Acta Cryst. (1952). **5**, 654

The Crystal and Molecular Structure of α-Glucose

By T. R. R. McDonald* and C. A. Beevers

Dewar Crystallographic Laboratory, Edinburgh University, Scotland

(*Received* 16 *February* 1952)

The crystal structure of α-glucose has been fully determined. In the elucidation of the structure, extensive use was made of the two-dimensional Patterson function. The refinement of the atomic parameters was carried out by Fourier and least-squares methods.

The stereochemical configuration of the molecule, as deduced on purely chemical grounds, has been directly confirmed. The pyranose ring is in the Sachse–Mohr *trans* form, and there is a *cis*-glycol grouping on C_1 and C_2. Some attempt was made to determine the interatomic bond distances with precision; there would appear to be a shortening of the glycosidic and primary hydroxyl bonds compared with the results obtained for C–O bonds in other investigations.

Introduction

The object of this investigation was to provide accurate information about the stereochemical configuration and molecular dimensions of α-glucose, in the hope that such information might provide a basis for an explanation of the chemical properties of the sugars in general. To date, only three compounds containing a sugar residue have been studied in any detail by X-ray methods, viz: glucosamine hydrobromide (Cox & Jeffrey, 1939), sucrose sodium bromide dihydrate (Beevers & Cochran, 1947), and cytidine (Furberg, 1950).

A preliminary account of our work has already appeared (McDonald & Beevers, 1950). In the present paper, an account will be given of the methods employed for obtaining the atomic parameters with a higher degree of accuracy. We would also draw attention to the use which has been made of the two-dimensional Patterson function in the determination of the structure. The method employed is generally applicable to crystal structures which have an apparent centre of symmetry in each projection, and which are built up of molecules which approximate to centro-symmetry, and should therefore prove useful in elucidating the structures of other pyranose sugars.

Experimental

(*a*) *Crystal data*

α-Glucose crystallizes from methyl alcohol as fine needles parallel to the *c* axis of the unit cell. The crystal habit is described by Groth (1906–19, vol. 3, p. 437). Crystal class: orthorhombic bisphenoidal. $a:b:c = 0{\cdot}704:1:0{\cdot}335$. Observed density, 1·56 g.cm.$^{-3}$. The unit-cell dimensions, obtained by the extrapolation method of Farquhar & Lipson (1946), are:

$$a=10{\cdot}36\pm0{\cdot}02,\quad b=14{\cdot}84\pm0{\cdot}02,\quad c=4{\cdot}97\pm0{\cdot}02\ \text{Å}\,.$$

These are in good agreement with the values obtained by previous workers: Hengstenberg & Mark (1930) $a=10{\cdot}40$, $b=14{\cdot}84$, $c=4{\cdot}99$ Å; Sponsler & Dore (1931) $a=10{\cdot}45$, $b=14{\cdot}85$, $c=4{\cdot}97$ Å. The axial ratios $a:b:c=0{\cdot}698:1:0{\cdot}335$ agree with those given by Groth.

The density calculated on the basis of four molecules per unit cell is 1·56 g.cm.$^{-3}$.

Systematic absences: $h00$, $0k0$, $00l$ for h, k, l odd. Space group $P2_12_12_1$.

(*b*) *Intensity measurements*

The reflexion intensities were obtained from Weissenberg and oscillation photographs taken about the three principal axes. A crystal about 1 mm. in length with a cross-section of 0·1 × 0·1 mm. was used for the *c*-axis photographs. The specimens used for the *a*- and *b*-axis photographs were 0·5 mm. in length and had approximately square cross-sections 0·3 × 0·3 mm. The intensities were measured by visual comparison with standard scales prepared by exposing the crystal to the X-ray beam for known periods of time. After the application of Lorentz, polarisation and Cox & Shaw factors, the intensities were brought on to the same scale by direct comparison of common spots. The method of Wilson (1942) was then used to bring the relative F^2 on to an approximately absolute scale. The F_o were later rescaled by direct comparison with F_c.

