	7·4	− 7·1
8	9	—	− 1·4		11	2	3·7	− 4·6		14	3	—	− 1·6
8	10	3·5	3·6		11	3	1·9	1·3		14	4	4·3	3·8
8	11	8·5	7·0		11	4	11·3	12·9		15	1	4·7	3·2
8	12	1·7	− 2·2		11	5	10·1	− 8·3		15	2	2·2	2·6
8	13	2·9	0·5		11	6	—	− 0·8		15	3	0·7	2·5
8	14	—	− 1·7		11	7	—	3·2		15	4	3·5	− 3·4
8	15	1·5	− 1·8		11	8	4·6	− 4·5		15	5	6·2	− 7·8
9	1	1·5	− 1·0		11	9	6·3	4·7		15	6	1·2	− 0·9
9	2	5·4	3·1		11	10	5·0	− 3·6		16	0	5·5	− 5·1
9	3	13·7	14·9		11	11	—	− 1·0		16	1	6·0	6·6
9	4	—	− 0·4		11	12	2·5	− 3·1		16	2	2·1	− 2·3
9	5	2·9	2·1		12	0	14·8	−14·5					

as were reflections with observed intensity equal to zero.

Discussion of the structure

Configuration of the molecule

The intramolecular bond lengths and bond angles are listed in Table 4; they are also shown in Fig. 5,

Table 4. *Intramolecular bond distances and bond angles*

Distance (Å)		Angles (°)	
C_1–C_2	1·506	O_1–C_1–N_1	119·6
C_2–C_3	1·544	O_1–C_1–C_2	124·9
C_3–C_4	1·563	N_1–C_1–C_2	115·5
C_5–C_6	1·502	C_1–C_2–C_3	111·3
		C_2–C_3–C_4	112·0
C_1–N_1	1·388	C_4–C_3–N_2	109·5
C_3–N_2	1·455	C_3–C_4–O_2	120·1
C_5–N_2	1·320	C_3–C_4–O_3	116·1
C_6–N_3	1·463	O_2–C_4–O_3	123·5
		C_3–N_2–C_5	123·5
C_1–O_1	1·218	N_2–C_5–O_4	123·9
C_4–O_2	1·204	N_2–C_5–C_6	116·6
C_4–O_3	1·260	O_4–C_5–C_6	119·6
C_5–O_4	1·227	C_5–C_6–N_3	111·1

which is a schematic drawing of the molecule viewed along the *a* axis of the crystal. For purposes of discussion the molecule may be considered to consist of three
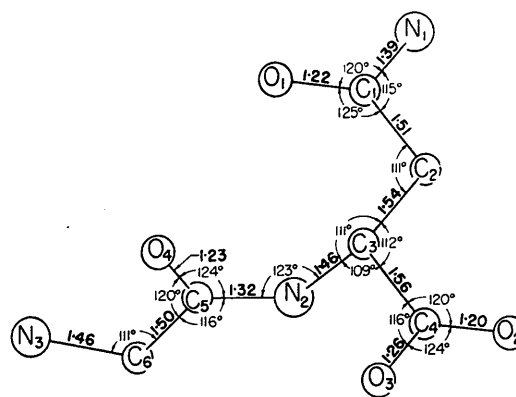

Fig. 5. A schematic drawing of the molecule viewed along the *a* axis of the crystal. Bond lengths are given in Ångström units.

simple groups of atoms: the peptide group and the amino nitrogen atom attached to it, the terminal amide group, and the carboxyl group. The characteristics of these groups will be considered separately.

