Acta Cryst. (1963). **16**, 969

An Application of a New Phase Determination Procedure to the Structure of Cyclo(hexaglycyl) Hemihydrate

By Isabella L. Karle and J. Karle

U.S. Naval Research Laboratory, Washington 25, D.C., U.S.A.

(*Received* 1 *December* 1962)

The crystal structure of the synthetic polypeptide cyclo(hexaglycyl) hemihydrate, $(CH_2CONH)_6 . \frac{1}{2}H_2O$, has been determined by a simple, direct phase determination method in which letter symbols are assigned to initial phases. All the final phases are expressible in terms of a few of the initially assigned letters. The virtue of this procedure is that it is very simple, may be performed by hand computation, and is completely general for all centrosymmetric space groups.

The crystals of cyclo(hexaglycyl) are triclinic in space group $P\bar{1}$ with cell dimensions

$$a = 12{\cdot}31,\ b = 15{\cdot}50,\ c = 19{\cdot}04\ \text{Å};\ \alpha = 94° 58',\ \beta = 121° 09',\ \gamma = 88° 05'\ \text{and}\ Z = 8\ .$$

There is a strong pseudo-cell with $c/4$. The molecules are in the form of 18-membered rings with four distinctly different configurations in the unit cell. Hydrogen bonding stabilizes the structure. The water molecule is associated with only one-half of the cyclo(hexaglycyl) molecules.

Introduction

Cyclo(hexaglycyl) is a cyclic polypeptide which was synthesized by Ballard, Bamford & Weymouth (1955). The great stability of this compound to both acids and alkalis hindered the examination of partial hydrolysates by paper chromatography. Bamford & Weymouth (1955), however, favored the formation of an 18-membered ring. In a previous crystallographic examination (Cant, 1956), the space group and cell dimensions were determined and the material was shown to be a hemihydrate with the molecular formula $(CH_2CONH)_6 . \frac{1}{2}H_2O$.

The present investigation was undertaken at the suggestion of Prof. D. Wrinch, through whom the crystals were obtained. Our particular interest was both in the configuration of the polypeptide and in the application of phase determining procedures for the solution of the structure. This investigation affords the opportunity to illustrate a new general scheme for phase determination in centrosymmetric crystals.

Experimental measurements

Crystals of $(CH_2CONH)_6 . \frac{1}{2}H_2O$ are colorless, small, and needle-shaped; they crystallize in the triclinic system. The diffraction results show a pronounced pseudo-cell with the following parameters:

	Pseudo-cell		Real cell
Cant (1956)	$a = 10{\cdot}66$ Å	$\alpha = 94° 58'$	$4a$
	$b = 15{\cdot}50$ Å	$\beta = 98° 40'$	b
	$c = 4{\cdot}76$ Å	$\gamma = 90° 01'$	$4c$
	with $Z = 2$		$Z = 32$
	Pseudo-cell		Real cell
Present	$a = 12{\cdot}31$ Å	$\alpha = 94° 58'$	a
	$b = 15{\cdot}50$ Å	$\beta = 121° 09'$	b
	$c = 4{\cdot}76$ Å	$\gamma = 88° 05'$	$4c$
	with $Z = 2$		$Z = 8$

$\varrho_c = 1{\cdot}505$ g. cm^{-3}, $\varrho_o = 1{\cdot}501$ g.cm^{-3} (Cant, 1956).

The present a axis was chosen in the direction of the old a–c to give the smallest possible real cell.

Three-dimensional data were collected, employing Cu $K\alpha$ radiation and a Weissenberg camera, only along the c axis, the needle axis. The crystals were too small to permit the collection of data along the other axes. The intensities were estimated visually by means of a calibrated film strip. These data were punched on IBM cards directly from the primary data sheets and processed on the IBM 7090 computer, the data reduction program of Norment (1962) being used. The computer output yielded directly the scaled and averaged data in structure factor and normalized structure factor form. For this crystal, each layer (along the c axis) was treated separately.

The characteristics of the intensity data were as follows:

$l = 4n$, strong reflections,
$l = 2n-1$, weak reflections,
$l = 4n-2$, absent.

For the pseudo-cell ($l = 4n$), there were 2532 data (2097 non-zero data). The statistical averages for the normalized structure factors were

$$\langle |E| \rangle = 0{\cdot}776,\ \langle |E^2-1| \rangle = 0{\cdot}957,\ \langle E^2-1 \rangle = -0{\cdot}009\ ,$$

as compared with 0·798, 0·968, and 0, respectively, for a centrosymmetric crystal. Furthermore, piezoelectric tests were negative. Accordingly, the crystal was assigned to the $P\bar{1}$ space group.

The distribution of the normalized structure factors for the pseudo-cell was 0·3% for $E > 3$, 4·6% for $E > 2$ and 32·3% for $E > 1$. These values are almost identical with those computed for a crystal with randomly distributed atoms, *viz.* 0·3%, 5·0% and 32·0%.

Reflections with an odd l index were weak and only those with $l = 1$ and 3 were collected. There were 901 non-zero data in these two layers.

