Acta Cryst. (1970). B26, 1373

The Crystal and Molecular Structure of Methyl β-Cellobioside–Methanol*

BY JOHN T. HAM AND DALE G. WILLIAMS

The Institute of Paper Chemistry, Appleton, Wisconsin, 54911, U.S.A.

(*Received* 23 *April* 1969)

Methyl β-cellobioside–methanol ($C_{13}O_{11}H_{24}.CH_3OH$) has space group $P2_1$ with two molecules per unit cell. The lattice parameters are $a=7{\cdot}652$, $b=25{\cdot}532$, $c=4{\cdot}496$ Å, and $\beta=101{\cdot}84°$ at the experimental temperature of about -193 °C. The structure was solved by means of the symbolic addition technique and Fourier syntheses. Block-diagonal least-squares refinement resulted in a final R index of 0·060 for 1724 reflections. The C(5)–C(6) and C(5′)–C(6′) bonds are shorter than the average carbon–carbon bond in the corresponding glucopyranose rings by 2·2 and 3·0 σ. The C(1)–O(1) and C(1′)–O(1′) equatorial glycosidic bonds are shorter than the average of the carbon–hydroxyl bonds by 7·2 and 9·3 σ. The valence angle at the O(1′) oxygen atom is 2·7° smaller than that of the O(1) oxygen atom. The structure contains a bifurcated intramolecular hydrogen bond between the two glucopyranose rings. The angle of twist between the least-squares planes formed by the two rings is 169·3°. The ring-to-ring conformation after Rees & Skerrett is (−25°, +142°). The average deviation between the carbon–carbon, ring oxygen–carbon, and glycosidic oxygen–carbon bond lengths which are common to methyl β-cellobioside and cellobiose is 0·005 Å.

Introduction

Methyl β-cellobioside was chosen for study because it has been used as a model compound for cellulose (Best & Green, 1968). Earlier related crystal structure studies included β-D-glucose (Ferrier, 1963) and cellobiose (Jacobson, Wunderlich & Lipscomb, 1961; Brown, 1966) which confirmed that the β-glucose residue has the pyranoid configuration and the all equatorial Cl chair conformation (Reeves, 1950), with the β 1–4 configuration between the two glucopyranose rings. The first cellobiose structure determination (Jacobson, Wunderlich & Lipscomb, 1961) showed a bent chain conformation with the two glucopyranose rings twisted relative to one another by roughly 150°, with an intramolecular hydrogen bond between O(3′) and O(5). A later structure refinement of cellobiose by Brown (1966) gave several bonds which deviated significantly from the expected values. Chu & Jeffrey (1968) recently completed a second refinement of the cellobiose structure using room temperature counter data and concluded that the C(1)–O(1) and C(1′)–O(1′) bonds were indeed short, but that the remaining bond lengths are normal, as was consistent with observations from other carbohydrate structure determinations (Berman, Chu & Jeffrey, 1967).

A first objective of determining the methyl β-cellobioside structure was a detailed comparison of the bond lengths and bond angles common to methyl β-cellobioside and cellobiose. A second objective was to compare the structural features of its two glycosidic bonds. Since cellobiose has been used as a model to speculate on the extent of rotation about the single bonds to the glycosidic bridge oxygen atom and on the nature of the intramolecular hydrogen bonding between the two glucopyranose rings in the cellulose monomer unit, a final reason for solving the methyl β-cellobioside structure was to determine the effect of the addition of the methyl glycosidic group on the chain conformation and on the intramolecular hydrogen bonding.

* A portion of a thesis submitted by John T. Ham in partial fulfillment of the requirements of The Institute of Paper Chemistry for the degree of Doctor of Philosophy from Lawrence University, Appleton, Wisconsin, January, 1969.

Crystal data

Suitable crystals were obtained after three recrystallizations from methanol of a sample of methyl β-cellobioside obtained from Best & Green (1968). A one-to-one solute–solvent complex was formed which gave the data listed in Table 1. The unit-cell dimensions were determined from the high-angle Cu $K\alpha_2$ and Cu $K\alpha_1$ spots on back-reflection photographs (Ham, 1969*a*, *b*). The crystal density was determined by flotation in a mixture of benzene and carbon tetrachloride.

Table 1. *Fundamental crystal properties**

Methyl β-cellobioside.methanol, $C_{13}O_{11}H_{24}.CH_3OH$, M.W. 388·37. Monoclinic, space group $P2_1$, from systematic extinctions ($0k0$) with $k=2n+1$.

Parameter	20 °C	−193 °C
a	7·636 (2) Å	7·652 (1) Å
b	25·978 (5) Å	25·532 (6) Å
c	4·567 (1) Å	4·496 (2) Å
β	101·55 (5)°	101·84 (5)°
V	887·5 (4) Å^3	859·6 (4) Å^3
Z	2	2
D_f	1·4535 (6) g.cm^{-3}	
D_x	1·4527 (7) g.cm^{-3}	1·4999 (7) g.cm^{-3}
μ Cu $K\alpha$	11·60 cm^{-1}	11·97 cm^{-1}

* The numbers in parentheses are the standard deviations referred to the last decimal place.