Determination of the approximate structure

The crystal structure was determined by the use of the Patterson projections. A preliminary survey of the reflexion intensities indicated that the plane of the molecules was very nearly perpendicular to the short *c* axis of the unit cell, a conclusion which was confirmed by a study of the three Patterson projections. The *a*- and *b*-axis projections (Figs. 1 and 2) indicate a layer structure parallel to *a b*, while from the *c*-axis projection (Fig. 3) the orientation of the molecules

---

* Now at Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England.

within the layers can be obtained. The circle of peaks round the origin of the latter is readily interpreted on the basis of a centro-symmetrical model of the pyranose ring, each peak corresponding to the superposition of several vectors of length 2·5 Å, the vector distance between alternate atoms of the ring.

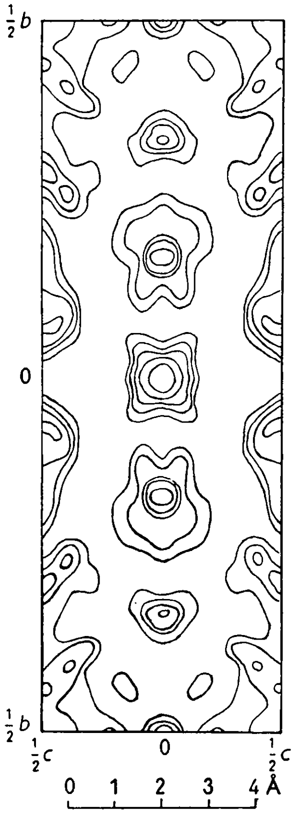

Fig. 1.

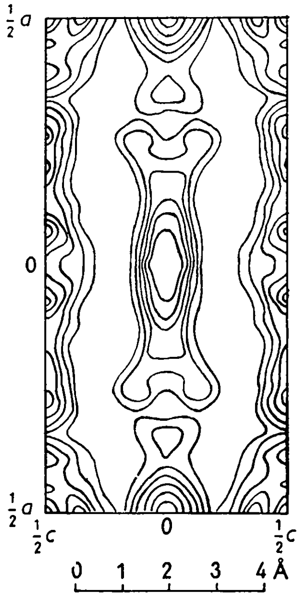

Fig. 2.

Fig. 1. *a*-axis Patterson projection.
Fig. 2. *b*-axis Patterson projection.

Locating the molecule in the unit cell proved to be a more difficult problem, but it was eventually solved by the following simple extension of Harker's principle (Harker, 1936) to the case of a centro-symmetrical molecule. A plane centro-symmetrical molecule (lower part of Fig. 4), when operated upon by an external centre of symmetry, gives rise to another molecule which is identical with the first as far as the orientation is concerned, and the vectors between the atoms of molecule (1) and their counterparts in molecule (2) are all parallel and equal in length to the vector distance between the molecular centres. Extending this principle to three dimensions, consider a centro-symmetrical molecule operated upon by a screw axis (upper part of Fig. 4).

The vectors between corresponding atoms, though not necessarily equal and parallel in space, will have equal and parallel components of projection on the plane perpendicular to the screw axis. As there are three mutually perpendicular screw axes in the space group $P2_12_12_1$, and since the glucose molecule approximates to centro-symmetry, we should expect to find one particularly strong peak on each of the three Patterson projections due to the coincidence of vectors of this type. It should be immediately possible, from

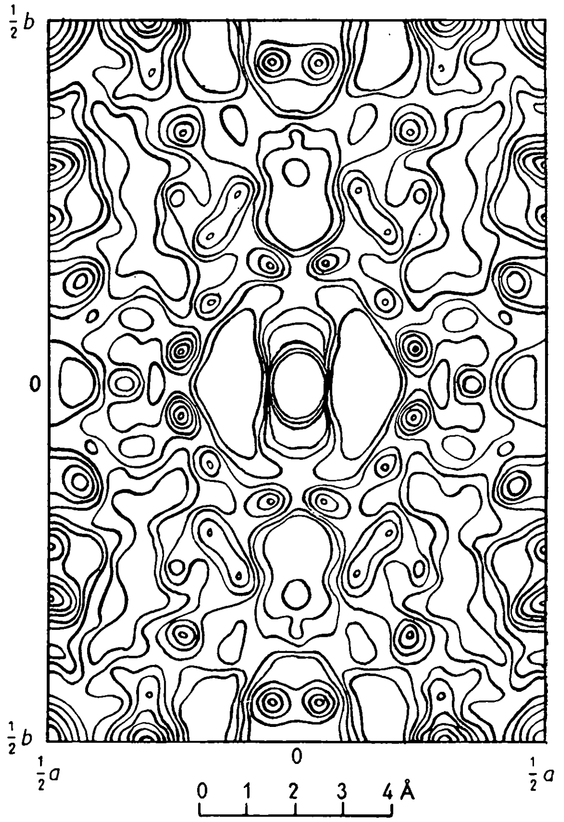

Fig. 3. *c*-axis Patterson projection.