Phase determination

The phase determining procedure for this crystal was applied only to the reflections in the pseudo-cell since the reflections for the layers with $l \neq 4n$ were either weak or absent. A number of starting phases could be obtained from the magnitudes of the intensities by the application of Σ_1 (Hauptman & Karle, 1953) and formula $B_{3,0}$ (3·1·2) (Karle & Hauptman, 1959), as described for example for *N*-benzyl-1,4-dihydronicotinamide (Karle, 1961), and the phase determination could be completed with the application of Σ_2 defined by

$$sE_{\mathbf{h}} \sim s \sum_{k} E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}} \,. \qquad (1)$$

However an alternative procedure which appears to have the merit of simplicity, reliability and general applicability has also been applied to these data. This procedure avoids, for example, the large-scale computations on a high-speed computer which are required for an application of formula $B_{3,0}$. The specifications involve the assignment of phases in order to fix the origin (Hauptman & Karle, 1953) and in addition the assignment of letters to represent the signs of a few specially chosen structure factors. These letters are carried through the phase determining procedure, employing Σ_2, as unknowns. If there are p letters assigned, there will be at most 2^p Fourier maps to consider. However, relationships among the letters occur as the phase determination proceeds and it is generally necessary to consider only at most a very few Fourier maps. Information regarding some signs obtained with great certainty from probability methods, a knowledge of the number of exceptions to be expected from Σ_2 by application of the probability theory, and various chemical considerations such as the presence or absence of an atom at the origin delimit even further the number of Fourier maps to be considered.

This procedure has evolved as a consequence of several observations which we have made in the course of applying direct phase determination procedures. If the data from the copper sphere of radiation are used for which $hkl \neq 0$, there are a great many Σ_2 relationships whose probability of correctness may be reliably computed to be very high. Very few unknown letters need to be assigned in order to carry through a determination employing Σ_2. Also *E*-maps (which are three-dimensional electron density maps computed with *E* values rather than *F* values for coefficients) (Karle, Hauptman, Karle & Wing, 1958) can quite adequately reveal the structure when the phases for only 10% or less of the larger $|E|$ values are known.

Since the phase determination procedure for cyclo-(hexaglycyl) was applied only to the pseudo-cell, the indices were transformed to correspond to this cell, *i.e.* only those data with $l = 4n$ were used and the l index was divided by four. The origin of the pseudo-cell was specified by assigning signs to three linearly independent reflections. In addition, the signs of four other reflections were specified with a letter. These seven assignments which form the basic starting set for applying Σ_2 are shown in Table 1.

Table 1. *Assignment of three origin-specifying reflections and four other reflections as a starting set for the application of* Σ_2

Sign	*hkl*	*E*	
+	$\bar{9}$,2,3	3·06	for origin specification
+	1,1,2	3·18	
+	$1\bar{4}$,4,1	3·67	
q	$1\bar{3}$,6,2	2·77	
r	$\bar{9}$,4,0	3·62	
s	0,9,2	2·30	
t	9,3,0	2·61	

The reflections chosen to have signs q, r, s, and t were selected on the basis that they had many relationships applicable to the Σ_2 formula among themselves and with the origin specifying reflections. In addition, they all have large *E* magnitudes, giving very high probabilities that the indications from the Σ_2 formula were correct. Two hundred and seventy signs for $|E| > 1{\cdot}5$ were determined in this manner. In the course of application, as many Σ_2 combinations as possible were used to determine each sign. It was found from the relationships among the letters that r and t both had to be positive in order to form a consistent set. In addition, q and s could not both be positive, since in this case all the signs would be

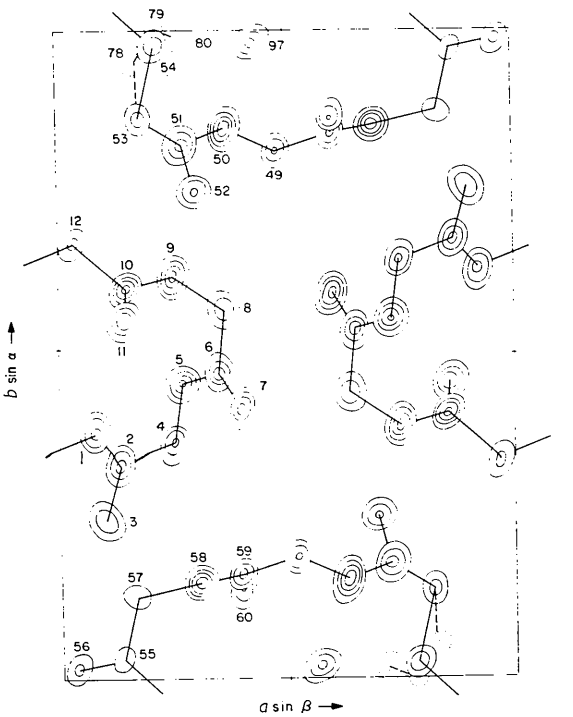

Fig. 1. Composite of the three-dimensional Fourier map for the pseudo-cell computed with 270 terms with *E* values as coefficients where $|E| > 1{\cdot}5$.

positive. Of the three remaining possibilities, $(q+, s-)$, $(q-, s+)$ and $(q-, s-)$, the first combination was preferred on the basis that it gave the fewest exceptions to Σ_2. Using this combination, a three-dimensional E-map (Fig. 1) was computed from 270 terms and the correct structure was readily obtained. A subsequent comparison of signs after refinement of the structure revealed that three of the 270 signs were incorrect.