Experimental

Two approximately cylindrical crystals were used to record the intensity data. The crystal rotated around the *a* axis was approximately 0·35 mm long and 0·24 by 0·24 mm in cross section. The crystal rotated around the *c* axis was approximately 0·55 mm long and 0·09 by 0·18 mm in cross section. The intensity data were recorded with an equi-inclination Weissenberg camera using Cu $K\alpha$ radiation and the multiple-film exposure technique. The crystals were kept near −193 °C during each exposure with a liquid nitrogen gas flow cryostat similar to the one designed by Richards (1964). The low crystal-temperature was employed to reduce molecular thermal motion and to extend the angular range of the observable data, and thus improve the precision of the parameters obtained from the refined structure. The delivery tube from the cryostat was jacketed with liquid nitrogen so that when the cold gas delivery nozzle was positioned 1 mm from the crystal, the end of the liquid nitrogen jacket was about 1 cm from the crystal. In order to keep the X-ray film at 20 °C, while the crystal was kept just above the boiling point of liquid nitrogen, a temperature control system was designed for the Weissenberg film holder (Ham, 1969*a*, *b*). It was possible to estimate 95·9 % of the theoretically recordable reflections visually by using one film packet, containing four films per layer line, and exposing for fourteen hours. When the reflection spots were partially separated, only the $K\alpha_1$ intensity was estimated, and the resulting intensity was multiplied by 1·1, 1·2, 1·3 or 1·4, depending on the degree of separation. When the $K\alpha_1$ and $K\alpha_2$ spots were completely separated, the $K\alpha_1$ intensity was multiplied by 1·5. The intensity data were corrected for spot shape, and Lorentz and polarization factors. No absorption correction was made. The resulting data were correlated and reduced to relative structure factor amplitudes using the method of Rollett & Sparks (1960). A total of 1724 individual structure factor amplitudes were observed, of which 71 had intensities below that of the background. A *K* curve was used to put the data on an absolute scale, in order to calculate normalized structure factor magnitudes, $|E|$. The values of the statistical averages for $\langle|E|\rangle$ and $\langle|E^2-1|\rangle$ for the noncentrosymmetric class of reflections (96 % of the total) were 0·880 and 0·747 as compared with the theoretical values of 0·866 and 0·736 for noncentrosymmetric space groups.

Solution of the structure

The noncentrosymmetric symbolic addition method developed by Karle & Karle (1966) revealed a fragment of the structure. Equation (1) was used to define 169 phases ($|E| \geq 1{\cdot}5$) in terms of the initial set of phases listed in Table 2.

$$\varphi_{\mathbf{h}} \simeq \langle \varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}} \rangle_{\mathbf{k}_r} \quad . \qquad (1)$$

This initial set of phases included three origin-determining phases (Hauptman & Karle, 1956), two symbolic phases to implement the use of equation (1), and the independently determined phase of 600, a structure invariant in the centrosymmetric projection (Killean, 1966) which was determined with 98 % probability.

Table 2. *Initial set of phases*

Reflection	Phase	*E*
$\bar{1}$ 0 1	0°	3·04
8 0 1	0	2·19
$\bar{1}$ 1 2	0	2·13
4, 17, 1	a	2·62
3 3 1	b	2·41
6 0 0	180	2·09

There were several indications that the symbolic phases, *a* and *b*, were related, the most frequently occurring being $a = 4b \pm \pi$. Thus, the 169 phases were approximately defined in terms of one unknown symbol, *b*. Three numerical sets of phases were calculated with the tangent formula

$$\tan \varphi \simeq \frac{\sum_{\mathbf{k}} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \sin (\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})}{\sum_{\mathbf{k}} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \cos (\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})}, \qquad (2)$$

by setting *b* equal to 45, 90 and 135°. Only positive imaginary phases were considered for *b* so that its value could specify one of the possible enantiomorphs. Using the tangent formula to extend the list of defined phases to

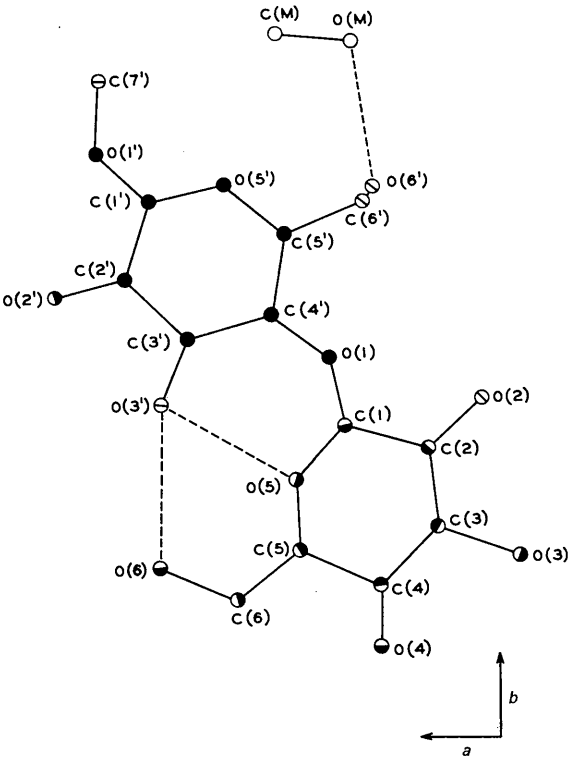


Fig. 1. Structure of methyl β-cellobioside–methanol.

those with a minimum $|E|$ of 1·40 yielded three numerical sets of 224 phases. All phase refinements with the tangent formula were accomplished with three iterations.

E maps were calculated for the three sets of phases. The resulting maps had a wide range of peak heights making it difficult to distinguish between atomic peaks and background peaks. A minimum peak magnitude was accepted which was low enough to give at least as many peaks as the total number of carbon and oxygen atoms in the asymmetric unit. All three maps showed puckered six-membered rings in the chair conformation and peaks were located equatorially at the C(1) and C(4) positions. The C(1) and C(4) positions were assumed because the C(1)–C(4) vector was almost parallel to the long axis of the unit cell, b, which was consistent with the molecular packing required by the unit-cell dimensions. The coordinates of the eight peak cluster in the E map based on $b=45°$ were all different from the corresponding peaks in the other two maps and their peak magnitudes were lower. Those in the two E maps with $b=90°$ and $b=135°$ differed only by three per cent in the y coordinate suggesting that either one might lead to the structure. The reality of each peak in the eight-atom cluster in these two maps was supported by the vectors in a sharpened three-dimensional Patterson map. This eight-atom cluster is represented in Fig. 1 by the solid circles.