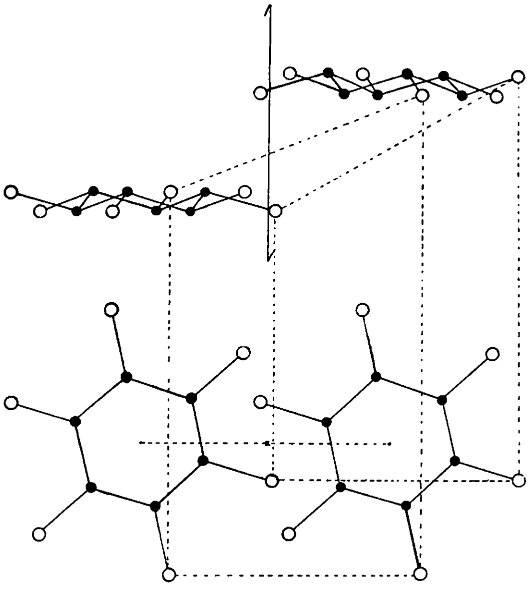

Fig. 4. Vectors between two centro-symmetrical molecules related by a screw axis (in space and in projection).

the positions of these peaks, to derive the coordinates of the centre of the molecule, for if the latter are x, y, z, the arrangement of screw axes in the space group demands that the peaks occur at the following positions in the three projections

a-axis	—	$\frac{1}{2}-2y$	$2z$
b-axis	$2x$	—	$\frac{1}{2}-2z$
c-axis	$\frac{1}{2}-2x$	$2y$	—

Though there were no outstandingly strong peaks on any of the projections, nevertheless it was a simple matter to find a unique set of three peaks with coordinates satisfying these conditions. Their coordinates are (in 60ths of the unit cell edges)

a-axis	—	$y=21$	$z=0$
b-axis	$x=4$	—	$z=30$
c-axis	$x=26$	$y=9$	—

placing the centre of the molecule at

$$x=2,\ y=4\tfrac{1}{2},\ z=0 \text{ 60ths} .$$

In a recent paper, Patterson (1949) has suggested the use of the series

$$\Sigma F^2(h_i) \exp\left[-2\pi i \Sigma h_i u_i\right]$$

for detecting centres of symmetry or pseudo-symmetry in crystal structures. In the present case, series of this type are identical with the $|F|^2$ series, and the analysis affords an interesting example of this alternative way of interpreting Patterson diagrams, strong peaks being regarded as corresponding to positions of approximate centro-symmetry in the structure.

Having determined the orientation and position of the symmetrical model, it still remained to introduce the element of asymmetry present in the real molecule. This was done by comparison of observed and calculated structure amplitudes, attempts at obtaining this information from the Patterson function meeting with no success.

Refinement of parameters

(*a*) *Fourier refinement*

The Fourier refinement of the x and y parameters proceeded along the usual lines, the value of $R_{hk0} = \Sigma||F_o|-|F_c|| \div \Sigma F_o$ being reduced to 0·30 after five successive approximations. When this stage was reached, no further shifts were indicated, and the refinement process was discontinued. Conditions were particularly favourable for the use of the Fourier method of refinement in this zone, all but two of the atoms being clearly resolved (see Fig. 2 of the previous communication (McDonald & Beevers, 1950)).

(*b*) *Least-squares refinement*

For the refinement of the $(0kl)$ and $(h0l)$ zones, use was made of the method of least-squares (Hughes, 1941) which has certain advantages over the ordinary Fourier method, at least when applied to poorly resolved projections.

The usual observational equations were set up, one for each F:

$$\left(\frac{\partial F_i}{\partial x}\right)_1 \delta x_1 + \left(\frac{\partial F_i}{\partial x}\right)_2 \delta x_2 + \ldots + \left(\frac{\partial F_i}{\partial x}\right)_n \delta x_n - \delta F_i = 0 \ ,$$

the δx_r being the required shifts in the atomic parameters, and $\delta F_i = F_o - F_c$.

