Acta Cryst. (1953). **6**, 531

The Crystal Structure of Glycyl-L-Tyrosine Hydrochloride*

BY D. W. SMITS AND E. H. WIEBENGA

*Laboratorium voor Anorganische en Physische Chemie der Rijks-Universiteit Groningen,
The Netherlands*

(*Received* 6 *December* 1952)

The crystal structure of glycyl-L-tyrosine hydrochloride has been determined, making use of the isomorphism of the hydrochloride with the corresponding hydrobromide. Both compounds crystallize with one molecule of water; space group $P2_1$, two molecules per unit cell. After a direct approximate determination of the [010] and the [001] projection, the atomic coordinates were refined, mainly by consecutive Fourier syntheses of these projections. An essential improvement of the agreement between the calculated and the observed $h0l$ structure factors was obtained by the introduction of an anisotropic temperature factor for the water molecule. The final values for the reliability factor are $R_{h0l} = 0{\cdot}14$ and $R_{hk0} = 0{\cdot}10$. The atomic distances are subject to a probable error of approximately 0·02 Å when both atoms are resolved in the projections; for the other distances this error is estimated to be twice as large. The glycyltyrosine molecule consists of a nearly planar chain, a benzene ring, which is rotated with respect to the plane of the chain, and a carboxyl group, almost perpendicular to the chain direction and in *cis* position with the peptide carbonyl group. A discussion of the intra- and intermolecular distances is given.

1. Introduction

The present structure determination of glycyl-L-tyrosine hydrochloride was started some years ago with the purpose of obtaining information about the atomic configuration in the amide group.

The dipeptide was kindly put at our disposal by Prof. Dr E. Havinga, Leiden. It was there involved in an investigation of the properties of tyrosine as a constituent of proteins and peptides (Bruigom, 1950).

During our investigation the results of the determinations of the structures of two other dipeptides and a similar compound were published, viz. β-glycylglycine (Hughes & Moore, 1949), N-acetylglycine (Carpenter & Donohue, 1950) and cysteylglycine. ½NaI (Dyer, 1951). The information, obtained from the former two of these investigations, has been used recently by Pauling *et al.* in their considerations of the structure of proteins. The configuration about the amide group in the glycyltyrosine molecule appears to be the same as that observed in glycylglycine and acetylglycine.

2. Experimental

Materials

Crystals of glycyl-L-tyrosine hydrochloride and of the isomorphous hydrobromide were obtained from their aqueous solutions as small needles with {100}, {110} and $\{1\bar{1}0\}$. A chemical analysis showed that they contain one molecule water per molecule of the salt.

Unit cell and space group

The cell dimensions were obtained from rotation, oscillation and Weissenberg photographs about the *b* and *c* axes. They were determined with some care, using a procedure similar to that of Weisz, Cochran & Cole (1948). The crystals are monoclinic; with λ(Cu $K\alpha$) = 1·5418 Å the following values for the lattice constants were obtained:

	Glycyltyrosine–hydrochloride†	Glycyltyrosine–hydrobromide
a (Å)	14·05 ±0·01	14·25 ±0·01
b (Å)	$9{\cdot}35_3 \pm 0{\cdot}00_5$	$9{\cdot}55_4 \pm 0{\cdot}00_5$
c (Å)	$5{\cdot}10_7 \pm 0{\cdot}00_3$	$5{\cdot}08_8 \pm 0{\cdot}00_3$
β	91° 20′±4′	91° 4′±4′
Vol.	$670{\cdot}_9 \pm 1$ Å^3	$692{\cdot}_5 \pm 1$ Å^3

A density of 1·439 g.cm.$^{-3}$ for the hydrochloride was found by floating the crystals in chloroform–nitrobenzene mixtures. With this value the number of molecules per cell was calculated to be 1·99.

No systematic extinctions other than $0k0$ for k odd were observed. Consequently the possible space-groups are $P2_1/m$, $P2_1$, $P2/m$, $P2$ and Pm. Of these $P2_1/m$ and $P2/m$ could be excluded because of their fourfold general equivalent positions, whereas $P2$ and Pm appeared to be incompatible with the [001] and [010] Patterson syntheses. This leaves as the correct space group $P2_1$, with the general equivalent positions $x, y, z;\ \bar{x}, \tfrac{1}{2}+y, \bar{z}$.

---

* The determination of this structure is described in more detail by Smits (1952).

† These unit-cell dimensions agree with those given by Tranter (1952) in a recent communication, which appeared after our work was finished.

Structure factors

The intensities of the $h0l$ and the $hk0$ reflexions of both compounds were measured on integrated Weissenberg photographs (Wiebenga & Smits, 1950; Smits & Wiebenga, 1953). The very weak reflexions were estimated visually on ordinary Weissenberg photographs of long exposure times. Each exposure was taken with nickel-filtered Cu K radiation, using three sheets of Ilford Ilfex X-ray film in the camera. A graphical correction for absorption was applied, which was less than 10% because of the small diameter (about 0·1 mm.) of the crystals used. After correction for Lorentz and polarization factors relative values of $|F_o|$ were obtained, which were put on an absolute scale during the structure determination.

For glycyltyrosine hydrochloride all 179 accessible structure factors $hk0$ and 145 out of the 187 accessible structure factors $h0l$ were observed; they are given in Tables 2 and 3. For the hydrobromide the $|F_o|$ values were obtained for 155 out of the 187 accessible $hk0$ reflexions and for 177 out of the 191 accessible $h0l$ reflexions..

3. Determination of the structure

Approximate determination of the [010] *projection*

In this projection the Br⁻ ion was located by means of a [010] Patterson synthesis of the hydrobromide, which showed very distinctly the Br–Br vector; evidently this also reveals the position of the Cl⁻ ion in the hydrochloride. An [010] projection of the electron-density distribution was then directly obtained in two different ways:

(*a*) For the hydrobromide the signs of 125 out of the 177 measured structure factors (representing 85% of $\Sigma|F_o|$) could be determined with reasonable certainty by comparing the magnitude of each structure factor with the calculated contribution F_{Br} of the Br⁻ ions and the average contribution of the other atoms, which was calculated in a way similar to that given by Wilson (1949). The experimental $|F_o|$ values were put on an approximately absolute scale by considering the reflexions which were not observed for the hydrochloride. For these reflexions the structure factors for the hydrobromide must be approximately equal to $(F_{\text{Br}}-F_{\text{Cl}})_{\text{calc.}}$. An isotropic temperature factor with $B = 3{\cdot}4$ Å^2 was chosen, which represents a mean value found for other comparable compounds. The Fourier synthesis based on the 125 terms of which the signs could be determined, was only in details different from the final one, given in Fig. 4.