The structure was completed with three-dimensional Fourier syntheses starting with the phases calculated from the eight-atom fragment observed in the E map based on $b=135°$. This fragment gave an $R=\sum(|F_o|-|F_c|)|/\sum|F_o|$ of 0·53. The first Fourier synthesis revealed eleven more connected peaks which completed a second six-membered puckered ring in the chair conformation. These eleven peaks are represented by the half-filled circles in Fig. 1. The addition of these peaks reduced R to 0·39. A second Fourier synthesis revealed the remaining five carbon and oxygen atoms in methyl β-cellobioside, which are represented by the dashed circles in Fig. 1, and gave an R of 0·28. A third Fourier synthesis revealed the solvent atom positions as two additional peaks whose magnitudes were in a ratio of approximately eight to six and which were roughly 1·41 Å apart, indicating a molecule of methanol. The addition of these two atoms, represented by the hollow circles in Fig. 1, reduced R to 0·173.

Before completing the structure with Fourier syntheses, the tangent formula recycling approach was tried with a seven-atom fragment. Six of these atoms were in the eight-atom E map fragment described above and the seventh one proved to be a spurious peak. Tangent formula refinement based on this fragment produced no new information in a second E map.

Structure refinement

The refinement was with a block-diagonal least-squares program written by Mair (1963) for the 40K IBM 1620 computer. Six cycles of refinement with isotropic tem-

Table 3. *Fractional atomic coordinates and thermal parameters for the carbon and oxygen atoms**

	x	y	z	B_{11}	B_{22}	B_{33}	B_{23}	B_{13}	B_{12}
C(1)	0·6350 (6)	0·1701 (2)	0·6311 (10)	0·00323	0·00041	0·01539	−0·00010	0·00315	0·00003
C(2)	0·4450 (5)	0·1499 (2)	0·6084 (10)	0·00325	0·00053	0·01260	0·00086	0·00532	0·00050
C(3)	0·4297 (6)	0·0954 (2)	0·4615 (10)	0·00295	0·00054	0·01241	0·00007	0·00430	−0·00081
C(4)	0·5756 (6)	0·0572 (2)	0·6103 (10)	0·00414	0·00051	0·01158	0·00014	0·00955	−0·00020
C(5)	0·7563 (5)	0·0849 (2)	0·6488 (10)	0·00303	0·00044	0·01401	−0·00038	0·00920	0·00042
C(6)	0·9062 (5)	0·0523 (2)	0·8312 (10)	0·00188	0·00041	0·01977	−0·00003	0·00329	0·00030
C(1′)	1·0495 (5)	0·3343 (2)	0·9114 (10)	0·00276	0·00045	0·01810	0·00148	0·00654	0·00031
C(2′)	1·1134 (5)	0·2802 (2)	0·8460 (9)	0·00229	0·00048	0·01359	0·00067	0·00190	−0·00005
C(3′)	0·9783 (6)	0·2369 (2)	0·8650 (11)	0·00481	0·00035	0·01689	−0·00009	0·00624	−0·00067
C(4′)	0·7890 (6)	0·2522 (2)	0·7029 (10)	0·00265	0·00043	0·01761	0·00073	0·00426	−0·00049
C(5′)	0·7475 (6)	0·3071 (2)	0·8044 (10)	0·00478	0·00030	0·01657	0·00019	0·01017	−0·00081
C(6′)	0·5635 (6)	0·3270 (2)	0·6668 (10)	0·00522	0·00051	0·01447	−0·00026	0·01284	0·00013
C(7′)	0·1465 (6)	0·4218 (2)	0·9358 (13)	0·00486	0·00059	0·03288	−0·00068	0·00821	−0·00016
O(1)	0·6574 (4)	0·2178 (1)	0·7829 (7)	0·00328	0·00044	0·01594	−0·00003	0·00875	−0·00021
O(2)	0·3193 (4)	0·1839 (1)	0·4298 (7)	0·00262	0·00056	0·02100	0·00233	0·00432	0·00044
O(3)	0·2586 (4)	0·0730 (1)	0·4645 (7)	0·00248	0·00065	0·01706	0·00026	0·00243	−0·00053
O(4)	0·5793 (4)	0·0134 (1)	0·4209 (7)	0·00580	0·00057	0·01929	−0·00152	0·01081	−0·00021
O(5)	0·7532 (4)	0·1331 (1)	0·8091 (7)	0·00390	0·00038	0·01322	0·00011	0·00415	0·00019
O(6)	1·0701 (4)	0·0814 (1)	0·8947 (7)	0·00286	0·00054	0·01742	−0·00108	0·00322	−0·00014
O(1′)	1·1657 (4)	0·3698 (1)	0·8254 (7)	0·00357	0·00039	0·02414	0·00023	0·00883	−0·00038
O(2′)	1·2776 (4)	0·2707 (1)	1·0606 (7)	0·00370	0·00049	0·02138	0·00176	−0·00157	−0·00025
O(3′)	1·0435 (4)	0·1920 (1)	0·7317 (8)	0·00350	0·00035	0·02658	0·00044	0·00727	0·00046
O(5′)	0·8755 (4)	0·3427 (1)	0·7260 (7)	0·00276	0·00044	0·01934	0·00099	0·00268	−0·00020
O(6′)	0·5444 (4)	0·3344 (1)	0·3442 (7)	0·00476	0·00055	0·01309	0·00011	0·00770	0·00003
C(M)	0·7369 (7)	0·4470 (2)	0·1961 (13)	0·00717	0·00064	0·03420	−0·00074	0·01731	−0·00082
O(M)	0·5713 (5)	0·4392 (1)	0·2865 (9)	0·00729	0·00053	0·04050	0·00020	0·01677	0·00028