The peptide group has the *trans* configuration. The five atoms C_3, N_2, C_5, C_6, and O_4 are coplanar within the accuracy of this determination. The maximum distance of any one of these atoms from the median plane of the group, as determined by least-squares

Table 5. *Bond lengths and bond angles of the peptide group found in crystalline peptides*

	Bond lengths (Å)				Bond angles (°)			
	N–C	C'–C	C'–O	C'–N	C–C'–O	N–C'–O	C–C'–N	C'–N–C
N-Acetylglycine (a)	1·45	1·50	1·24	1·32	121·0	121·3	117·7	119·6
β-Glycylglycine (b)	1·48	1·53	1·23	1·29	121·0	125·0	114·0	122·0
α-Glycylglycine (c)	1·47	1·56	1·24	1·32	121·1	124·2	114·4	119·3
N,N'-Diglycylcystine (d)	1·48	1·56	1·21	1·35	120·6	125·3	113·2	121·6
Glycyl-L-asparagine	1·46	1·50	1·23	1·32	119·6	123·9	116·6	123·5

(a) Carpenter & Donohue, 1950. (b) Hughes & Moore, 1949. Two-dimensional data only. (c) Hughes & Biswas, 1954.
(d) Yakel & Hughes, 1952.

treatment of the atomic coordinates, is less than 0·01Å. Probably the hydrogen atom attached to the nitrogen atom is also coplanar with the other atoms of the group. The dimensions of the peptide group in glycylasparagine are in good agreement with those of other simple peptides and related compounds. In Table 5 the bond lengths and bond angles obtained in this investigation are listed together with those derived from previous investigations. In all of these structures the C–N peptide bond shows resonance shortening of about 0·15 Å, and the bond angles around the carbon atom are those to be expected: angle N–C'–O is the largest, angle C–C'–N the smallest. In all but one of these structures the peptide group is reported to be planar within the limits of error of the determination. In the one exception, diglycylcystine, a rotation of only 6° around the C–N bond was found. The resonance shortening of the C–N peptide bond and the associated planarity of the peptide group have been discussed elsewhere (Corey & Pauling, 1953).

The amino nitrogen atom, N_3, also lies near the plane of the peptide group so that it approaches closely (2·68 Å) the oxygen atom O_4; the C_6–N_3 bond forms an angle of about 7° with this plane. The length of the C_6–N_3 bond, 1·46 Å, is slightly but probably not significantly shorter than the average of the lengths found for corresponding bonds in investigations of comparable accuracy (threonine, 1·49 Å; alanine, 1·50 Å; diglycylcystine, 1·46 Å; serine, 1·49 Å).

Regarding the amino group as a whole, there is good evidence that it is present as a charged NH_3^+ group. The arrangement of the oxygen atoms around the nitrogen atom suggests the formation of three hydrogen bonds, and in the difference Fourier the three peaks near the position of the nitrogen atom indicate the presence of three hydrogen atoms. Further evidence for the dipolar configuration of the molecule is furnished by the dimensions of the carboxyl group. One would expect that the presence of a hydrogen atom covalently bonded to a carboxyl oxygen atom would be revealed by a C–O distance close to 1·31 Å, 0·05 Å longer than observed.

In the terminal amide group the four atoms O_1, C_1, N_1, and C_2 are coplanar, the maximum deviation from the best median plane being 0·015 Å. This group, like the peptide amide group, shows resonance shortening of the C–N bond, though to a lesser degree; the C–N bond length (1·39 Å) corresponds to about 20%

double-bond character. In view of this small amount of double-bond character of the C–N bond, the angles adjacent to the C=O double bond would be expected to be nearly equal, with angle O_1–C_1–C_2 slightly smaller than the other. Actually, a difference of 5° in the opposite sense is observed. This difference must probably be accepted as significant, although no satisfactory explanation is apparent.

The structures of only two other compounds containing the terminal amide group have been published. One paper describes a two-dimensional determination of the crystal structure of glutamine (Cochran & Penfold, 1952). The dimensions found for the amide group in this molecule are surprising: the length of the C–N bond (1·28 Å) indicates a large amount of double-bond character (about 80%), contrasting with the nearly equal bond angles around the carbon atom. The other reported determination is that of the structure of acetamide (Senti & Harker, 1940), which has clarified the configuration of the compound but has not furnished accurate dimensions of the terminal amide group.