(*b*) A more complete Fourier synthesis was obtained by comparing the structure factors of the hydrobromide with those of the isomorphous hydrochloride. Making use of the relation

$$F_{\text{GT.HBr}}-F_{\text{GT.HCl}} = F_{\text{Br}}-F_{\text{Cl}}$$

the signs of 154 out of the 177 structure factors (94% of $\Sigma|F_o|$) for the hydrobromide, and 111 out of the 145 structure factors (86% of $\Sigma|F_o|$) for the hydrochloride could be determined. Fig. 1 shows the contour-map which was thus obtained for the hydrobromide.

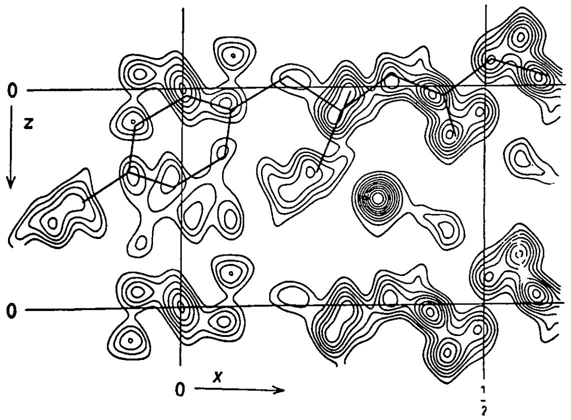


Fig. 1. [010] Fourier synthesis of glycyltyrosine hydrobromide, obtained by application of the isomorphous-replacement method. Contours are at equal intervals on an arbitrary scale; in the Br⁻ ion every fourth contour line is drawn.

As is seen, a model of the molecule could be fitted satisfactorily on this map. During the subsequent refinement of the parameters this first interpretation of the [010] projection proved to be essentially correct.

The x and z coordinates were first refined by the Fourier method. To this end the coordinates of those atoms which were not resolved in the projection were taken from the model and the signs of some more structure factors could be determined. From the [010] Fourier syntheses thus obtained, the coordinates of the resolved atoms were taken. A further refinement of these coordinates and those of the non-resolved atoms was then tried for the hydrochloride, making use of structure-factor maps (Bragg & Lipson, 1936). The progress made in this way, however, was rather slow and it was decided to postpone further refinement of the x, z coordinates until information about the [001] projection had been obtained.

Determination of the [001] *projection*

By comparison of the $hk0$ structure factors of the hydrobromide and the hydrochloride, the phases of the structure factors for this non-centrosymmetric projection could be approximately determined, apart from the ambiguity α or $2\pi-\alpha$. Taking both phases for each reflexion, a Fourier synthesis for each compound was obtained, representing the superposition of the structure and the inverted one. Although the errors in the phases were rather large (see Table 3) because of the uncertainty in the value of the temperature factor, possible differences between corresponding coordinates in the two compounds, and the experimental errors in the measured intensities, the agreement between the 'double' contour maps obtained for both salts, as well as their general appearance, was surprisingly good.

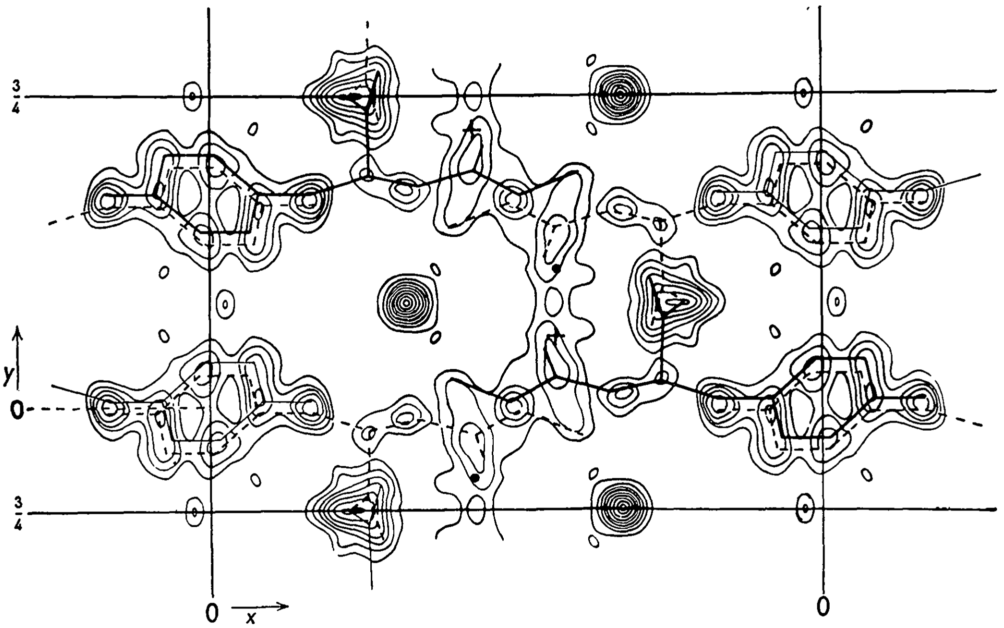


Fig. 2. [001] 'double' Fourier synthesis of glycyltyrosine hydrochloride, obtained by application of the isomorphous-replacement method. The inverted structure is indicated by broken lines; ● and + show the position of the water molecule in the structure and in the inverted structure respectively. Contours are at equal intervals on an arbitrary scale; in the Cl⁻ ion every second contour line is drawn.

The [001] 'double' Fourier synthesis is shown for the hydrochloride in Fig. 2. Variations of the model were now considerably restricted by the y coordinates of the resolved atoms C_8 and N_1,* and by those of the benzene-ring atoms. This made it possible to take the y coordinates of the other atoms, except the oxygen atom of the water molecule, from the model. The position of the water molecule was determined by some trial, after which the $hk0$ structure factors were calculated for the hydrochloride, giving a reliability factor

$$R \equiv \Sigma||F_c|-|F_o||\div\Sigma|F_o|$$

of 0·25.

The x, y coordinates were then refined by successive Fourier syntheses in the usual way. Fortunately only three atoms (C_8, C_{10}, O_3) were not resolved in the [001] projection. Generally the 'double shift rule' was applied, correction for termination of the series was made by the 'back shift' method (Booth, 1946).

Further refinement of the parameters

After this refinement of the x, y coordinates, which seemed rather complete, we returned to the [010] projection. With the x coordinates taken from the [001] projection, the z coordinates for the resolved atoms from the last [010] Fourier synthesis, and the other z coordinates from the model, the $h0l$ structure factors were recalculated. They showed an agreement with the observed structure factors which was better than that hitherto obtained.