* The estimated standard deviations of the fractional coordinates are given in the parentheses. They refer to the last decimal positions of the respective values. The expression for the temperature factor which is consistent with the B values is

$$\exp\left[-(h^2B_{11}+k^2B_{22}+l^2B_{33}+klB_{23}+hlB_{13}+hkB_{12})\right].$$

perature parameters and six cycles with anisotropic parameters reduced R to 0·095 and 0·085 respectively. At this point, Fourier difference syntheses clearly located the hydrogen atoms with reasonable bond lengths and angles, and reduced R to 0·069. The atomic scattering factors for carbon and oxygen atoms were those computed from the Hartree–Fock–Slater wave function by Hanson, Herman, Lea & Skillman (1964). Those for the hydrogen atoms were by Stewart, Davidson & Simpson (1965).

Four cycles of refinement with anisotropic thermal parameters, fixed hydrogen atom positions and unit weights reduced R to 0·063. The weighting scheme provided by Mair's (1963) least-squares program,

$$w=1/\{1+[(|F_o|-b)/a]^2\}\,, \qquad (4)$$

with the constants $b=9{\cdot}56$ and $a=10{\cdot}00$ derived empirically after methods by Cruickshank (1961), was introduced and five more cycles of anisotropic refinement reduced R to its final value of 0·060. The final positional coordinates and thermal parameters are listed in Tables 3 and 4 and the structure factor data in Table 5.

Since there were several spurious peaks in the E map which led to the structure solution, it is of interest to discover just how accurate the tangent formula phase estimates were. This was done by comparing the phases obtained with the tangent formula with the phases calculated from the refined structure. Fig. 2 shows the distribution of the discrepancies between these two sets of phases for the 224 reflections used to construct the E map. Only nine of the 224 E values were from the centrosymmetric class. The unusually high average of 43° (Karle & Karle, 1966) is consistent with the large number of spurious peaks. The poor quality of the E map is probably due to the use of reflection 801 to determine the origin. This reflection entered into only two $\sum_2$ interactions for which the variance was less than 0·5. It was chosen because it was the only centrosymmetric reflection with $|E|$ greater than 1·50 and which was not in the parity group represented by reflection $\bar{1}01$.

The failure of the tangent formula recycling approach is probably related to the large average error in the initial set of tangent formula phases.

Description of the structure

As expected, methyl β-cellobioside is composed of two glucopyranose rings and the configuration at both C(1) and C(1′) is β. Projections of the structure are shown in Fig. 3.

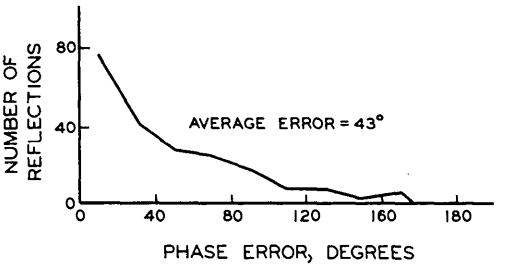


Fig. 2. Distribution of the error in the tangent formula phases.