Structure determination

An examination of the E-map for the pseudo-cell (Fig. 1) revealed all the atoms. The rapidly oscillating background seldom exceeded ± 15 (on an arbitrary scale). In one molecule the peak values for the atoms were all >40 and the average value was 70. In the other molecule, which appears disordered in the pseudo-cell (see dotted lines in Fig. 1), eight atoms had peak values of >40 with an average of 66, one was weaker with a value of about 30, and the six atoms involved in the 'disorder', which should have appeared about half-weight, had values between 30 and 40. No other peaks, except one for the H_2O molecule, appeared in the E-map. The coordinates of the atoms in the pseudo-cell as determined from the initial E-map are listed in Table 2.

Table 2. *The fractional coordinates of the atoms in the pseudo-cell as determined from the initial E-map*

In the last column the value of $z/4$ is listed so that it may be compared directly to the z coordinate of the real cell

Atom	x	y	z	$z/4$
N(1)	0·077	0·370	0·644	0·161
C(2)	0·123	0·322	0·890	0·222
O(3)	0·100	0·242	0·900	0·225
C(4)	0·250	0·355	0·217	0·304
N(5)	0·273	0·443	0·206	0·302
C(6)	0·355	0·460	0·122	0·281
O(7)	0·405	0·407	0·038	0·259
C(8)	0·366	0·560	0·099	0·275
N(9)	0·255	0·610	0·983	0·246
C(10)	0·146	0·595	0·679	0·170
O(11)	0·141	0·541	0·508	0·127
C(12)	0·027	0·663	0·627	0·157
C(49)	0·480	0·807	0·113	0·278
N(50)	0·374	0·843	0·874	0·219
C(51)	0·276	0·820	0·576	0·144
O(52)	0·308	0·750	0·475	0·119
C(53)	0·178	0·858	0·378	0·095
N(54)	0·216	0·978	0·416	0·104
C(55)	0·142	0·034	0·442	0·111
O(56)	0·047	0·008	0·380	0·095
C(57)	0·180	0·120	0·513	0·128
N(58)	0·308	0·145	0·541	0·135
C(59)	0·395	0·160	0·828	0·207
O(60)	0·400	0·141	0·133	0·283
N(78)	0·200	0·967	0·467	0·617*
C(79)	0·216	0	0·383	0·596*
O(80)	0·283	0·983	0·283	0·571*
O(97)	0·433	0·978	0·333	0·083
H_2O				

* These coordinates are $z/4+\frac{1}{2}$ in order to compare directly with the real cell.

The molecules are in the form of puckered eighteen-membered rings with planar peptide groups and they lie on centers of symmetry of the pseudo-cell. Several different configurations of the molecule exist in the unit cell. The H_2O molecule is associated with only one-half of the cyclo(hexaglycyl) molecules. In the real cell, where the c axis is four times as long as in the pseudo-cell, the molecule near $b=\frac{1}{2}$ (Fig. 1) merely repeats itself four times in the z direction as illustrated in Fig. 2. In the molecule near $b=0$, all atoms except N(54), N(78), C(55), C(79), O(56) and O(80) repeat themselves four times, to a good approximation, as illustrated in Fig. 3.

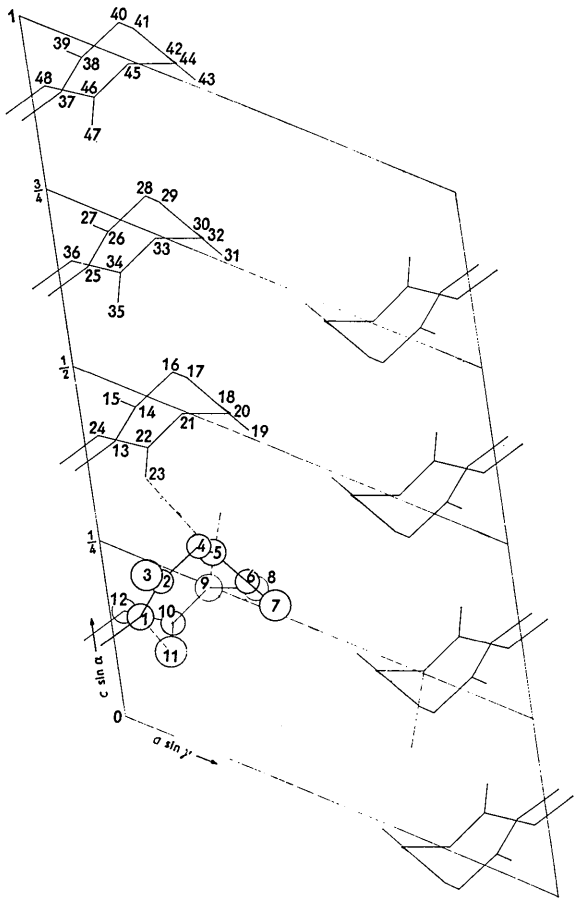

Fig. 2. Drawing of the real cell showing the positions of the molecules near $b=\frac{1}{2}$. The dotted lines indicate hydrogen bonds in the first quarter of the cell.