Further refinement of the parameters was attempted partly by the Fourier method and partly by trial. It appeared, however, that in this way for the $h0l$ structure factors no R value lower than approximately 0·18 could be obtained. It was suspected that this rather high value might be caused by an anisotropic thermal motion of the water molecule which was not taken into account when calculating the structure

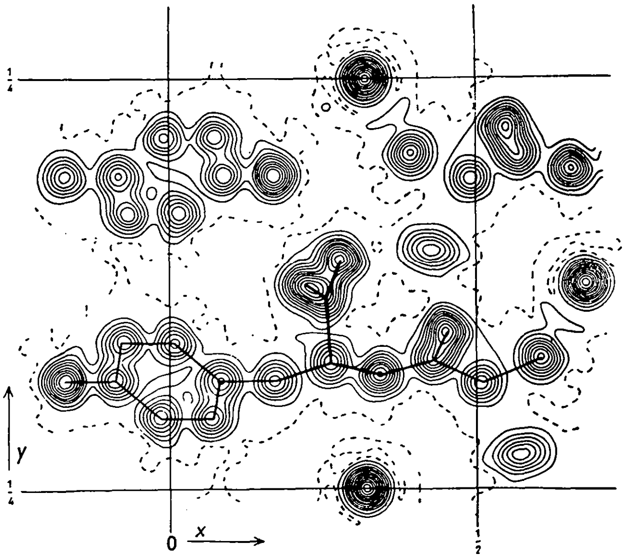


Fig. 3. Final [001] electron-density projection of glycyltyrosine hydrochloride. Contours are at intervals of 1 e.Å^{-2}, the one-electron contour being broken. In the Cl⁻ ion contours are at 1, 2, 4, 6... e.Å^{-2}.

---

* For the way of indicating the atoms see Fig. 5.

Table 1. *Final atomic coordinates*

	x	y	z		x	y	z		x	y	z
C_1	0·912	0·517	0·378	C_9	0·255	0·718	0·168	O_4	0·452	0·633	0·224
C_2	0·921	0·610	0·178	C_{10}	0·430	0·565	0·043	O_5	0·427	0·833	0·644
C_3	0·008	0·609	0·028	C_{11}	0·510	0·508	0·874				
C_4	0·082	0·516	0·107					N_1	0·344	0·529	0·941
C_5	0·070	0·424	0·312	O_1	0·829	0·516	0·517	N_2	0·605	0·568	0·955
C_6	0·985	0·424	0·460	O_2	0·278	0·811	0·016				
C_7	0·173	0·515	0·946	O_3	0·220	0·748	0·386	Cl	0·320	0·250	0·540
C_8	0·262	0·560	0·107								

factors. A strong indication for this was the fact that even in the last [010] Fourier synthesis the maximum representing the water molecule was not circularly symmetric and very low.

An anisotropic temperature factor $\exp[-(\alpha+\beta\cos^2\varphi)\sin^2\theta/\lambda^2]$ (Helmholz, 1936) for the water molecule was now introduced (φ represents the angle between the direction of maximum vibration and the normal on (hkl)). In the [010] Fourier projection an angle of 29° between the direction of maximum vibration and the x axis was measured. For α we maintained the value 3·0 Å^2 used in the isotropic temperature factor for the other atoms; for β we accepted after some trial a value of 14·0 Å^2. In this way the value of R_{h0l} was reduced to 0·14.

In the [001] Fourier synthesis only a slight indication of an anisotropic thermal motion of the water molecule was observed. When applying the anisotropic temperature factor for the water molecule in the calculation of the $hk0$ structure factors, R_{hk0} was, however, also somewhat reduced and decreased to 0·10.

The difference, which remains between the R values for both projections, can probably be explained by the fact that in the non-centrosymmetric case the differences $||F_c|-|F_o||$ are smaller than the moduli of the vector differences $\mathbf{F}_c-\mathbf{F}_o$ (see also Shoemaker, Donohue, Schomaker & Corey, 1950).

In Table 1 the finally accepted atomic coordinates are collected. Figs. 3 and 4 show the final Fourier syntheses of the [001] and [010] projections respectively. The height of the maxima and the shape of the contour lines are practically identical with those in the corresponding F_c syntheses, which are not shown. Tables 2 and 3 show the agreement between calculated and observed F values.

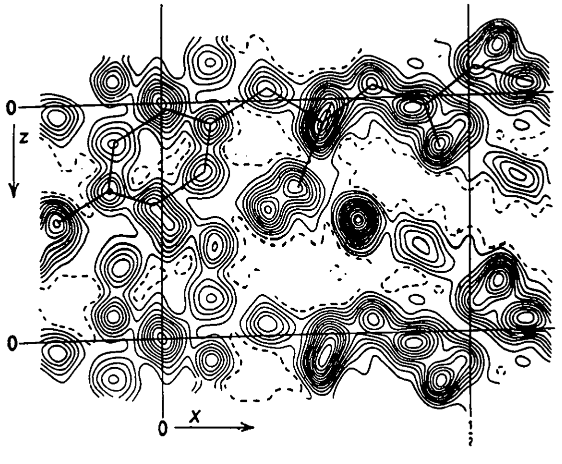


Fig. 4. Final [010] electron-density projection of glycyltyrosine hydrochloride. Contours are at intervals of 1 e.Å^{-2}, the two-electron contour being broken. In the Cl$^-$ ion contours are at 2, 3, 4, 6, 8... e.Å^{-2}.

4. Discussion

Accuracy of the atomic coordinates

The accuracy of the final coordinates was estimated by comparing the values of the x coordinates obtained from the [010] and [001] projections respectively. This gave for the probable error in the atomic distances a value of approximately 0·02 Å when the atoms are resolved in both projections. For the other distances it is very difficult to make an estimate of their accuracy. Considering the very close resemblance of the final F_c and F_o syntheses, however, it is likely that their inaccuracy is less than twice the inaccuracy for the former distances.

Shape and dimensions of the glycyltyrosine molecule

Fig. 5 shows the molecule as it is seen in the [001] projection.

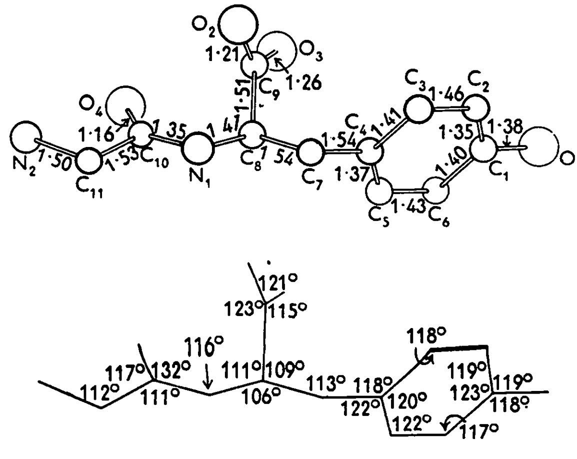


Fig. 5. Intramolecular bond lengths and bond angles.