Table 4. *Hydrogen atom positions (from difference maps)*

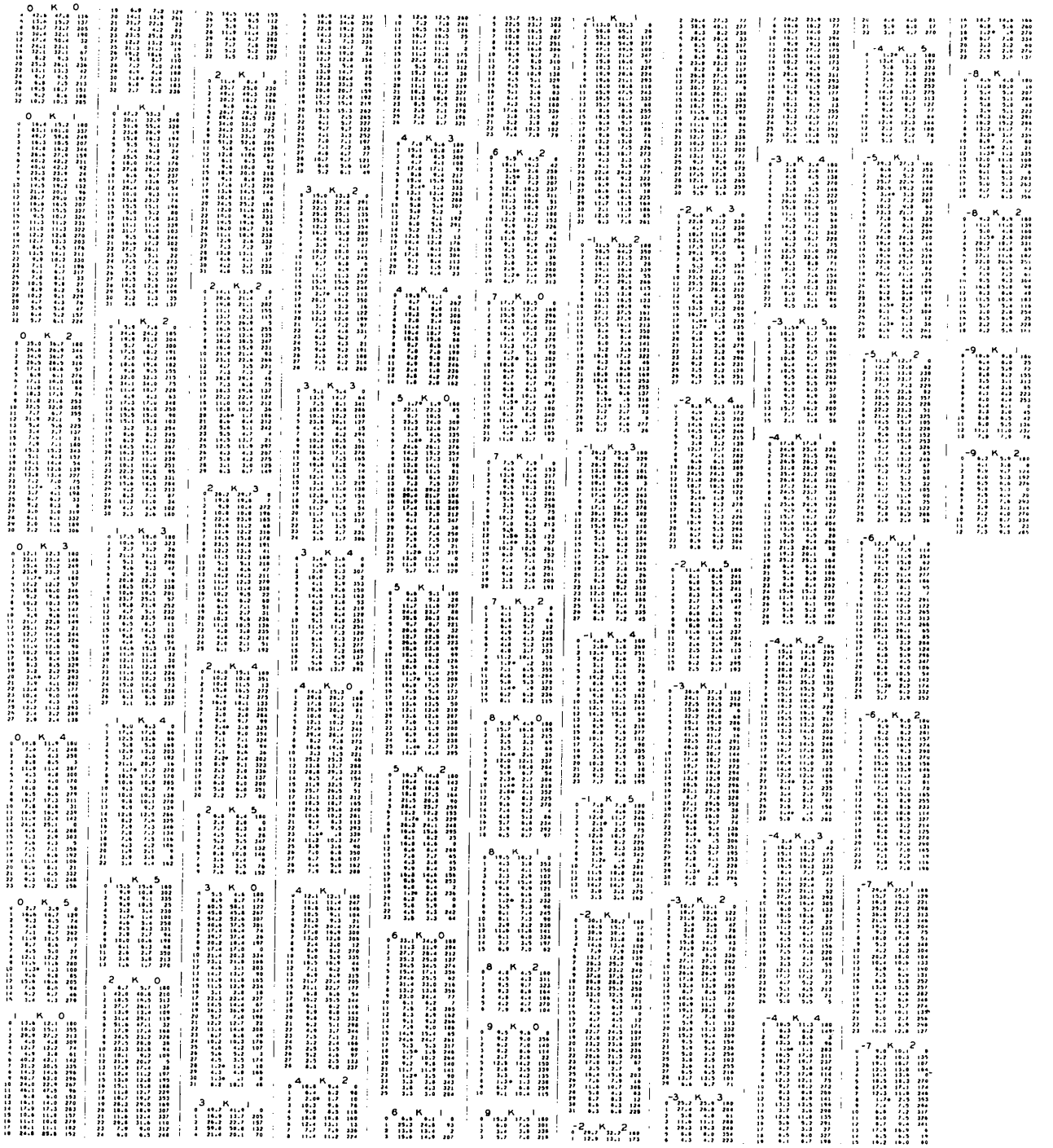
	Fractional coordinates			Peak height	Unrefined bond
	x	y	z	(e.Å^{-3})	length
H(1)	0·676	0·175	0·425	0·9	1·05Å
H(2)	0·400	0·149	0·826	0·9	1·10
H(3)	0·458	0·100	0·245	0·8	1·05
H(4)	0·548	0·047	0·805	0·9	0·98
H(5)	0·788	0·094	0·455	0·9	0·98
H(6–1)	0·866	0·043	1·050	0·8	1·11
H(6–2)	0·930	0·020	0·740	0·7	0·96
H(1′)	1·040	0·338	1·140	0·9	1·05
H(2′)	1·120	0·282	0·635	0·9	0·96
H(3′)	0·990	0·229	1·095	0·8	1·04
H(4′)	0·760	0·252	0·470	0·9	1·03
H(5′)	0·750	0·307	1·010	0·8	0·92
H(6′–1)	0·540	0·362	0·740	0·8	0·98
H(6′–2)	0·474	0·293	0·685	0·7	1·12
H(7′–1)	1·158	0·424	1·146	0·8	0·93
H(7′–2)	1·224	0·452	0·850	0·7	1·09
H(7′–3)	1·022	0·439	0·880	0·7	1·03
H(M–1)	0·752	0·438	−0·045	0·7	0·96
H(M–2)	0·738	0·483	0·080	0·7	1·06
H(M–3)	0·790	0·409	0·190	0·6	1·05
H(O2)	0·233	0·190	0·494	0·6	0·79
H(O3)	0·190	0·075	0·334	0·7	0·70
H(O4)	0·587	−0·012	0·470	0·6	0·74
H(O6)	1·148	0·073	0·708	0·6	1·15
H(O2′)	1·288	0·241	1·160	0·8	0·88
H(O3′)	1·004	0·166	0·826	0·6	0·87
H(O6′)	0·506	0·308	0·240	0·6	0·84
H(OM)	0·550	0·408	0·270	0·7	0·81

The bond lengths and bond angles involving carbon and oxygen atoms are listed in Table 6. The unrefined bond lengths involving hydrogen atoms are included in Table 4. Standard deviations are 0·006 Å for the carbon–carbon bonds, 0·005 Å for the carbon–oxygen bonds (0·006 Å for the two involving methyl carbon atoms), and 0·3° for the valence bond angles.

The average carbon–carbon bond length associated with the unprimed ring is 1·527 Å, with a spread of 0·020 Å, and that associated with the primed ring is

Table 5. *Observed and calculated structure factor magnitudes and their phase angles*

The column headings are: k, $|F_o|$, $|F_c|$, and phase angle. * refers to the unobserved reflections.



1·521 Å, with a spread of 0·028 Å. Although this slight difference is only 1σ, a similar discrepancy was observed for cellobiose by Chu & Jeffrey (1968) (*i.e.* 1·526 to 1·519 Å).

It has been pointed out that each carbon atom in a sugar molecule has a unique steric environment of adjacent bonds which could account for the differences between bond lengths which are larger than their standard deviations (Berman, Chu & Jeffrey, 1967). In this structure the spread of the carbon–carbon bond lengths in both glucopyranose rings is significantly reduced by removing the exterior bond, C(5)–C(6), from the average, and the difference between this bond and the average of the ring carbon–carbon bonds is of high significance. These results are shown in Table 7. This slight shortening of the C(5)–C(6) bond with respect to the average ring carbon–carbon bond has not been remarked upon previously. Although it is observed to a greater or less degree in the compounds listed, which have a close relationship to methyl β-cellobioside, some of the observed shortening may be due to the effect of thermal motion on these room temperature data.

The average length of the carbon–hydroxyl bonds is 1·428 Å with a spread of 0·030 Å. In both the primed and unprimed rings, the two ring oxygen–carbon bonds are nearly equal: C(1)–O(5) = 1·432 *versus* C(5)–O(5) = 1·430 and C(1′)–O(5′) = 1·434 *versus* C(5′)–O(5′) = 1·432 Å. The glycosidic bond lengths, C(1)–O(1) and C(1′)–O(1′), are shortened by 0·038 and 0·049 Å with respect to the mean of 1·428 Å. These differences are highly significant ($7{\cdot}2\sigma$ and $9{\cdot}3\sigma$ respectively) in agreement with the results summarized by Berman, Chu & Jeffrey (1967) for other pyranose sugars with an equatorial carbon–oxygen bond at the C(1) carbon atom.