The H_2O molecule and atoms N(54), N(78), C(55), C(79), O(56) and O(80) are the ones mainly involved in the apparent 'disorder' in the pseudo-cell. Their positions in the real cell were deduced from the distribution of the intensities as a function of the l index. The condition that all reflexions with $l=4n-2$ be absent, and those with $l=2n-1$ be weak is satisfied by the set of coordinates $\pm(x, y, z)$ and $\pm(x, y, z+\frac{1}{4})$ for atoms involved in the 'disorder'. The additional

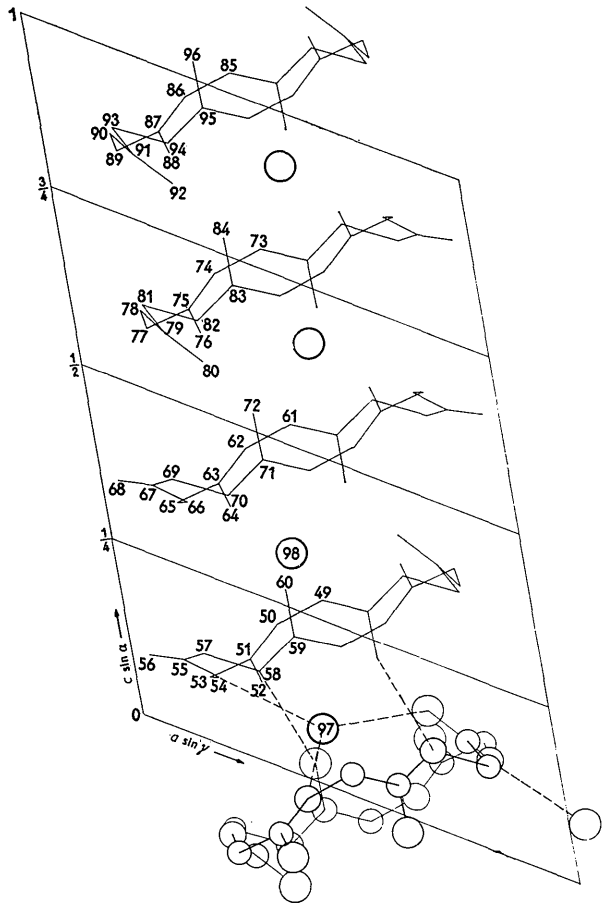

Fig. 3. Drawing of the real cell showing the configurations and positions of the molecules near $b=0$. The dotted lines indicate hydrogen bonds in the first quarter of the cell.

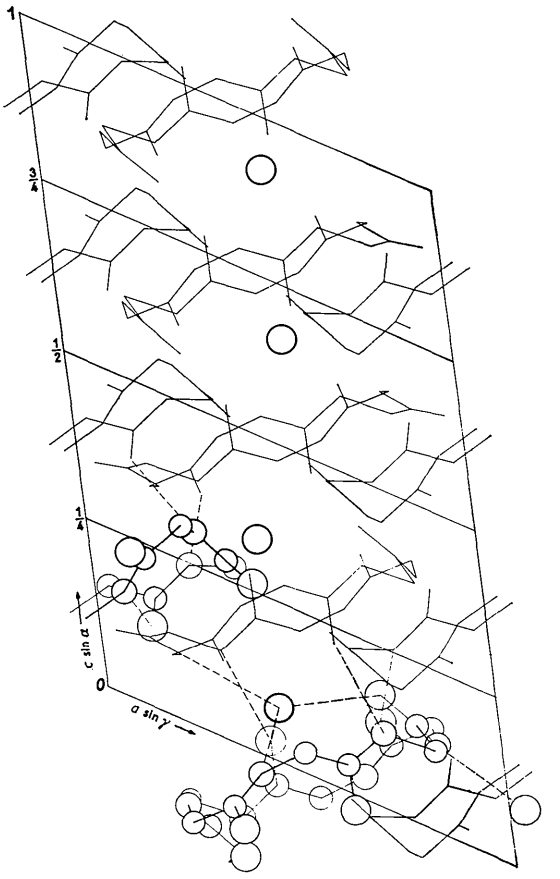

Fig. 4. Drawing of the real cell showing the complete contents.

space group limitation that centers of symmetry exist at 0, 0, 0; $\frac{1}{2}$, 0, 0; 0, 0, $\frac{1}{2}$ and $\frac{1}{2}$, 0, $\frac{1}{2}$ (for the molecules near $b=0$), led to the arrangement shown in Fig. 3. Three different configurations of the molecules near $b=0$ are necessary. The uppermost molecule (lying on a center of symmetry) has the two end peptide groups pointing inwards. The next molecule has one end peptide group pointing inwards and the other outwards (no symmetry center is required for this molecule). The central molecule (also lying on a center of symmetry) has both end peptide groups pointing outwards. With this arrangement, all intermolecular distances are >3 Å, except where hydrogen bonding takes place.