In glycylasparagine the atoms of the carboxyl group and the α-carbon atom are planar within the accuracy of the determination; the maximum deviation for any atom from the best median plane is 0·03 Å. The group is asymmetric, the lengths of the C–O bonds being 1·26 and 1·22 Å, respectively. Although this difference in bond lengths does not exceed twice the limit of error of the determination, it is, probably significant. Oxygen atom O_2 of the shorter C–O bond is participating in only one hydrogen bond, whereas O_3 takes part in three. Probably this distribution of hydrogen bonds causes the resonance structure with the double bond between the carbon atom and O_2 to be slightly more stable than the other. The magnitudes of the angles around the carbon atom are also in accord with this explanation, angle C_3–C_4–O_2 being larger than angle C_3–C_4–O_3. This sort of dissymmetry of the carboxyl group has been previously reported and discussed for other amino acids and related compounds (alanine, Levy & Corey, 1941; Donohue, 1950; β-glycylglycine, Hughes & Moore, 1949; N-acetylglycine, Carpenter & Donohue, 1950; hydroxy-L-proline, Donohue & Trueblood, 1952a, b). In contrast, Cochran & Penfold (1952) report that in glutamine the bond angles of the carboxyl group are symmetrical and that the shorter C–O bond involves

the oxygen to which the greater number of hydrogen bonds are directed.

The planes of the terminal amide group and the carboxyl group are nearly coincident, the angle between them being only 3°. The carbon atoms of these groups, C_1, C_2, C_3, C_4, form an extended chain in which the C–C bond lengths probably do not deviate significantly from the average of 1·54 Å. It is interesting that this planarity of the succinamic acid part of the molecule fits well into the steric pattern observed by MacGillavry, Hoogschagen & Sixma (1948) and by Morrison & Robertson (1949) for the related dicarboxylic acids succinic acid, adipic acid, and sebacic acid, all of which contain an even number of carbon atoms and are nearly planar. Glutamine (Cochran & Penfold, 1952), glutamic acid (Dawson, 1953), glutaric acid (Morrison & Robertson, 1949), and pimelic acid (MacGillavry et al., 1948), which contain an odd number of carbon atoms, are all nonplanar.

The planar succinamic acid part of the molecule and the planar peptide group, approximately perpendi-cular (85°) to each other, are easily recognized in the scale drawing of the molecule (Fig. 6).
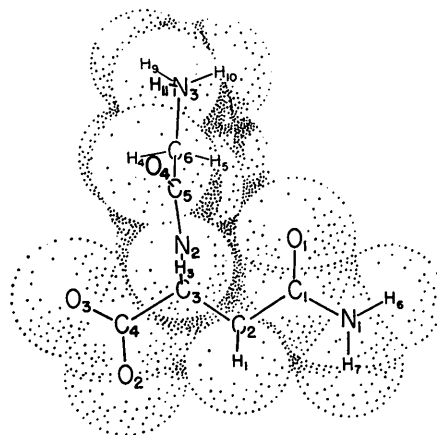