The most tedious part of the work in a least-squares refinement is the normalization of the observational equations. It is found in practice, however, that at no stage in the refinement is it necessary to normalize the equations completely, for if a very large number of equations are set up, the off-diagonal terms of the normalized matrix can, to a first approximation, be neglected, and the solutions can be obtained directly from the diagonal coefficients and the constant terms. The solutions then reduce to

$$\delta x_r = \frac{\sum\limits_{i=1}^{m} \left(\frac{\partial F_i}{\partial x}\right)_r \delta F_i}{\sum\limits_{i=1}^{m} \left(\frac{\partial F_i}{\partial x}\right)_r^2} \ .$$

The use of these modified expressions, which have been suggested by different authors (Cruickshank, 1950; Qurashi, 1949), is necessary if the work is to be made at all practicable. The approximation results in little loss of accuracy, at least in three-dimensional work, as has recently been demonstrated by Shoemaker, Donohue, Schomaker & Corey (1950). Indeed, these authors suggest that in future work with the three-dimensional least-squares method, the off-diagonal coefficients may be completely neglected. It must be observed, however, that the approximation is not valid if two atoms overlap in projection. In such cases, the shifts obtained for these atoms are unreliable and should be determined from projections in which the atoms are free from overlap. In order to increase the validity of the approximation, the zones were combined in pairs, making some two or three hundred equations available for the determination of each shift. These simplified expressions involve the coefficients of the δx_r only; an immediate and important consequence of this is that the variables become virtually independent of one another and this means that, in practice, for purposes of refinement, the atoms can be treated separately.

In the initial stage of the refinement, a much more approximate method was used. In setting up the observational equations, only those F's were employed which showed the worst discrepancies between observed and calculated values. The equations were fully normalized, and then solved by an iterative process. This method, which is relatively rapid, tended to give shifts which were too large, but the atomic parameters were found to converge rapidly to their final values on repeated application of the method. The state of refinement ultimately reached by the use of these methods can be indicated by quoting the value of the R factor for each zone:

$$R(hk0)=20\cdot0\%,\quad R(h0l)=20\cdot9\%,\quad R(0kl)=15\cdot5\%\,.$$

In calculating these figures, a temperature factor of $\exp(-1\cdot3\sin^2\theta)$ was applied to the F_c, and the relative F_o were scaled to them by direct comparison. For a detailed comparison of F_o and F_c, reference should be made to a thesis for the degree of Ph.D. (McDonald, 1950).

Hydrogen atoms

It was observed at this stage that many of the worst discrepancies between F_o and F_c occurred in the low orders. The inference that the hydrogen atoms were making significant contributions to the structure factors was confirmed by computing (F_o-F_c) syntheses for the three projections. The c-axis projection is reproduced here (Fig. 5). The positions of the hydrogen atoms, marked as crosses, were calculated from the oxygen and carbon atom positions by assuming tetrahedral valency angles and a C–H bond length of 1·1 Å. In the case of the hydrogen bonds, the hydrogen atoms were simply placed half-way between the oxygen atoms. The correspondence between the theoretical positions and the electron-density maxima is quite striking. In the case of the other two projections, the correspondence is less marked owing to the poorer resolution, but the general agreement is quite unmistakable. This evidence, while demonstrating that the positions postulated for the hydrogen atoms are approximately correct, is not sufficiently definite to render their precise location possible, or to justify any attempts at refinement.

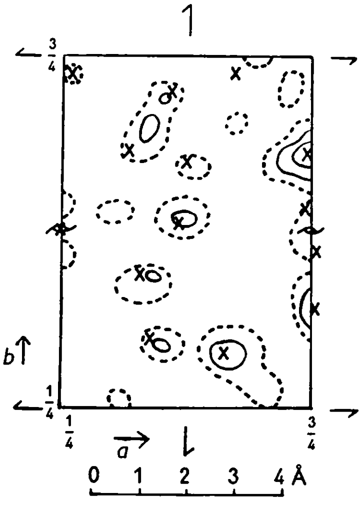

Fig. 5. (F_o-F_c) synthesis on ab. Contours at ½, 1, 1½ e.Å^{-2}. Theoretical hydrogen-atom positions are marked with crosses.