The *chain* N_2–C_{11}–C_{10}–N_1–C_8–C_7 is approximately planar. A similar planar chain has been found in β-glycylglycine (Hughes & Moore, 1949) and N-acetylglycine (Carpenter & Donohue, 1950); cysteylglycine.$\frac{1}{2}$NaI (Dyer, 1951), on the contrary, shows a

Table 2. *Comparison of observed and calculated structure factors hOl*

The calculation of F_c is based on the final atomic coordinates (Table 1)

h0l	F_o	F_c
000		+308
100	9·9	−9·1
200	3·5	−4·1
300	21·2	+21·4
400	15·5	+13·9
500	8·9	−9·2
600	1·7	−2·4
700	3·1	+3·6
800	21·1	−20·5
900	4·4	+6·1
10,0,0	2·3	+3·6
11,0,0	8·0	+6·7
12,0,0	22·2	+20·1
13,0,0	7·0	+8·4
14,0,0	3·0	−2·4
15,0,0	5·3	+4·2
16,0,0	3·9	+2·4
17,0,0	7·0	−6·1
18,0,0	2·2	+3·0
$\overline{17}$,0,1	2·3	+2·1
$\overline{16}$,0,1	< 2·1	−1·2
$\overline{15}$,0,1	< 2·4	+1·6
$\overline{14}$,0,1	6·3	+6·7
$\overline{13}$,0,1	11·0	−13·6
$\overline{12}$,0,1	< 2·7	−0·4
$\overline{11}$,0,1	13·7	+15·1
$\overline{10}$,0,1	11·3	−9·5
$\bar{9}01$	6·6	−7·9
$\bar{8}01$	18·4	+18·7
$\bar{7}01$	< 2·0	+2·0
$\bar{6}01$	24·9	−24·2
$\bar{5}01$	27·7	+25·5
$\bar{4}01$	6·9	−6·4
$\bar{3}01$	11·4	−12·0
$\bar{2}01$	14·6	+17·0
$\bar{1}01$	6·3	−3·8
001	17·8	+13·9
101	31·4	−32·6
201	11·1	+11·9
301	10·4	+10·5
401	21·4	+25·7
501	25·6	−25·7
601	18·1	−21·3
701	11·8	+12·2
801	19·5	+21·9
901	22·0	−24·4
10,0,1	< 2·5	+0·1
11,0,1	13·6	+15·8
12,0,1	< 2·7	+0·7
13,0,1	5·3	−5·8
14,0,1	8·1	+9·0
15,0,1	4·7	+3·4
16,0,1	4·9	−3·4
17,0,1	< 1·6	+2·6
$\overline{17}$,0,2	4·1	+3·0
$\overline{16}$,0,2	5·8	+5·0
$\overline{15}$,0,2	3·1	−2·6
$\overline{14}$,0,2	6·0	−4·5
$\overline{13}$,0,2	9·1	+8·1
$\overline{12}$,0,2	10·5	+11·4
$\overline{11}$,0,2	3·8	−3·5
$\overline{10}$,0,2	8·7	+11·3
$\bar{9}02$	< 2·5	−1·5
$\bar{8}02$	< 2·4	+1·6
$\bar{7}02$	7·0	+8·8
$\bar{6}02$	7·4	+8·0
$\bar{5}02$	24·4	+20·0
$\bar{4}02$	2·5	+1·8
$\bar{3}02$	39·6	+41·2
$\bar{2}02$	55·1	−50·9
$\bar{1}02$	11·9	+13·5
002	53·1	+50·7
102	8·6	−12·2
202	2·3	−4·6
302	14·9	+14·4
402	10·0	−12·6
502	7·6	−7·7
602	37·7	+43·3
702	16·9	+15·5
802	7·7	−9·7
902	3·6	+0·7
10,0,2	7·5	+8·7
11,0,2	5·4	−4·7
12,0,2	7·6	+8·0
13,0,2	8·2	−8·9
14,0,2	5·9	−4·4
15,0,2	< 2·1	+2·6
16,0,2	5·4	−4·1
17,0,2	4·0	−4·5
$\overline{16}$,0,3	4·6	−5·8
$\overline{15}$,0,3	4·3	+4·9
$\overline{14}$,0,3	6·7	+5·9
$\overline{13}$,0,3	6·3	−6·2
$\overline{12}$,0,3	7·3	−5·3
$\overline{11}$,0,3	7·6	+6·8
$\overline{10}$,0,3	3·9	−4·1
$\bar{9}03$	< 2·7	+0·9
$\bar{8}03$	7·5	+8·8
$\bar{7}03$	12·4	−13·3
$\bar{6}03$	6·9	−8·1
$\bar{5}03$	21·2	+23·3
$\bar{4}03$	11·7	−12·2
$\bar{3}03$	7·4	−6·8
$\bar{2}03$	12·9	+15·9
$\bar{1}03$	2·0	−4·0
003	2·9	−3·1
103	2·9	+2·8
203	4·2	−4·8
303	6·1	+6·6
403	6·4	+9·0
503	< 2·3	+0·5
603	14·4	−19·2
703	16·2	+17·9
803	3·8	−1·3
903	5·5	−4·1
10,0,3	7·7	+7·5
11,0,3	< 2·7	+0·6
12,0,3	5·0	−5·6
13,0,3	< 2·3	+0·8
14,0,3	7·8	+9·4
15,0,3	3·9	−4·1
$\overline{14}$,0,4	2·0	−2·3
$\overline{13}$,0,4	4·5	+4·4
$\overline{12}$,0,4	7·4	−6·9
$\overline{11}$,0,4	4·8	+3·2
$\overline{10}$,0,4	< 2·6	+3·1
$\bar{9}04$	6·6	−7·3
$\bar{8}04$	6·7	−4·6
$\bar{7}04$	11·6	+12·2
$\bar{6}04$	13·2	+13·0
$\bar{5}04$	5·3	+4·6
$\bar{4}04$	12·2	+15·0
$\bar{3}04$	10·9	+14·3
$\bar{2}04$	6·2	−6·8
$\bar{1}04$	2·5	+2·4
004	2·5	+5·2
104	9·4	−9·6
204	< 2·6	+2·9
304	< 2·6	+1·9
404	13·9	−17·7
504	< 2·7	+0·4
604	9·4	+13·7
704	8·6	+11·7
804	7·6	−7·8
904	9·4	+10·2
10,0,4	6·1	−4·7
11,0,4	< 2·3	−0·8
12,0,4	< 2·0	+3·8
13,0,4	< 1·7	−1·4
14,0,4	2·9	−4·6
$\overline{11}$,0,5	< 1·6	+2·1
$\overline{10}$,0,5	1·9	−3·0
$\bar{9}05$	< 2·1	+1·5
$\bar{8}05$	5·7	+6·3
$\bar{7}05$	5·9	−5·0
$\bar{6}05$	6·2	−6·6
$\bar{5}05$	8·9	+10·2
$\bar{4}05$	6·4	−7·2
$\bar{3}05$	7·4	+8·5
$\bar{2}05$	< 2·6	+2·4
$\bar{1}05$	7·5	−7·2
005	< 2·6	−0·3
105	6·5	+5·7
205	2·6	−2·2
305	9·0	−8·6
405	9·7	+11·7
505	< 2·5	−2·9
605	< 2·4	−5·1
705	7·4	+8·8
805	< 2·2	−1·7
905	4·0	+4·1
10,0,5	2·4	+4·9
11,0,5	< 1·4	+0·5
$\bar{7}06$	4·6	+5·2
$\bar{6}06$	< 1·6	−2·4
$\bar{5}06$	< 1·7	−0·7
$\bar{4}06$	< 1·9	+2·1
$\bar{3}06$	< 1·9	−1·6
$\bar{2}06$	5·7	−8·6
$\bar{1}06$	< 2·0	+5·2
006	< 2·0	+2·1
106	< 2·0	−0·5
206	5·6	+7·9
306	< 1·9	+3·0
406	< 1·8	+0·9
506	< 1·6	−1·3
606	< 1·4	+1·1
706	2·4	−2·1