There are no significant differences between the carbon–carbon, ring oxygen–carbon, and glycosidic oxygen–carbon bond lengths which are common to methyl β-cellobioside and cellobiose (Chu & Jeffrey, 1968). The average difference is 0·005 Å ($0{\cdot}7\sigma$) and the largest difference is 0·015 Å ($1{\cdot}8\sigma$). The slightly larger average

Table 6. *Bond lengths and bond angles**

i	*j*	*L*(*ij*)	δ†
C(1)	C(2)	1·526 (6) Å	−01
C(2)	C(3)	1·535 (6)	−15
C(3)	C(4)	1·528 (6)	+06
C(4)	C(5)	1·530 (6)	+02
C(5)	C(6)	1·515 (6)	+04
C(1)	O(1)	1·390 (5)	+07
C(1)	O(5)	1·432 (5)	−07
C(2)	O(2)	1·416 (5)	00
C(3)	O(3)	1·431 (5)	−04
C(4)	O(4)	1·410 (5)	+10
C(5)	O(5)	1·430 (5)	+06
C(6)	O(6)	1·434 (5)	−18
C(1′)	C(2′)	1·513 (6)	+01
C(2′)	C(3′)	1·529 (6)	−10
C(3′)	C(4′)	1·533 (6)	−03
C(4′)	C(5′)	1·526 (6)	+01
C(5′)	C(6′)	1·505 (6)	−04
C(1′)	O(1′)	1·379 (5)	+02
C(1′)	O(5′)	1·434 (5)	+01
C(2′)	O(2′)	1·439 (5)	−16
C(3′)	O(3′)	1·429 (5)	−19
C(4′)	O(1)	1·437 (5)	+09
C(5′)	O(5′)	1·432 (5)	+05
C(6′)	O(6′)	1·440 (5)	−17
C(7′)	O(1′)	1·435 (6)	—
C(M)	O(M)	1·422 (6)	—

i	*j*	*k*	Angle (*ijk*)	δ†
O(5)	C(1)	C(2)	107·4 (3)°	+0·9
C(1)	C(2)	C(3)	108·8 (3)	−0·5
C(2)	C(3)	C(4)	113·9 (4)	−4·4
C(3)	C(4)	C(5)	108·5 (3)	+2·5
C(4)	C(5)	O(5)	110·3 (3)	+0·2
C(2)	C(1)	O(1)	110·4 (3)	−1·4
O(5)	C(1)	O(1)	107·6 (3)	−0·2
C(1)	C(2)	O(2)	111·0 (3)	−1·0
O(2)	C(2)	C(3)	109·1 (3)	+4·5
C(2)	C(3)	O(3)	110·3 (3)	+1·7
O(3)	C(3)	C(4)	109·2 (3)	−2·3
C(3)	C(4)	O(4)	110·2 (3)	−2·1
O(4)	C(4)	C(5)	107·8 (3)	+1·6
C(4)	C(5)	C(6)	112·0 (3)	−1·0
C(6)	C(5)	O(5)	106·8 (3)	−1·4
C(5)	C(6)	O(6)	111·0 (3)	+1·2
O(5′)	C(1′)	C(2′)	108·7 (3)	+0·6
C(1′)	C(2′)	C(3′)	113·6 (4)	−3·6
C(2′)	C(3′)	C(4′)	111·8 (4)	0·0
C(3′)	C(4′)	C(5′)	109·1 (4)	+3·2
C(4′)	C(5′)	O(5′)	108·3 (3)	+0·9
C(2′)	C(1′)	O(1′)	107·0 (3)	+3·2
O(5′)	C(1′)	O(1′)	108·3 (3)	−1·3
C(1′)	C(2′)	O(2′)	107·2 (3)	+3·4
O(2′)	C(2′)	C(3′)	110·3 (3)	−3·8
C(2′)	C(3′)	O(3′)	105·0 (3)	+2·1
O(3′)	C(3′)	C(4′)	112·5 (3)	0·0
C(3′)	C(4′)	O(1)	111·5 (3)	−2·5
O(1)	C(4′)	C(5′)	106·7 (3)	−0·3
C(4′)	C(5′)	C(6′)	114·9 (4)	−1·9
C(6′)	C(5′)	O(5′)	108·3 (3)	−1·9
C(5′)	C(6′)	O(6′)	110·6 (3)	+1·2
C(1)	O(5)	C(5)	111·1 (3)	+1·3
C(1′)	O(5′)	C(5′)	111·3 (3)	+2·2
C(1)	O(1)	C(4′)	115·8 (3)	+0·3
C(1′)	O(1′)	C(7′)	113·1 (3)	—

* The estimated standard deviations are in parentheses and refer to the last decimal positions of the respective values.
† δ is the deviation in thousandths of an angstrom or in degrees from the comparable data of cellobiose (Chu & Jeffrey, 1968).

deviation of 0·012 Å (1·7σ) between the carbon–hydroxyl bonds which are common to methyl β-cellobioside and cellobiose (Chu & Jeffrey, 1968) may be attributed to differences in the hydrogen bonding schemes in the two crystals.