The effect of including the hydrogen atoms in the structure-factor calculations is shown in Table 1. The improvement in the agreement may be judged from the fact that the value of R is reduced from 0·17 to 0·12 for the 42 planes most affected. The coordinates of the hydrogen atoms used in these calculations are

Table 1. *Effect of hydrogen atoms on calculated structure factors*

hkl	F_o	F_c (without H)	F_c (with H)
200	45	48	51
400	43	−48	−42
600	2	3½	− 2
002	71	76	68
202	23	27	21
101	19½	8	11
301	7½	5½	3½
501	9½	− 6	−10
401	21½	21½	19½
102	30	−29	−27
060	15	−11	−14
080	20	19	21½
062	24½	31½	29½
051	41	−51	−47
021	5	½	3
041	28	−28	−25½
061	50	−48	−53
063	22	25	23
032	6½	− 4½	− 2
220	27	−27½	−31½
240	18½	15	18
260	26	23	26
280	0	− 1½	0
420	35½	−28½	−32
440	17½	−18½	−14½
210	15½	−10½	−12
250	11½	−14½	−10½
450	18½	21½	20
610	18	18	16½
630	13½	−12½	−14
110	15½	−10	− 7½
150	33	31½	33
310	20	28½	21½
330	14	−11½	−13½
350	24	18	23½
370	22½	−25	−23½
550	23½	−30	−27½
120	33½	32½	31
140	14½	−21	−23
1,10,0	0	7	4½
520	17½	−13	−14½
560	3	− 2	− 3½
$\Sigma\lvert F_o-\lvert F_c\rvert\rvert$		152	110

Table 2. *Hydrogen atom coordinates*

	x (Å)	y (Å)	z (Å)
H_1	2·71	0·62	2·97
H_2	0·31	4·38	3·68
H_3	0·92	11·49	0·68
H_4	7·68	3·39	0·61
H_5	7·77	0·00	4·09
H_6	2·71	1·70	−0·14
H_7	1·06	2·67	1·38
H_8	−0·90	2·31	−1·03
H_9	−0·74	0·85	1·68
H_{10}	−0·13	−0·13	−1·16
H_{11}	0·10	−1·67	1·48
H_{12}	−1·24	−1·85	0·36

Table 3. *Coordinates of carbon and oxygen atoms*

	x (Å)	*y* (Å)	*z* (Å)
O_1	1·60	1·47	3·27
O_2	1·16	3·83	4·61
O_3	9·04	3·19	0·85
O_4	8·09	0·48	0·37
O_5	1·36	0·23	0·18
O_6	0·53	12·36	4·70
C_1	1·66	1·50	4·58
C_2	0·85	2·62	0·30
C_3	9·73	2·23	0·03
C_4	9·43	0·84	0·60
C_5	0·04	14·68	4·90
C_6	10·19	13·24	0·42

All coordinates ±0·02 Å.

given in Table 2. The final coordinates of the carbon and oxygen atoms are given in Table 3.

The final *F* agreement for the three zones, as expressed by the reliability factor, is as follows:

$$R(hk0)=18{\cdot}4\,\%,\ R(h0l)=20{\cdot}4\,\%,\ R(0kl)=14{\cdot}7\,\%\ .$$

A number of *F*(*hkl*) were calculated as a check on the structure determination, and the agreement with the observed *F*'s was found to be very satisfactory.

Discussion of results

(*a*) *Dimensions of the molecule*

The various hypotheses which have been put forward concerning the configuration of the pyranose ring and the experimental evidence in support of them have been discussed elsewhere (Cox, Goodwin & Wagstaff, 1935; Reeves, 1950). We shall confine ourselves here to a discussion of the results which have been obtained in the present investigation.

It has been definitely established that, in α-glucose at least, the pyranose ring has the Sachse *trans* configuration, and that, in agreement with the chemical evidence, there is a *cis*-glycol grouping on C_1 and C_2 (Fig. 6). The interatomic distances are shown in Table 4. The C–C bond lengths agree well with one another and with the accepted value, the mean being 1·54 Å, and the maximum deviation from the mean 0·02 Å. The length of the C–O bond in aliphatic compounds may be taken as 1·43 Å (from electron-diffraction results for alcohols and ethers, see Allen & Sutton, 1950). Five of the C–O bond lengths in glucose agree well with this figure, but two, viz., C_1–O_1 and C_6–O_6 are short by 0·1 Å.