Table 3. *Comparison of observed and calculated hk*0 *structure factors and phase angles*

The calculation of F_c and α_c is based on atomic coordinates differing slightly from those of Table 1

*hk*0	F_o	α_o (°)	F_c	α_c (°)
000			308	
100	9·9	180	9·3	180
200	3·5	180	3·5	180
300	21·2	0	22·0	0
400	15·5	0	15·7	0
500	8·9	180	10·2	180
600	1·7	180	2·5	180
700	3·1	?	3·7	0
800	21·1	180	19·5	180
900	4·4	0	5·8	0
10,0,0	2·3	0	2·2	0
11,0,0	8·0	180	6·3	0
12,0,0	22·2	0	20·7	0
13,0,0	7·0	0	9·4	0
14,0,0	3·0	180	2·3	180
15,0,0	5·3	0	4·4	0
16,0,0	3·9	?	2·8	0
17,0,0	7·0	180	5·2	180
18,0,0	2·2	?	3·3	0
110	18·3	33	19·4	−25
210	7·4	81	10·4	+90
310	19·0	122	17·4	+163
410	45·3	180	43·7	−170
510	48·4	37	45·1	−11
610	7·7	?	5·5	+117
710	16·3	180	17·8	−161
810	15·0	0	12·4	+60
910	18·7	76	18·3	−14
10,1,0	6·9	180	7·1	−142
11,1,0	14·0	99	12·2	+84
12,1,0	18·7	62	15·0	−19
13,1,0	10·1	122	10·2	+146
14,1,0	2·3	?	1·9	+83
15,1,0	6·8	54	7·0	+19
16,1,0	1·9	73	2·7	−151
17,1,0	1·8	?	3·3	−163
18,1,0	2·1	?	3·5	−14
020	44·6	52	40·9	+54
120	29·5	56	31·8	−17
220	48·2	0	50·9	+11
320	30·1	180	29·9	−166
420	24·4	180	25·0	−170
520	13·0	60	14·4	+22
620	7·3	120	3·2	−46
720	12·2	?	8·1	−168
820	5·2	0	4·8	−8
920	21·1	180	19·7	+164
10,2,0	5·0	0	3·4	−95
11,2,0	18·5	75	18·6	+15
12,2,0	12·0	33	11·4	+23
13,2,0	4·0	?	4·4	+135
14,2,0	6·9	0	6·8	−22
15,2,0	3·6	180	2·9	+160
16,2,0	4·7	61	2·9	+157
17,2,0	4·2	39	3·9	−20
130	17·9	77	13·8	−76
230	35·1	0	35·5	+158
330	3·1	180	2·7	+31
430	23·4	49	21·3	−48
530	13·4	180	14·2	−172
630	21·1	118	19·7	+178
730	29·2	27	29·7	+11
830	3·6	131	3·9	−114
930	6·9	54	5·3	+90
10,3,0	18·3	0	17·1	+2
11,3,0	5·6	112	5·5	+108
12,3,0	6·0	?	5·6	−35
13,3,0	7·6	0	8·2	+31
14,3,0	5·0	?	3·8	+68
15,3,0	5·8	180	6·3	−150
16,3,0	2·7	0	3·5	−10
17,3,0	2·6	?	3·3	+7
040	51·5	0	52·7	+46
140	20·7	126	19·5	−87
240	42·2	168	43·1	−171
340	2·7	143	2·0	+27
440	11·2	180	9·7	−126
540	18·9	145	14·9	+176
640	22·9	0	19·6	−22
740	18·2	?	15·9	+150
840	11·0	180	9·1	−166
940	5·4	38	6·4	−87
10,4,0	4·3	?	3·9	+31
11,4,0	9·5	180	11·6	+159
12,4,0	13·2	0	11·9	+17
13,4,0	3·7	?	3·9	−86
14,4,0	7·9	180	6·9	+172
15,4,0	3·0	?	3·5	−179
16,4,0	3·5	?	2·7	−85
17,4,0	3·0	180	3·4	−162
150	7·5	136	9·5	+148
250	16·8	0	17·4	−6
350	13·2	119	14·5	+177
450	12·1	180	10·3	−134
550	10·9	46	10·6	+45
650	16·7	75	17·3	−41
750	15·1	123	15·8	+161
850	4·4	66	3·3	+96
950	11·6	5	10·2	−3
10,5,0	3·9	180	3·3	−167
11,5,0	3·6	?	3·8	+149
12,5,0	9·9	51	8·8	+14
13,5,0	6·0	103	4·9	−149
14,5,0	1·7	?	1·5	−162
15,5,0	5·5	86	6·9	+7
16,5,0	3·3	97	4·0	−155
060	13·9	150	11·9	+146
160	2·3	90	2·7	−140
260	13·4	58	13·1	−52
360	23·2	180	23·8	−172
460	10·2	124	9·2	−123
560	2·9	120	2·8	+117
660	2·6	180	1·0	−156
760	4·2	?	4·9	−90
860	5·0	0	2·9	+35
960	11·5	180	13·2	−162
10,6,0	3·4	?	3·9	−145
11,6,0	11·8	32	11·5	−10
12,6,0	7·1	111	6·7	+103
13,6,0	6·5	?	7·1	−147
14,6,0	4·9	37	3·9	−50
15,6,0	5·4	180	5·0	−158
170	14·4	0	14·3	+14
270	14·8	129	14·5	+161
370	3·8	?	2·9	−60
470	18·3	0	17·7	0
570	6·8	180	6·5	+141
670	8·4	143	8·2	−170
770	8·6	0	7·3	−26
870	2·3	52	1·5	−122
970	7·2	142	7·8	−173
10,7,0	3·9	0	2·0	−47
11,7,0	9·2	180	8·1	−135
12,7,0	3·0	180	1·9	+175
13,7,0	1·8	0	2·7	+19
14,7,0	4·3	?	3·7	−141
080	13·7	0	12·9	+23
180	3·1	?	4·0	+70
280	5·3	180	3·1	−178
380	3·1	118	3·0	−132
480	5·7	180	3·9	−109
580	7·7	180	8·3	−158
680	14·7	41	14·2	+24
780	9·7	?	8·8	−153
880	9·7	180	9·3	−157
980	7·2	8	8·1	−11
10,8,0	4·0	?	2·3	−21
11,8,0	5·0	126	4·5	+144
12,8,0	5·0	59	4·5	+25
13,8,0	5·3	?	6·3	−1
190	5·1	180	4·7	−170
290	4·9	0	6·0	+42
390	5·6	?	5·6	+4
490	8·4	180	8·4	+162
590	3·2	0	3·8	−10
690	8·4	55	6·8	−9
790	6·1	108	5·1	+174
890	4·6	?	5·0	−148
990	3·4	?	3·1	−40
10,9,0	4·3	180	4·2	−154
11,9,0	2·9	?	2·8	−123
0,10,0	13·0	99	14·1	+104
1,10,0	10·1	77	9·5	+29
2,10,0	6·2	0	6·4	+22
3,10,0	6·3	180	5·5	−173
4,10,0	2·7	0	2·5	+51
5,10,0	3·6	?	2·5	−82
6,10,0	2·1	180	2·1	+137
7,10,0	2·7	?	2·0	−117
8,10,0	4·8	0	4·6	−27
9,10,0	6·4	180	6·5	180
1,11,0	6·0	55	5·6	+50
2,11,0	6·3	130	5·6	−154
3,11,0	2·2	?	1·3	−148
4,11,0	7·5	42	7·1	−8
5,11,0	6·7	136	5·9	+146
6,11,0	2·6	?	2·9	−114
7,11,0	3·6	0	3·9	0
0,12,0	5·9	63	7·0	+19
1,12,0	5·0	0	5·9	+94