The difference of 0·011 Å between the C(1)–O(1) and C(1′)–O(1′) bond lengths is significant only at the 85% confidence level (1·5σ), but is consistent with results obtained with β-D-glucose (Chu & Jeffrey, 1968), cellobiose (Chu & Jeffrey, 1968), and methyl β-maltoside

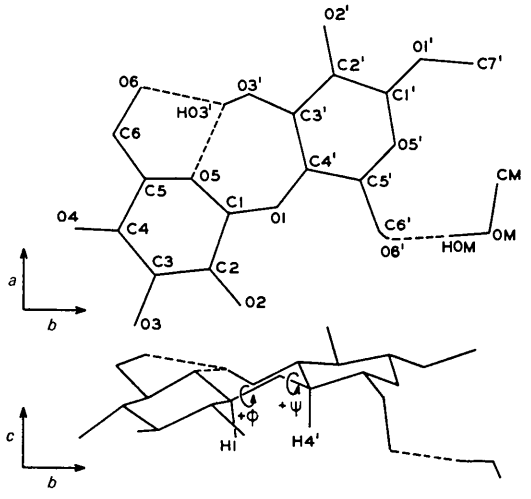


Fig. 3. Nomenclature and projected views of the refined structure of methyl β-cellobioside.methanol.

(Chu & Jeffrey, 1967), as shown in Table 8. The average glycosidic bond length with a glucosyl group is 0·015 Å (3·5σ) longer than the average of the glycosidic bonds with either a methyl group or a hydrogen atom.

In methyl β-cellobioside, the valence bond angle at the glycosidic oxygen atom with the methyl glycosidic group (113·1°) agrees with the corresponding angles found in the crystal structures of methyl-β-maltose (113·2°) (Chu & Jeffrey, 1967) and methyl β-xyloside (113·4°) (Brown, 1966). The valence bond angle at the glycosidic oxygen atom between the two glucopyranose rings is larger (115·8°) and agrees with the corresponding angle found in cellobiose (116·1°) (Chu & Jeffrey, 1968). Since the bulky nature of the glucosyl group compared with the methyl group increases the angle at the glycosidic oxygen atom by 2·7° and increases the length of the glycosidic bond by 0·011 Å, the methyl glycosidic bond is not a perfect model for the glycosidic bond in the cellulose monomer, even though the methyl group blocks the reducing end group.

The carbon valence bond angles interior to the rings average 110·0° with a spread of 6·5°. The average deviation from the corresponding angles in cellobiose (Chu & Jeffrey, 1968), which average 110·0° with a spread of 4·0°, is 1·7°. The carbon valence bond angles exterior to the rings average 109·8° with a spread of 9·9°. The average deviation from the corresponding angles in the cellobiose structure (Chu & Jeffrey, 1968), which average 109·5° with a spread of 8·2°, is 1·9°.

The molecular packing is determined by the hydrogen-bonding network, indicated by the dashed lines in

Table 7. *Shortening of the* C(5)–C(6) *bond associated with glucopyranose rings*

Compound	Average ring C–C	Range ring C–C	C(5)–C(6)	Significance*
Methyl β-cellobioside				
unprimed ring	1·530 Å	1·526–1·535 Å	1·515 (6) Å	2·3
primed ring	1·525	1·513–1·533	1·505 (6)	3·0
β-D-Glucose				
(Chu & Jeffrey, 1968)	1·521	1·511–1·529	1·513 (5)	1·5
Cellobiose				
(Chu & Jeffrey, 1968)				
unprimed ring	1·528	1·520–1·534	1·519 (5)	1·9
primed ring	1·523	1·514–1·530	1·501 (5)	3·8
Methyl α-D-glucoside				
(Berman & Kim, 1968)	1·523	1·509–1·531	1·506 (4)	3·8
α-D-Glucose				
(Brown & Levy, 1965)	1·526	1·519–1·534	1·510 (3)	4·8

* Significance expressed as the average carbon–carbon bond length in the ring minus the C(5)–C(6) bond length divided by the σ of this difference.

Table 8. *Equatorial* C(1)–O(1) *bond lengths on the glucopyranose ring*

Glycosidic group	Compound	C(1)–O(1)	Average length for the glycosidic group
H	β-D-Glucose	1·383 (4) Å	1·382 (3) Å
H	Cellobiose	1·381 (5)	
Methyl	Methyl β-maltoside	1·375 (8)	1·377 (5)
Methyl	Methyl β-cellobioside	1·379 (5)	
Glucosyl	Methyl β-cellobioside	1·390 (5)	1·394 (3)
Glucosyl	Cellobioside	1·397 (4)	

Fig. 4, and by the tendency of the nonpolar groups of the system to seek the adjacent positions, indicated by the dotted line triangles in Fig. 4.

All eight of the hydroxyl groups in the asymmetric unit take part in the hydrogen bonding (Table 9), with all the hydroxyl oxygen atoms except O(4) functioning as both donors and acceptors. The ethereal oxygen atoms O(1),O(1′) and O(5′) do not take part in the hydrogen bonding. Two major hydrogen bonding sequences are an infinite donor–acceptor chain linking O(6) and O(3) in the **c** direction and a seven-membered finite chain which winds along the screw axis in the negative **b** direction with the sequence: O(4) → O(M) → O(6′) → O(2′) → O(2) → O(3′) → O(5). The average length of the bonds in these two sequences is 2·633 and 2·727 Å respectively. The last link in the finite chain [O(3′) → O(5)] is an intramolecular hydrogen bond. The O(4) → O(M) → O(6′) link forms a continuous bonded chain of alternate methyl β-cellobioside and methanol molecules along the screw axis.