Fig. 6. A scale drawing of the molecule viewed in a direction perpendicular to its succinamic acid part. Hydrogen atoms H_2 and H_8 are not visible and are therefore omitted from the skeleton formula.
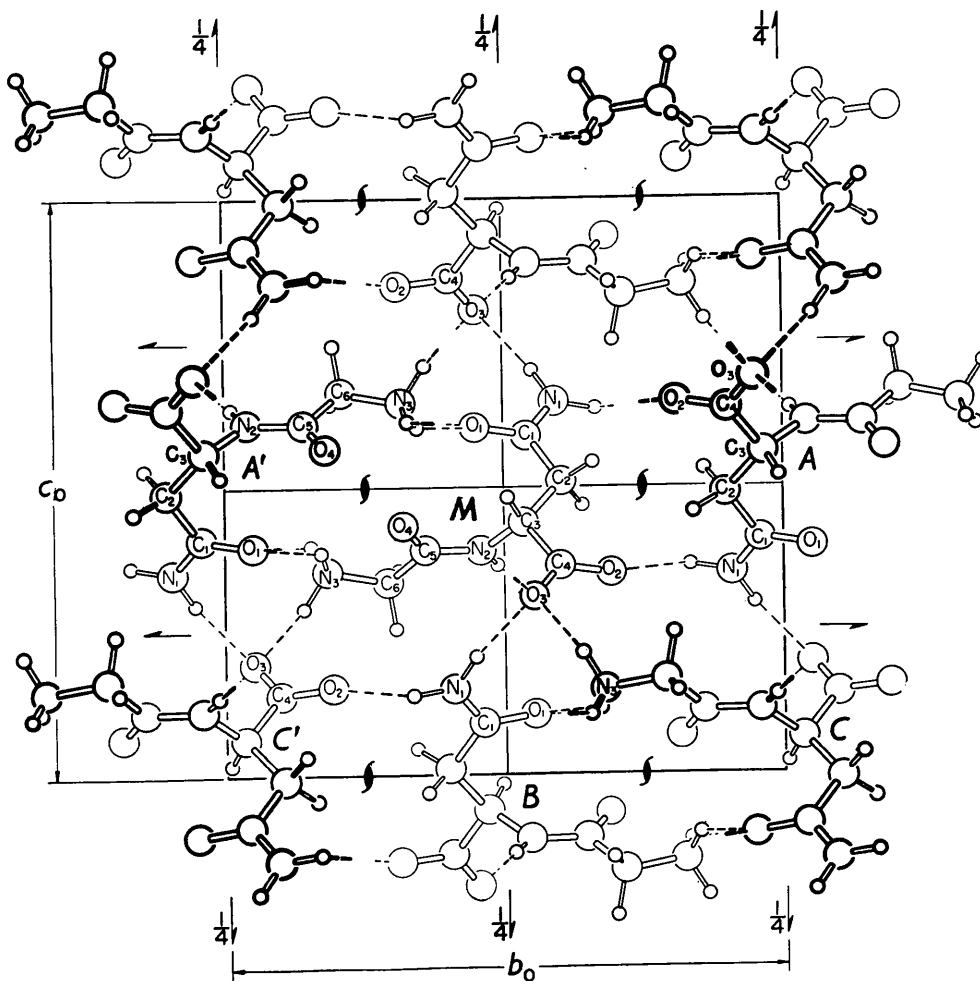

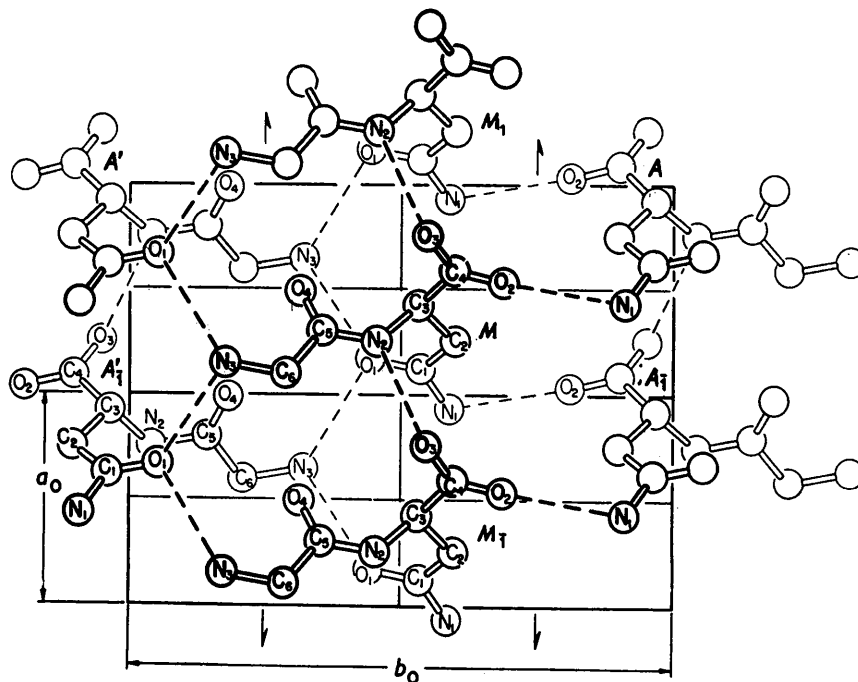Fig. 7. A projection of the structure viewed along the a axis of the crystal.