considerable deviation from a planar configuration of the chain atoms.

The *carboxyl group* is situated with atom C_8 in a plane almost perpendicular to the direction of the chain. The carboxyl group, and the carbonyl group of the amide configuration are in *cis* position, which is different from the situation in glycylglycine and acetylglycine, where they are *trans*.

Owing to the mutual repulsion of O_3 and C_5, the plane of the *benzene ring* is different from that of the chain.

The interatomic distances and valence angles (see Table 4) will now be discussed in some more detail.

Table 4. *Intramolecular bond lengths and bond angles*

C_1-C_2	1·35 Å	$C_1-C_2-C_3$	119°
C_2-C_3	1·46	$C_2-C_3-C_4$	118
C_3-C_4	1·41	$C_3-C_4-C_5$	120
C_4-C_5	1·37	$C_4-C_5-C_6$	122
C_5-C_6	1·43	$C_5-C_6-C_1$	117
C_6-C_1	1·40	$C_6-C_1-C_2$	123
C_1-O_1	1·38	$C_2-C_1-O_1$	119
C_4-C_7	1·54	$C_6-C_1-O_1$	118
C_7-C_8	1·54	$C_3-C_4-C_7$	118
C_8-C_9	1·51	$C_5-C_4-C_7$	122
C_9-O_2	1·21	$C_4-C_7-C_8$	113
C_9-O_3	1·26	$C_7-C_8-C_9$	109
C_8-N_1	1·47	$C_8-C_9-O_2$	123
N_1-C_{10}	1·35	$C_8-C_9-O_3$	115
$C_{10}-O_4$	1·16	$O_2-C_9-O_3$	121
$C_{10}-C_{11}$	1·53	$C_9-C_8-N_1$	111
$C_{11}-N_2$	1·50	$C_7-C_8-N_1$	106
		$C_8-N_1-C_{10}$	116
		$N_1-C_{10}-O_4$	132
		$N_1-C_{10}-C_{11}$	111
		$O_4-C_{10}-C_{11}$	117
		$C_{10}-C_{11}-N_2$	112

(*a*) The atoms of the benzene ring belong essentially to the plane

$$0{\cdot}792x+y+0{\cdot}479z = 0{\cdot}631\ .$$

The largest deviation from this plane is only 0·014 Å.

The six C–C distances in the ring show a variation from 1·35 to 1·46 Å; it is questionable, however, whether the deviations from a regular hexagon are significant or not. It may be remarked that similar deviations have recently been reported for 1,4-dimethoxybenzene (Goodwin, Przybylska & Robertson, 1950). The atoms O_1 and C_7, adjacent to the benzene ring, belong to the same plane, apart from small deviations of 0·015 and 0·029 Å respectively. The distance C_1-O_1 (1·38 Å) agrees with the distance between an aromatic carbon atom and a phenolic hydroxyl group observed in other compounds. Examples are 1·36 Å in α-resorcinol (Robertson, 1936), 1·36 and 1·37 Å in β-resorcinol (Robertson & Ubbelohde, 1938), and 1·36 Å in some addition complexes of quinol (Palin & Powell, 1947). An interesting exception is the value of 1·47 Å observed for *p*-aminophenol (Brown, 1951), which is even higher than the standard length of the single carbon–oxygen bond (1·43 Å, Pauling, 1940).

(*b*) The carboxyl group is also essentially flat and belongs to the plane

$$x-0{\cdot}007y+0{\cdot}156z = 0{\cdot}275\ .$$

The distances C_9-O_3. and C_9-O_2 (1·26 and 1·21 Å respectively) and the valence angles agree with those observed for carboxyl groups in other compounds. It must be kept in mind, however, that the probable error in the positions of C_9, O_2 and O_3 is rather large because of the fact that neither of these atoms is resolved in both projections. The intermolecular distances indicate that the hydrogen atom is connected to O_3.

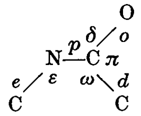
(*c*) The chain atoms N_1, C_{10}, O_4 and C_{11} belong in very close approximation to the plane

$$-0{\cdot}019x+y-0{\cdot}360z = 0{\cdot}543\ .$$

The atoms N_2 and C_8, however, show significant deviations (0·24 and 0·20 Å respectively) from this plane. The plane through N_1, C_8 and C_7 makes an angle of 12° with the plane through N_1, C_{10}, O_4 and C_{11}.