Refinement

A least-squares refinement of the pseudo-cell was performed on 28 atoms, with one-half weight assigned to atoms N(54), N(78), C(55), C(79), O(56), O(80) and O(97) (the water oxygen). After three cycles of Busing & Levy's (1959) program with isotropic temperature factors, the R index for the 2097 non-zero data for the pseudo-cell was 16·5%. The bond distances in the molecule near $b=\frac{1}{2}$ all seemed to refine properly. However, in the molecule near $b=0$, some of the bond distances in the 'disordered' part of the molecule did not appear to refine, possibly because of the close proximity of the 'disordered' atoms to each other or because the other atoms did not repeat themselves exactly in a fourfold fashion in the real cell.

Next the coordinates of all 98 atoms in the real cell were refined using the observed data for the weak odd layers as well as for the layers with $l=4n$. The least-squares program for the IBM 7090, as revised by Van den Hende (1961), had the capacity for this large number of atoms and data. The refinement proceeded at a slow rate, about 0·5% improvement per cycle. The large number of weak data and the large number of parameters are no doubt responsible for the slow convergence. At $R=18·2\%$ for all the non-zero data, the coordinates still showed appreciable changes. It was decided at this point to report the approximate structure and to defer the more complete refinement to a future time.

The fractional coordinates and the isotropic tem-

Table 3. *The fractional coordinates and isotropic temperature factors of the atoms in the real cell of cyclo(hexaglycyl)*

The structure has not been completely refined

Atom	x	y	z	B (Å^2)	Atom	x	y	z	B (Å^2)
N(1)	0·0793	0·3709	0·1614	4·1	C(49)	0·4794	0·7965	0·2792	4·1
C(2)	0·1339	0·3209	0·2285	4·1	N(50)	0·3671	0·8460	0·2192	3·8
O(3)	0·0955	0·2485	0·2276	6·5	C(51)	0·2832	0·8154	0·1432	4·2
C(4)	0·2510	0·3578	0·3050	4·0	O(52)	0·2941	0·7417	0·1188	6·5
N(5)	0·2744	0·4481	0·3012	3·5	C(53)	0·1750	0·8760	0·0811	4·0
C(6)	0·3510	0·4667	0·2753	3·1	N(54)	0·2177	0·9729	0·1067	2·5
O(7)	0·4066	0·4088	0·2577	4·6	C(55)	0·1370	0·0318	0·1136	3·5
C(8)	0·3707	0·5628	0·2696	4·2	O(56)	0·0280	0·0150	0·1029	6·1
N(9)	0·2570	0·6135	0·2518	3·7	C(57)	0·1725	0·1251	0·1274	5·0
C(10)	0·1524	0·5984	0·1756	3·9	N(58)	0·3010	0·1467	0·1350	3·4
O(11)	0·1385	0·5394	0·1254	3·9	C(59)	0·3999	0·1569	0·2114	3·7
C(12)	0·0460	0·6681	0·1609	4·7	O(60)	0·4006	0·1343	0·2719	5·0
N(13)	0·0755	0·3677	0·4121	3·7	C(61)	0·4813	0·8038	0·5297	4·4
C(14)	0·1319	0·3201	0·4773	4·8	N(62)	0·3667	0·8520	0·4695	4·1
O(15)	0·9001	0·2489	0·4782	6·3	C(63)	0·2889	0·8192	0·3949	4·2
C(16)	0·2492	0·3569	0·5556	4·1	O(64)	0·2973	0·7438	0·3688	6·2
N(17)	0·2703	0·4440	0·5515	3·9	C(65)	0·1772	0·8804	0·3316	3·7
C(18)	0·3522	0·4659	0·5261	3·7	N(66)	0·2195	0·9756	0·3554	2·9
O(19)	0·4073	0·4087	0·5082	5·3	C(67)	0·1408	0·0335	0·3606	3·9
C(20)	0·3714	0·5592	0·5204	4·6	O(68)	0·0313	0·0160	0·3498	6·4
N(21)	0·2580	0·6120	0·5013	3·3	C(69)	0·1744	0·1252	0·3759	4·3
C(22)	0·1527	0·5980	0·4252	3·5	N(70)	0·3011	0·1469	0·3850	3·5
O(23)	0·1395	0·5393	0·3759	3·8	C(71)	0·4023	0·1593	0·4626	3·2
C(24)	0·0488	0·6678	0·4114	5·2	O(72)	0·4008	0·1402	0·5227	5·2
N(25)	0·0712	0·3607	0·6639	3·4	C(73)	0·4827	0·8027	0·7789	4·5
C(26)	0·1303	0·3154	0·7275	4·3	N(74)	0·3664	0·8515	0·7200	3·8
O(27)	0·0922	0·2428	0·7280	6·8	C(75)	0·2921	0·8148	0·6449	4·3
C(28)	0·2472	0·3527	0·8071	4·5	O(76)	0·3011	0·7419	0·6184	5·9
N(29)	0·2689	0·4406	0·8013	3·7	C(77)	0·1698	0·8597	0·5893	5·0
C(30)	0·3484	0·4620	0·7754	4·1	N(78)	0·1642	0·9453	0·6145	4·2
O(31)	0·4062	0·4068	0·7569	5·3	C(79)	0·2211	0·0081	0·5965	2·9
C(32)	0·3680	0·5554	0·7724	4·1	O(80)	0·2954	0·9790	0·5765	5·1
N(33)	0·2529	0·6082	0·7513	3·3	C(81)	0·1921	0·1038	0·6136	4·0
C(34)	0·1499	0·5976	0·6766	3·4	N(82)	0·3038	0·1472	0·6344	2·9
O(35)	0·1350	0·5359	0·6267	4·0	C(83)	0·4026	0·1582	0·7121	3·0
C(36)	0·0414	0·6600	0·6635	5·1	O(84)	0·4016	0·1413	0·7727	5·2
N(37)	0·0745	0·3638	0·9133	4·1	C(85)	0·4806	0·7955	0·0284	4·1
C(38)	0·1322	0·3163	0·9786	3·8	N(86)	0·3670	0·8459	0·9696	3·4
O(39)	0·0979	0·2423	0·9773	6·9	C(87)	0·2869	0·8115	0·8932	4·0
C(40)	0·2491	0·3539	0·0564	4·4	O(88)	0·2979	0·7397	0·8684	6·3
N(41)	0·2731	0·4447	0·0509	3·5	C(89)	0·1685	0·8551	0·8380	4·3
C(42)	0·3472	0·4632	0·0246	3·5	N(90)	0·1664	0·9461	0·8640	3·8
O(43)	0·4055	0·4069	0·0065	4·6	C(91)	0·2177	0·0052	0·8489	3·9
C(44)	0·3674	0·5590	0·0215	3·7	O(92)	0·2985	0·9790	0·8238	4·6
N(45)	0·2516	0·6099	0·0018	3·8	C(93)	0·1916	0·1025	0·8627	4·5
C(46)	0·1495	0·5980	0·9270	3·9	N(94)	0·3036	0·1469	0·8845	2·9
O(47)	0·1340	0·5359	0·8762	4·1	C(95)	0·4002	0·1558	0·9609	3·5
C(48)	0·0390	0·6608	0·9129	4·6	O(96)	0·4014	0·1355	0·0218	5·1
					O(97)	0·4375	0·9815	0·0960	4·9
					O(98)	0·4387	0·9799	0·3454	5·0