Fig. 8. A projection of half of the molecules in the unit cell viewed along the c axis of the crystal.

The environment of the molecule

Two views of the structure are shown in Figs. 7 and 8. In Fig. 7 the content of a unit cell, including the hydrogen atoms, is projected along the a axis. Fig. 8 is a projection on the (001) face of half of the unit cell with respect to c; the hydrogen atoms are omitted for clarity. The four equivalent molecules are labelled M, A, B, C, the parameters of the molecule M being those given in Table 1, column (3). The molecules A' and C' are related to A and C, respectively, by the unit vector translation $-b$, and subscripts 1 and $\bar{1}$ are used to signify molecules displaced by the unit vector translations a and $-a$, respectively.

The sterically simple configuration of the molecule, shown in Fig. 6 and described in the preceding section, permits fairly good packing. In the direction of the a axis the planar succinamic acid chains of identical molecules M, M_1, etc. in adjacent unit cells fit on top of each other, as shown in Fig. 9, a packing drawing of three such chains viewed in the direction of the b axis of the crystal. Oxygen atom O_1 is not visible and has therefore been omitted from the skeleton formula. The perpendicular distance between the planes is only 3·44 Å; this near approach is possible because the hydrogen atoms of the methylene groups fit neatly into interstices. In the packing drawing the oxygen atoms appear to be very loosely packed; the validity of this implication will be discussed later. The other planar part of the molecule, the peptide group, is tilted at an angle of about 45° to the a axis, and in the stacking of the molecules the oxygen atom O_4 in molecule M is brought in close contact with atoms of the molecule M_1 above (Fig. 8). The molecules

within the stacks are held together by intermolecular hydrogen bonds between the peptide nitrogen atoms N_2 and the carboxyl oxygen atoms O_3.

Four such equivalent stacks of molecules M, A, B, C, run through the unit cell in the direction of the a axis. Stacks M and A are related by a screw axis at $y = \frac{3}{4}, z = \frac{1}{4}$ parallel to a. The stacks are tied together
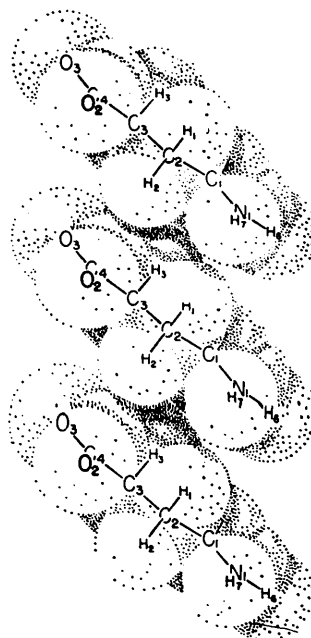

Fig. 9. A scale drawing showing the packing of the succinamic acid parts of three molecules viewed in the direction of the b axis.

by hydrogen bonds between oxygen atoms O_2 and nitrogen atoms N_1 as clearly shown in Figs. 7 and 8. These hydrogen bonds join the succinamic acid chains in a helix around the screw axis. Stack M is related to a second stack of molecules A' through another screw axis parallel to a at $y = \frac{1}{4}, z = \frac{1}{2}$. The amino groups of the one stack are situated nearly equidistant from two oxygen atoms O_1 of the terminal amide groups of the other, and two non-equivalent sets of hydrogen bonds formed by these atoms knit the stacks together in an infinite zig-zag chain (Fig. 8). In the opposing piles of peptide groups oxygen atom O_4 of molecule M fits into the space between atoms C_5, C_6, and N_3, of molecule A'; it does not make contact with its equivalent oxygen atom O_4 (3·42 Å).

The stacks of molecules M and A form a continuous layer parallel to the b direction in the crystal. This layer is related to the equivalent one, consisting of stacks of molecules B and C, by the screw axes parallel to b and c. These two layers are held together by two sets of hydrogen bonds, both involving the same oxygen atom O_3. The atoms C_6 and N_3 of one layer (molecule C) make near contacts with O_2 and N_1 of the other (molecules M and A, respectively).