The distances and valence angles in the peptide configuration C_8, N_1, C_{10}, O_4, C_{11} agree on the whole with the values observed in other compounds (see Table 5). Unfortunately, C_{10} is not resolved in either projection, which makes the accuracy of its coordinates somewhat more doubtful than that of the other atoms. Inaccuracies in the position of C_{10} may be the explanation for the observed short distance $C_{10}-O_4$ (1·16 Å) as compared with the standard length $1{\cdot}21_5$ Å of the 'pure' double C=O bond (Pauling, 1940). On the other hand, small lengths for resonating C–O

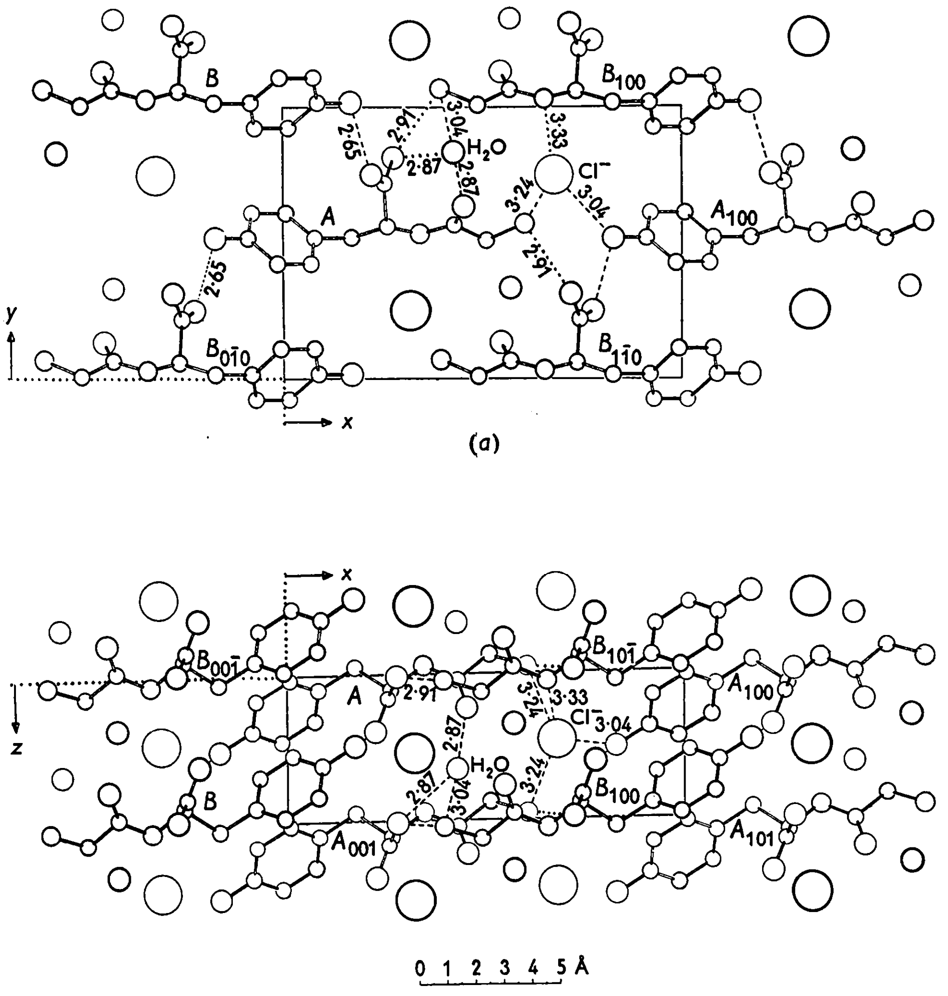
Table 5. *Comparison of bond lengths and bond angles about the amide group of various compounds*



	p (Å)	*o* (Å)	*d* (Å)	*e* (Å)	π (°)	ω (°)	δ (°)	ε (°)	Ref.†
Glycyltyrosine.HCl.H_2O	1·35	1·16	1·53	1·47	117	111	132	116	
β-Glycylglycine	1·29	1·23	1·53	1·48	121	114	125	122	1
Cysteylglycine.½NaI	1·32	1·21	1·54	1·33	109	126	107	139	2
Acetylglycine*	1·32	1·24	1·50	1·45	121	118	121	120	3
Diketopiperazine	1·33	1·25	1·47	1·41	—	—	—	—	4
Cyanuric acid*	1·35	1·21	—	—	—	115	122	125	5

* Determined by three-dimensional Fourier synthesis; standard deviation in the bond lengths smaller than 0·02 Å.

† (1) Hughes & Moore, 1949; (2) Dyer, 1951; (3) Carpenter & Donohue, 1950; (4) Corey, 1938; (5) Wiebenga, 1952.



(a)

(b)

Fig. 6. (a) [001] projection of the crystal structure. A: molecule with atomic coordinates x, y, z from Table 1; B: molecule with atomic coordinates $\bar{x}, \frac{1}{2}+y, 1-z$; Cl^-: coordinates $x = 0{\cdot}680$, $y = 0{\cdot}750$, $z = 0{\cdot}460$; H_2O: coordinates $x = 0{\cdot}427$, $y = 0{\cdot}833$, $z = 0{\cdot}644$. The other molecules shown in the figure are obtained from A and B by application of unit translations, which are indicated by indices. (b) [010] projection of the crystal structure.

bonds have also been reported in other, carefully investigated, structures, e.g. 1·19 Å in N-acetylglycine (Carpenter & Donohue, 1950), 1·21 Å in DL-alanine (Donohue, 1950), 1·21 and 1·20 Å in maleic acid (Shahat, 1952). This may indicate that a smaller value for the 'pure' C=O distance should be accepted.

For the distance C_{11}–N_2 (CH_2–NH_3^+) at the end of the chain we find a value of 1·50 Å, which is somewhat larger than the sum of the single-bond radii (1·47 Å) of carbon and nitrogen. This has also been observed in β-glycylglycine (Hughes & Moore, 1949), DL-alanine (Donohue, 1950), L-threonine (Shoemaker, Donohue, Schomaker & Corey, 1950), α-DL- and β-DL-methionine (Mathieson, 1952), L-glutamine (Cochran & Penfold, 1952), hexamethylenediamine (Binnie & Robertson, 1949) and hexamethylenediamine dihydrochloride (Binnie & Robertson, 1950).

The other bond lengths and the valence angles (Table 5) are in agreement with expectation. Short distances are observed between the following atoms belonging to the same molecule but not connected by a covalent bond:

$$N_2\text{–}O_4\ 2{\cdot}65\ \text{Å},\quad O_2\text{–}O_4\ 3{\cdot}12\ \text{Å}\quad \text{and}\quad O_3\text{–}C_4\ 3{\cdot}22\ \text{Å}$$

Arrangement of the molecules in the crystal; intermolecular distances

Fig. 6 shows the [001] and [010] projections of the structure. It is seen that the molecules are all parallel to the x axis. They are arranged in sheets parallel to (010), which consist alternatively of molecules A or molecules B. In each layer the molecules—strictly speaking the positive ions—are connected by the Cl^- ions, and via hydrogen bonds by the water molecules. Different layers are connected by OH · · · O and NH · · · O hydrogen bonds, and by the Cl^- ions.

On the other hand, the benzene rings are arranged in layers parallel to (100); these layers separate the regions which contain the dipole groups. The benzene rings of the molecules *A* are approximately perpendicular to those of the molecules *B*.

Most of the intermolecular connections are indicated in Fig. 6 by dashed or dotted lines. Dotted lines indicate bonds directed towards a molecule which is concealed behind a molecule shown in the figure. In Table 6 all intermolecular distances smaller than 3·6 Å are collected. Some of these will now be discussed.