perature factors of the 98 atoms in the asymmetric unit of the real cell are listed in Table 3.

Discussion of the structure

In view of the relatively large uncertainties in the atomic parameters, there is little point in discussing individual bond distances. Accordingly, only the averages of similar bonds and angles for each group of four molecules (those near $b=0$ and those near $b=\frac{1}{2}$) are listed in Table 4. The notation (N(1′)–C(2′) for example means the average of the bond lengths N(1)–C(2), N(13)–C(14), N(25)–C(26) and N(37)–C(38), where these groups of atoms are approximately related by a shift of $c/4$. Bond distances and angles for the 24 different peptide groups were averaged to obtain a composite peptide group which is illustrated in Fig. 5. The values are in good agreement with those found in other peptides. All the peptide groups are approximately planar with an average deviation of 0·03 Å. Presumably this deviation will be decreased with further refinement. The dihedral angles between planes through adjacent peptide groups range from 90° to 116°.

The molecules are held together by a network of hydrogen bonds. In the set of molecules near $b=\frac{1}{2}$,

Table 4. *Average bond distances and angles*

Distance		Angle	
N(1′)–C(2′)	1·34 Å	C′(12′)–N(1′)–C(2′)	121°
C(2′)–O(3′)	1·23	N(1′)–C(2′)–O(3′)	123
C(2′)–C(4′)	1·52	N(1′)–C(2′)–C(4′)	119
C(4′)–N(5′)	1·44	O(3′)–C(2′)–C(4′)	118
		C(2′)–C(4′)–N(5′)	112
N(5′)–C(6′)	1·34	C(4′)–N(5′)–C(6′)	120
C(6′)–O(7′)	1·24	N(5′)–C(6′)–O(7′)	121
C(6′)–C(8′)	1·52	N(5′)–C(6′)–C(8′)	118
C(8′)–N(9′)	1·48	O(7′)–C(6′)–C(8′)	121
		C(6′)–C(8′)–N(9′)	112
N(9′)–C(10′)	1·34	C(8′)–N(9′)–C(10′)	118
C(10′)–O(11′)	1·22	N(9′)–C(10′)–O(11′)	125
C(10′)–C(12′)	1·57	N(9′)–C(10′)–C(12′)	112
C(12′)–N′(1′)	1·47	O(11′)–C(10′)–C(12′)	123
		C(10′)–C(12′)–N′(1′)	109
C(49′)–N(50′)	1·50	C′(59′)–C(49′)–N(50′)	107
N(50′)–C(51′)	1·32	C(49′)–N(50′)–C(51′)	121
C(51′)–O(52′)	1·22	N(50′)–C(51′)–O(52′)	124
C(51′)–C(53′)	1·55	N(50′)–C(51′)–C(53′)	119
		O(52′)–C(51′)–C(53′)	117
C(53′)–N(54′)	1·48	C(51′)–C(53′)–N(54′)	112
N(54′)–C(55′)	1·34	C(53′)–N(54′)–C(55′)	120
C(55′)–O(56′)	1·28	N(54′)–C(55′)–O(56′)	121
C(55′)–C(57′)	1·52	N(54′)–C(55′)–C(57′)	118
		O(56′)–C(55′)–C(57′)	121
C(57′)–N(58′)	1·47	C(55′)–C(57′)–N(58′)	110
N(58′)–C(59′)	1·34	C(57′)–N(58′)–C(59′)	119
C(59′)–O(60′)	1·22	N(58′)–C(59′)–O(60′)	124
C(59′)–C′(49′)	1·54	N(58′)–C(59′)–C′(49′)	115
		O(60′)–C(59′)–C′(49′)	120