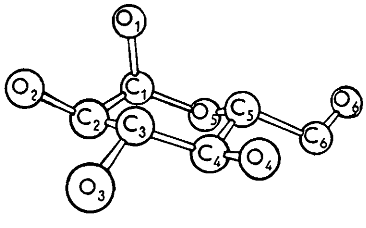

Fig. 6. Perspective drawing of α-glucose molecule.

Table 4. *Bond-lengths in the glucose molecule*

C_1–C_2	1·54 Å	C_1–O_1	1·32 Å
C_2–C_3	1·56	C_2–O_2	1·41
C_3–C_4	1·53	C_3–O_3	1·44
C_4–C_5	1·54	C_4–O_4	1·40
C_5–C_6	1·53	C_5–O_5	1·40
		C_1–O_5	1·42
		C_6–O_6	1·32

Table 5. *Bond angles in the glucose molecule*

O_5–C_1–C_2	110°	C_5–C_4–O_4	112°
C_1–C_2–C_3	104	C_3–C_4–O_4	110½
C_2–C_3–C_4	110	C_4–C_3–O_3	108
C_3–C_4–C_5	108	C_2–C_3–O_3	102
O_5–C_5–C_4	109½	O_5–C_1–O_1	112
C_5–O_5–C_1	111	O_1–C_1–C_2	115
O_5–C_5–C_6	110	C_1–C_2–O_2	109
C_4–C_5–C_6	112	C_3–C_2–O_2	109
C_5–C_6–O_6	113		

A rigorous estimate of the accuracy of the bond lengths has not been made, but we believe that the probable errors in the bond-lengths and coordinates are about 0·03 Å and 0·02 Å respectively. The observed shortening of C_1–O_1 and C_6–O_6 is therefore probably significant. A discussion of these anomalous bond lengths is hardly profitable at present, and should await supporting evidence from further work. However, it may be remarked that the glycosidic group $(OH)_1$ shows a peculiar chemical reactivity, and the primary hydroxyl group is to a certain extent anomalous in its chemical behaviour, as is shown by its preferred oxidation and substitution by certain reagents. Though there may be other explanations, it is perhaps not impossible that this chemical reactivity is due, in part at least, to some abnormality in the lengths of the bonds in question. The probable significance of these results is discussed in greater detail elsewhere (McDonald, 1950).

Unfortunately, there are few data relating to the exact bond lengths in the sugars with which to compare our results. It is interesting to note, however, that the recent work on cytidine (Furberg, 1950) shows some evidence for a shortening of the primary alcoholic bond, the length quoted being 1·35 Å, compared with 1·44 Å, the mean of the other C–O bonds in the sugar residue.

Table 5 shows the bond angles in the glucose molecule. In general they are quite close to the tetrahedral angle of 109½°, but in one or two cases they appear to be significantly different.

(*b*) *Intermolecular bonding*

The hardness and high density of the crystals of the sugars suggest that van der Waals forces play only a small part in bonding the molecules together, and that the packing arrangement is dictated mainly by the hydrogen bonding. This is found to be the case in glucose, the molecules being held together by a complete system of strong hydrogen bonds. The five available hydrogen atoms per molecule are fully utilized in the formation of intermolecular bonds; four of them are accounted for by oxygen atoms 2, 3, 4

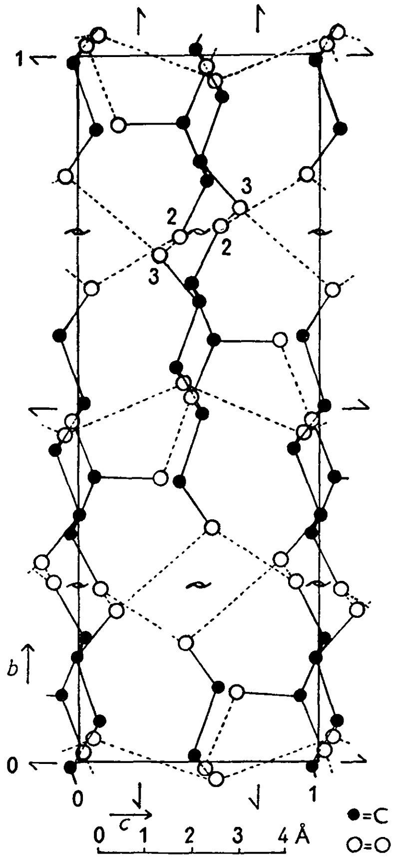

Fig. 7. Unit-cell contents projected down *a* axis. Hydrogen bonds are shown as broken lines.