Some intermolecular van der Waals' contacts are given in Table 6. Data concerning the hydrogen bonds

Table 6. *Some intermolecular interatomic distances*

From atom x in molecule M	to atom y	in molecule	Distance x–y (Å)
C_3	C_1	M_1	3·92
C_4	C_2	M_1	3·76
C_4	C_1	M_1	4·10
C_5	C_6	M_1	4·00
O_4	C_5	M_1	3·90
O_4	C_6	M_1	3·24
O_4	C_5	A'	3·56
O_4	C_6	A'	3·45
O_4	N_3	A'	2·97
C_6	N_1	B	3·76
C_6	O_2	C'	3·28
N_3	O_2	C'	3·14

are summarized in Table 7, in which are listed the lengths of the N...O vectors and the angles which they make with the directions of the C-N and N-H bonds. In structures of this sort the shape of the molecule is an important factor in determining the details of the molecular arrangement, and little quantitative significance can be ascribed to the lengths and the angles associated with the hydrogen bonds (Donohue, 1952; Corey & Pauling, 1953). In addition, uncertainties arise from the fact that the positions of the hydrogen atoms could not be fixed with any degree of accuracy. In the case of the hydrogen bond $N_2 \ldots O_3$ the hydrogen atom was assumed to be situated in the plane of the peptide group and the angle $H-N_2-C_5$ was taken to be 120°. There is little doubt that this choice is appropriate for the additional reason that the N–H bond forms an angle of only 7° with the direction of the N...O vector. This deviation arises largely from the fact that the oxygen atom O_3 does not lie precisely in the plane of the peptide amide group. The positions of the hydrogen atoms in the terminal amide group are less certain. In the assignment of their positions the assumption was made that these hydrogen atoms are co-planar with the amide group; but this assumption is open to some doubt because of the small amount of double-bond character associated with the C–N bond.

Around the charged amino nitrogen atom N_3 the arrangement of the hydrogen atoms is even more uncertain. This nitrogen atom is surrounded by not less than six oxygen atoms (one of them in the same molecule) at distances short enough to suggest the possibility of hydrogen bonds, the longest distance being 3·14 Å. No three of the oxygen atoms, however, are arranged tetrahedrally with respect to the C–N covalent bond, but two of them, O_3 in C' and O_1 in A'_1, are nearly at the corners of a tetrahedron and are so close to N_3 (2·75 and 2·86 Å) that they undoubtedly take part in hydrogen bonds. If, however, the amino group were oriented in such a way that two of the N–H bonds pointed directly at these two oxygen atoms the third N–H bond would form an angle of about 40° with the vector from N_3 to O_1 in A'. This latter arrangement was the one which was assumed in the early stages of the investigation. However, in previous studies of amino acids and peptides it has been found that all hydrogen atoms attached to a singly bonded nitrogen atom take part in hydrogen bonds involving short N...O distances and H–N...O angles generally less than 25°. By rotating the amino group around the C–N bond, an orientation was indeed found which allows the formation of three hydrogen bonds, two of which form H–N...O angles slightly

Table 7. *Hydrogen bond distances and angles*

From atom x in molecule M	to atom y	in molecule	Distance N $\cdots$ O (Å)	Angle O $\cdots$ N–C (°)	Angle O $\cdots$ N–H (°)
N_1	O_2	$A\bar{1}$	2·88	108·9	23
N_1	O_3	$B(a)$	2·93	111·3	12
N_2	O_3	$M\bar{1}$	2·88	$\begin{cases} 112·5(b) \\ 123·6(c) \end{cases}$	7
N_3	O_1	$A'_{\bar{1}}$	2·86	116·8	26
N_3	O_1	A'	3·03	132·8	28
N_3	O_3	C'	2·75	117·9	13

(a) Molecule displaced by $\bar{c}$ in respect to L. (b) Angle $O_3 \cdots N_2-C_3$. (c) Angle $O_3 \cdots N_2-C_5$.