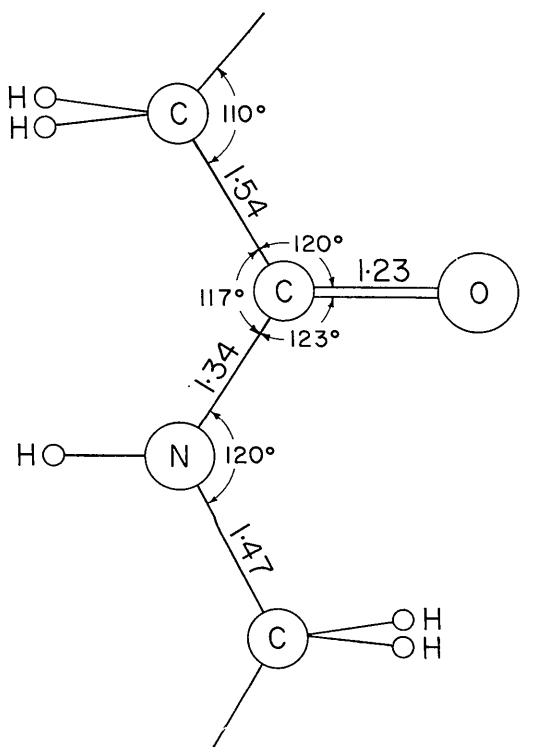

Fig. 5. Composite peptide group formed by averaging bond distances and angles for the twenty-four peptide groups in the asymmetric unit.

each molecule has two internal hydrogen bonds; see for example N(1) · · · O(11). On the other hand, the set of molecules near $b=0$ does not have any internal bonds, but four bonds are formed to each interleaving H_2O molecule, two with N–H groups and two with C=O groups, such as N(54) · · · O(97), N(86) · · · O(97), O(96) · · · O(97) and O′(92) · · · O(97). In addition, hydrogen bonds between N–H and the carbonyl oxygens are formed between adjacent molecules in the *a*, *b* and *c* directions, as for example, N(5) · · · O(23), N(9) · · · O(64), N(58) · · · O(96) and N′(90) · · · O(56). Each hydrogen attached to a nitrogen and each hydrogen in the water molecules is involved in hydrogen bonding. Eight carbonyl oxygens in the molecules near $b=\frac{1}{2}$ are not involved in hydrogen bonding. They are O(3), O(7), O(15), O(19), O(27), O(31), 0(39) and O(43). The distances and angles for representative hydrogen bonds are listed in Table 5. These values are similar to those found in other peptides (see for example, Fuller, 1959).

Table 5. *Representative hydrogen-bond distances and angles illustrated in Figs. 2–4*

There are several other sets of hydrogen bonds in the complete cell very similar to these

N(1) · · · O(11)	2·96 Å	C(2)–N(1) · · · O(11)	136°
		C′(48)–N(1) · · · O(11)	103
N(5) · · · O(23)	2·95	C(4)–N(5) · · · O(23)	102
		C(6)–N(5) · · · O(23)	137
N(9) · · · O(64)	2·73	C(8)–N(9) · · · O(64)	114
		C(10)–N(9) · · · O(64)	127
N(58) · · · O(96)	2·98	C(57)–N(58) · · · O(96)	135
		C(59)–N(58) · · · O(96)	107
N′(90) · · · O(56)	2·81	C′(89)–N′(90) · · · O(56)	113
		C′(91)–N′(90) · · · O(56)	120
N(54) · · · O(97)	2·83	C(53)–N(54) · · · O(97)	104
		C(55)–N(54) · · · O(97)	136
N(86) · · · O(97)	2·84	C(87)–N(86) · · · O(97)	150
		C(85)–N(86) · · · O(97)	93
O(97) · · · O(96)	2·91	O(96) · · · O(97) · · · O′(92)	83
O(97) · · · O′(92)	2·85		

The last point of interest is the existence of conformational isomers. The molecules being in the form of 18-membered rings exhibit a certain amount of freedom in the ring configuration while preserving planar peptide groups. There are, in fact, four distinctly different configurations of cyclo(hexaglycyl) in each unit cell. Conformational isomers have been previously observed for such molecules as cyclononylamine hydrobromide (Bryan & Dunitz, 1960) which also has a fairly large ring.