and 6, each of which has two extra-molecular bonds, while the remaining hydrogen atom appears to take part in forming a bond between the ring oxygen atom of one molecule and the glycosidic atom of its neighbour. The lengths of the hydrogen bonds are as follows:

O_1–O_5	2·86 Å
O_2–O_3	2·78 Å
O_2–O_6	2·74 Å
O_3–O_6	2·70 Å
O_4–O_4	2·74 Å

The next closest approach of any two oxygen atoms in the structure is 3·65 Å. There is thus no ambiguity in the choice of hydrogen bonds.

The manner in which the molecules are packed together in the crystal structure is illustrated by Figs. 7 and 8 of this paper, and by Fig. 1 of the previous communication (McDonald & Beevers, 1950). The layer structure can be seen most clearly in the *a*-axis projection (Fig. 7). The molecules forming the layers are held together by hydrogen-bond formation between atoms O_2 and O_3, while the remaining hydrogen bonds serve to link the different layers together. The intermolecular bonding is equally strong in all directions, in agreement with the absence of any marked anisotropy in the physical properties of the crystal.

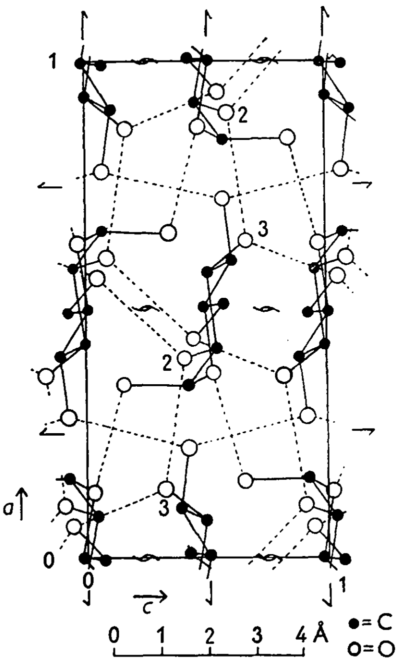

Fig. 8. Unit-cell contents projected down *b* axis. Hydrogen bonds are shown as broken lines.

One of us (T.R.R.McD.) is indebted to the Department of Scientific and Industrial Research for a maintenance allowance enabling him to take part in this work.

References

Allen, P. W. & Sutton, L. E. (1950). *Acta Cryst.* **3**, 46.
Beevers, C. A. & Cochran, W. (1947). *Proc. Roy. Soc.* A, **190**, 257.
Cox, E. G., Goodwin, T. H. & Wagstaff, A. I. (1935). *J. Chem. Soc.* p. 1495.
Cox, E. G. & Jeffrey, G. A. (1939). *Nature, Lond.* **143**, 894.
Cruickshank, D. W. J. (1950). *Acta Cryst.* **3**, 10.
Farquhar, M. C. M. & Lipson, H. (1946). *Proc. Phys. Soc.* **58**, 200.
Furberg, S. (1950). *Acta Cryst.* **3**, 325.
Groth, P. (1906–19). *Chemische Krystallographie.* Leipzig: Engelmann.
Harker, D. (1936). *J. Chem. Phys.* **4**, 381.
Hengstenberg, J. & Mark, H. (1930). *Z. Krystallogr.* **72**, 301.
Hughes, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.
McDonald, T. R. R. (1950). Ph.D. Thesis, University of Edinburgh.
McDonald, T. R. R. & Beevers, C. A. (1950). *Acta Cryst.* **3**, 394.
Patterson, A. L. (1949). *Acta Cryst.* **2**, 339.
Qurashi, M. M. (1949). *Acta Cryst.* **2**, 404.
Reeves, R. E. (1950). *J. Amer. Chem. Soc.* **72**, 1499.
Shoemaker, D. P., Donohue, J., Schomaker, V. & Corey, R. B. (1950). *J. Amer. Chem. Soc.* **72**, 2328.
Spônsler, O. L. & Dore, W. H. (1931). *J. Amer. Chem. Soc.* **53**, 1639.
Wilson, A. J. C. (1942). *Nature, Lond.* **150**, 152.