Table 6. *Intermolecular distances smaller than* 3·6 Å

$Cl\ldots O_1(A_{100})$	3·04 Å	$H_2O\ldots C_{10}(A_{001})$	3·23 Å
$Cl\ldots N_2(A)$	3·24	$H_2O\ldots C_{11}(B_{10\bar{1}})$	3·25
$Cl\ldots N_2(A_{001})$	3·24	$H_2O\ldots C_{10}(B_{100})$	3·34
$Cl\ldots N_1(B_{100})$	3·33	$H_2O\ldots C_{11}(A_{001})$	3·45
$Cl\ldots O_4(A)$	3·57	$O_1(A)\ldots O_3(B_{0\bar{1}0})$	2·65
$H_2O\ldots O_4(A)$	2·87	$O_1(A)\ldots O_2(B_{0\bar{1}0})$	3·43
$H_2O\ldots O_2(A_{001})$	2·87	$O_2(A)\ldots N_2(B_{10\bar{1}})$	2·91
$H_2O\ldots N_2(B_{100})$	3·04	$O_3(A)\ldots C_4(B)$	3·35
$H_2O\ldots O_3(A)$	3·26	$O_3(A)\ldots C_6(B)$	3·42
$H_2O\ldots O_4(B_{100})$	3·34	$O_2(A)\ldots O_3(A_{00\bar{1}})$	3·35
$H_2O\ldots N_1(A_{001})$	3·45	$O_4(A)\ldots C_{11}(B_{10\bar{1}})$	3·58
$H_2O\ldots O_4(A_{001})$	3·51	$C_3(A)\ldots C_6(A_{00\bar{1}})$	3·39
$H_2O\ldots C_{11}(B_{100})$	3·07	$C_3(A)\ldots C_5(B_{00\bar{1}})$	3·58

(*a*) The Cl^- ion is surrounded by the atoms N_2 ($-NH_3^+$) of molecules *A* and A_{001}, by the oxygen atom O_1 (phenolic hydroxyl group) of A_{100} and by N_1 (amide nitrogen) of molecule B_{100}. The distances $Cl^- \cdots N_2$ are both 3·24 Å, which is practically equal to the sum of the ionic radii of Cl^- (1·81 Å) and NH_4^+ with 4-coordination (1·41 Å). The distance between Cl^- and O_1 (3·04 Å) and the angle C_1–$O_1 \cdots Cl^-$ (122°) indicate the formation of a hydrogen bond, such that the hydrogen atom of the phenolic hydroxyl group is directed approximately towards the Cl^- ion. Similar $Cl^- \cdots O$ distances (2·99 and 3·05 Å) were observed in hydroxylammonium chloride by Jerslev (1948).

(*b*) The intermolecular distances further suggest that each molecule of a given layer is connected by both an $OH \cdots O$ (2·65 Å) and a $NH \cdots O$ (2·91 Å) hydrogen bond to molecules of the next and to molecules of the preceding layer. The four hydrogen bonds starting from molecule *A* are $O_3(A) \cdots O_1(B)$, $O_2(A) \cdots N_2(B_{10\bar{1}})$, $O_1(A) \cdots O_3(B_{0\bar{1}0})$ and $N_2(A) \cdots O_2(B_{1\bar{1}\bar{1}})$. Since it was reasonable to assume that the O_1 hydrogen is directed towards Cl^-, the hydrogen atom involved in the bond $O_3 \cdots O_1$ must belong to O_3. The angle C_9–$O_3 \cdots O_1$ (119°), and the bond length C_9–O_3 (1·26 Å), which is larger than that of C_9–O_2 (1·21 Å), are compatible with this view.

Each nitrogen atom N_2 ($-NH_3^+$) is approximately tetrahedrally surrounded by C_{11}, two Cl^- ions and an oxygen atom O_2; probably the hydrogen atoms are directed towards the Cl^- ions and the oxygen atom.

(*c*) Considering the intermolecular distances of 2·87 Å between the water molecule and the atoms $O_4(A)$ and $O_2(A_{001})$ it may be assumed that the hydrogen atoms of the water molecule form two $OH \cdots O$ hydrogen bonds to glycyltyrosine molecules of one layer. These bonds make an angle of 123°.

The short distance (3·04 Å) between the water molecule and nitrogen atom N_2 ($-NH_3^+$) of molecule B_{100} may be caused by a relatively weak electrostatic interaction of the water dipole and the $-NH_3^+$ group.

The Patterson and Fourier syntheses were computed on punched-card machines. We wish to express our gratitude to Theodorus Niemeijer N. V. for putting their I. B. M. equipment generously at our disposal and to Mr M. R. van der Velde for his invaluable assistance and advice.

We are also indebted to Messrs L. Dijkema, H. Dijkhuizen and J. de Boer for their assistance during part of this structure determination.

The Netherlands Organization for Pure Research (Z. W. O.) supported indirectly part of this work.

References

BINNIE, W. P. & ROBERTSON, J. M. (1949). *Acta Cryst.* **2**, 180.

BINNIE, W. P. & ROBERTSON, J. M. (1950). *Acta Cryst.* **3**, 424.

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

BRAGG, W. L. & LIPSON, H. (1936). *Z. Krystallogr.* **95**, 323.

BROWN, C. J. (1951). *Acta Cryst.* **4**, 100.

BRUIGOM, E. S. (1950). Thesis, Leiden.

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

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

DYER, H. B. (1951). *Acta Cryst.* **4**, 42.

GOODWIN, T. H., PRZYBYLSKA, M. & ROBERTSON, J. M. (1950). *Acta Cryst.* **3**, 279.

HELMHOLZ, L. (1936). *J. Chem. Phys.* **4**, 316.

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

JERSLEV, B. (1948). *Acta Cryst.* **1**, 21.

MATHIESON, A. McL. (1952). *Acta Cryst.* **5**, 332.

PALIN, D. E. & POWELL, H. M. (1947). *J. Chem. Soc.* p. 208.

PAULING, L. (1940). *The Nature of the Chemical Bond.* Ithaca: Cornell University Press.

ROBERTSON, J. M. (1936). *Proc. Roy. Soc.* A, **157**, 79.

ROBERTSON, J. M. & UBBELOHDE, A. R. (1938). *Proc. Roy. Soc.* A, **167**, 122.

SHAHAT, M. (1952). *Acta Cryst.* **5**, 763.

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

SMITS, D. W. (1952). Thesis, Groningen.

SMITS, D. W. & WIEBENGA, E. H. (1953). *J. Sci. Instrum.* To be published.

TRANTER, T. C. (1952). *Acta Cryst.* **5**, 843.

WEISZ, O., COCHRAN, W. & COLE, W. F. (1948). *Acta Cryst.* **1**, 83.

WIEBENGA, E. H. (1952). *J. Amer. Chem. Soc.* **74**, 6165.

WIEBENGA, E. H. & SMITS, D. W. (1950). *Acta Cryst.* **3**, 265.

WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318.