The existence of conformational isomers, intramolecular hydrogen bonding and the occurrence of carbonyl oxygens which do not participate in hydrogen bonding make it rather difficult to predict the structure of a polypeptide.

The authors are indebted to Dr Dan Appelman of the U.S. Geological Survey and Mr Paul Smith of

the U.S. Naval Research Laboratory for performing the piezoelectric tests. We also thank Dr Hillyer Norment and Mr Peter Gum for their invaluable aid in facilitating the calculations for this paper.

References

BALLARD, D. G. H., BAMFORD, C. H. & WEYMOUTH, F. J. (1955). *Proc. Roy. Soc.* A, **227**, 155.
BAMFORD, C. H. & WEYMOUTH, F. J. (1955). *J. Amer. Chem. Soc.* **77**, 6368.
BRYAN, R. F. & DUNITZ, J. D. (1960). *Helv. Chim. Acta*, **43**, 3.
BUSING, W. R. & LEVY, H. A. (1959). *A Crystallographic Least Squares Refinement Program for the IBM* 704. Oak Ridge, Tennessee: Oak Ridge National Laboratory.
CANT, E. M. (1956). *Acta Cryst.* **9**, 681.
FULLER, W. (1959). *J. Phys. Chem.* **63**, 1705.
HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal.* A. C. A. Monograph No. 3. Brooklyn: Polycrystal Book Service.
KARLE, I. L., HAUPTMAN, H., KARLE, J. & WING, A. B. (1958). *Acta Cryst.* **11**, 257.
KARLE, I. L. (1961). *Acta Cryst.* **14**, 497.
KARLE, J. & HAUPTMAN, H. (1959). *Acta Cryst.* **12**, 404.
NORMENT, H. G. (1962). U.S. Naval Research Laboratory Report No. 5739. Washington, D.C., U.S.A.
VAN DEN HENDE, J. H. (1961). *Crystallographic Structure Factor and Least Square Refinement Program for the IBM* 7090 *Computer*. Linden, New Jersey: Esso Research & Engineering Company.

Acta Cryst. (1963). **16**, 975

The Crystal Structure of Potassium Tetraborate Tetrahydrate

BY MASSIMO MAREZIO, H. A. PLETTINGER AND W. H. ZACHARIASEN

Department of Physics, University of Chicago and Argonne National Laboratory, U.S.A.

(*Received* 28 *December* 1962)

The compound $K_2B_4O_7.4H_2O = K_2[B_4O_5(OH)_4].2H_2O$ is orthorhombic, space group $P2_12_12_1$, with four molecules in a cell of dimensions $a=12{\cdot}899$, $b=11{\cdot}774$, $c=6{\cdot}859$ Å. The positions of all atoms (including the hydrogen atoms) have been determined. The final R-index is 0·040, and the bond lengths have been determined to an accuracy of 0·006 Å for K–O bonds, 0·008 Å for B–O bonds and 0·11 Å for H–O bonds. The structure contains the complex ion $[B_4O_5(OH)_4]^{2-}$ consisting of two boron–oxygen triangles and two tetrahedra which have only corners in common. Detailed results are reported for bond distances, and bond angles and for the anisotropic thermal motions.

Introduction

This paper reports the results of a crystal structure examination of the compound $K_2B_4O_7.4H_2O = K_2[B_4O_5(OH)_4].2H_2O$. The investigation is part of a systematic study of borate structures which was begun in this laboratory in the nineteen thirties, was interrupted by the war, and has been resumed in recent years.

Single crystals were prepared by slow evaporation from aqueous solutions. The formula was confirmed by chemical analysis which gave the following result:

	Exp.	Theor.
K	24·4%	25·6%
B	14·7	14·2
H_2O	21·2	23·6

A direct density determination gave $\varrho=1{\cdot}919$ g.cm^{-3}.

The crystals were found to be orthorhombic with cell dimensions

$$a=12{\cdot}899\pm0{\cdot}002,\ b=11{\cdot}774\pm0{\cdot}002,$$
$$c=6{\cdot}859\pm0{\cdot}001\ \text{Å}.$$

The cell contains four molecules, the calculated density being 1·898 g.cm^{-3}.

The only systematic absences are the odd orders of pinacoidal reflections. Accordingly one is led to the space group symmetry $P2_12_12_1$ with all atoms in general positions:

$$(x, y, z)\,(\tfrac{1}{2}-x, \bar{y}, \tfrac{1}{2}+z)\,(\tfrac{1}{2}+x, \tfrac{1}{2}-y, \bar{z})\,(\bar{x}, \tfrac{1}{2}+y, \tfrac{1}{2}-z)$$

and 75 positional degrees of freedom.

With the aid of a counter spectrometer and Cu $K\alpha$ radiation the intensities were measured for all (867) reflections $HK0$, $HK1$, $H0L$, $H1L$, $0KL$, $1KL$. The specimen was a crystal ground into a nearly perfect sphere of radius $R=0{\cdot}0259\pm0{\cdot}0015$ cm.

Determination of the structure

According to the space group symmetry $P2_12_12_1$ the projection of the structure on each of the coordinate planes has a center of symmetry. Because of the relatively short c period, it was to be expected that the projection on the XY plane would give the best