greater than 25°. This orientation of the amino group agrees somewhat better with the positions of the three hydrogen peaks in the difference Fourier synthesis around the amino nitrogen atom. Furthermore, this distribution of hydrogen bonds seems to be a very satisfactory one for holding the structure firmly together. For these reasons we prefer this hydrogen bond system, and in the preceding discussion of the structure this arrangement has been assumed to be the correct one. The corresponding parameters, x, y, z, for the hydrogen atoms are

$$H_9, 0.145, 0.138, 0.290; H_{10}, -0.005, 0.152, 0.397;$$
$$H_{11}, 0.330, 0.170, 0.387.$$

In addition to the directed hydrogen bonds, electrostatic attraction between the positively charged nitrogen atom and the surrounding oxygen atoms, all of which carry a partial negative charge, will also contribute to the coherence of the structure.

The distribution of the hydrogen bonds among the oxygen atoms is very unusual; oxygen atom O_3 takes part in three, O_1 in two, O_2 in one, and O_4 in none at all. The three atoms O_1, O_2, O_3, are, as has been mentioned already, very loosely packed perpendicular to the plane of the succinamic acid chains. Around atoms O_1 and O_3 the hydrogen bonds are arranged so as to hold them firmly in place. The single hydrogen bond of atom O_2, however, which lies approximately in the plane of the chain, does not impose restrictions on the thermal vibration of the atom perpendicular to the plane of the chain; a motion in this direction also causes a minimum of distortion of bond angles. Very pronounced temperature anisotropy in this direction was indeed indicated by the difference Fourier. Although carbonyl oxygen atom O_4 does not participate in any hydrogen bond, no pronounced temperature anisotropy is expected or observed because it makes fairly good contacts with the surrounding atoms.

References

ALBRECHT, G. & COREY, R. B. (1939). J. Amer. Chem. Soc. 61, 1087.

BEEVERS, C. A. & ROBERTSON, J. H. (1950). Acta Cryst. 3, 164.

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

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

COCHRAN, W. & PENFOLD, B. R. (1952). Acta Cryst. 5, 644.

COREY, R. B. & PAULING, L. (1953). Proc. Roy. Soc. B, 141, 10.

DAWSON, B. (1953). Acta Cryst. 6, 81.

DONOHUE, J. (1950). J. Amer. Chem. Soc. 72, 949.

DONOHUE, J. (1952). J. Phys. Chem. 56, 502.

DONOHUE, J. & TRUEBLOOD, K. N. (1952a). Acta Cryst. 5, 414.

DONOHUE, J. & TRUEBLOOD, K. N. (1952b). Acta Cryst. 5, 419.

HUGHES, E. W. & BISWAS, A. B. (1954). To be published.

HUGHES, E. W. & MOORE, W. J. (1949). J. Amer. Chem. Soc. 71, 2618.

KATZ, L., PASTERNAK, R. A. & COREY, R. B. (1952). Nature, Lond. 170, 1066.

LEVY, H. A. & COREY, R. B. (1941). J. Amer. Chem. Soc. 63, 2095.

MACGILLAVRY, C. H., HOOGSCHAGEN, G. & SIXMA, F.L.J. (1948). Rec. Trav. chim. Pays-Bas, 67, 869.

MORRISON, J. D. & ROBERTSON, J. M. (1949). J. Chem. Soc. p. 980.

PATTERSON, A. L. (1935). Z. Kristallogr. 90, 517.

SENTI, F. & HARKER, D. (1940). J. Amer. Chem. Soc. 62, 2008.

SHOEMAKER, D. P., BARIEAU, R. E., DONOHUE, J. & LU, C.-S. (1953). Acta Cryst. 6, 241.

SHOEMAKER, D. P., DONOHUE, J., SCHOMAKER, V. & COREY, R. B. (1950). J. Amer. Chem. Soc. 72, 2328.

STEWARD, F. C. & THOMPSON, J. F. (1952). Nature, Lond. 169, 739.

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

YAKEL, JR., H. L. & HUGHES, E. W. (1952). J. Amer. Chem. Soc. 74, 6302.