These two hydrogen bonding sequences are linked by a weaker intramolecular hydrogen bond, O(3′) →

Table 9. *Hydrogen bonding*

i	*j*	*k*	*l*	*L*(*jl*)	*L*(*jk*)	Angle (*ijk*)	Angle (*ijl*)	Angle (*jkl*)
C(4)	O(4)	H(04)	O(M) (*a*)	2·695 Å	0·743 Å	117·6°	101·5°	157·5°
C(M)	O(M)	H(0M)	O(6′) (–)	2·702	0·814	106·4	104·8	163·5
C(6′)	O(6′)	H(06′)	O(2′) (*b*)	2·714	0·838	114·2	107·8	141·4
C(2′)	O(2′)	H(02′)	O(2) (*c*)	2·748	0·875	117·7	121·2	170·8
C(2)	O(2)	H(02)	O(3′) (*d*)	2·741	0·788	115·9	105·9	165·5
C(3′)	O(3′)	H(03′)	O(5) (–)	2·762	0·874	102·7	91·4	134·1
C(3′)	O(3′)	H(03′)	O(6) (–)	2·914	0·874	102·7	133·2	135·5
C(6)	O(6)	H(06)	O(3) (*c*)	2·648	1·145	108·8	113·4	169·1
C(3)	O(3)	H(03)	O(6) (*d*)	2·678	0·704	119·3	105·8	159·7

Intermolecular nonbonded distances less than 3·3 Å

i	*j*	*L*(*ij*)
C(4)	O(M) (*a*)	3·281 Å
O(6′)	C(6′) (*e*)	3·084
O(6′)	C(5′) (*e*)	3·215

Symmetry code

–	*x*	*y*	*z*
a	1−*x*	*y*−½	1−*z*
b	*x*−1	*y*	*z*−1
c	*x*+1	*y*	*z*
d	*x*−1	*y*	*z*
e	*x*	*y*	*z*−1

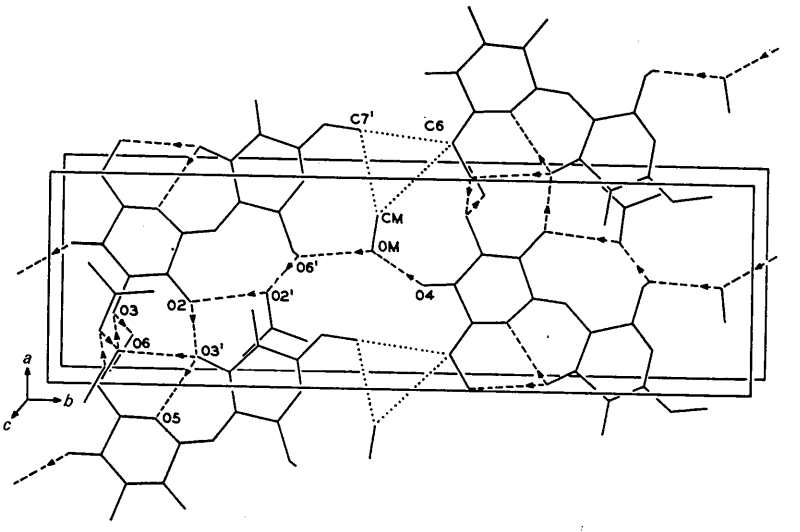


Fig. 4. Hydrogen bonding and molecular packing.

O(6) (2·914 Å). This hydrogen bond and the intramolecular hydrogen bond, O(3′) → O(5) (2·762 Å), are bifurcated. Although the O(3′) → O(6) bond is weaker than the O(3′) → O(5) bond, the angles formed by the hydrogen atom with the two oxygen atoms are nearly identical (135·5 and 134·1°). Both are smaller than the corresponding interior angles of the seven other hydrogen bonds in this structure, whose average is 161·5°. In cellobiose (Chu & Jeffrey, 1968), the corresponding intramolecular hydrogen bond O(3′) → O(5) (2·767 Å), is not bifurcated, since the O(3′) → O(6) distance is 3·120 Å.

Proton magnetic spectra of methyl β-cellobioside indicate that there is intramolecular hydrogen bonding in this molecule when it is dissolved in dimethyl sulfoxide (Ham, 1969*a*, *b*) and that this intramolecular hydrogen bonding is different from that of cellobiose in dimethyl sulfoxide solution (Casu, Reggiani, Gallo & Vigevani, 1966).

The two methyl groups and the C(6) methylene group form a nonpolar triangle [C(7′)–C(M)–C(6)] (Fig. 4) with comparatively short van der Waals separations, *i.e.* C(M)–C(7′) = 3·621, C(7′)–C(6) = 3·541, and C(6)–C(M) = 3·852 Å.

The cellulose monomer distance, O(4)–O(1′), in methyl β-cellobioside is 10·144 Å, which is 2·4% shorter than that calculated for cellobiose, 10·391 Å (Chu & Jeffrey, 1968). This difference is a consequence of the different intramolecular hydrogen bonding systems resulting in different conformation angles between the pyranose rings.

The methyl β-cellobioside molecule has the overall Cl chair conformation. All the substituents on the pyranose rings are in the equatorial positions. The individual conformation angles (Reeves, 1950) are listed in Table 10. The conformation angles in the glucopyranose rings range from 45·0 to 70·1°. The same angles in cellobiose (Chu & Jeffrey, 1968) range from 48·0 to 65·7° and differ from those in methyl β-cellobioside by an average of 5·0°.

Table 10. *Conformation angles*

Bond	Angle	δ*
I. Within the pyranose rings		
C(1) → C(2)	+58·7°	+4·8
C(2) → C(3)	−51·0	−6·3
C(3) → C(4)	+47·7	+4·3
C(4) → C(5)	−53·9	+2·0
C(5) → O(5)	+67·4	−7·6
O(5) → C(1)	−69·1	+3·4
C(1′) → C(2′)	+51·0	+6·8
C(2′) → C(3′)	−45·0	−5·7
C(3′) → C(4′)	+48·2	−0·2
C(4′) → C(5′)	−59·9	+8·8
C(5′) → O(5′)	+70·1	−9·2
O(5′) → C(1′)	−64·3	−0·8
II. Outside the pyranose rings		
O(5)–[C(5) → C(6)]–O(6)	+52·4	−3·7
O(5′)–[C(5′) → C(6′)]–O(6′)	−55·1	+120·6
C(4′)–[O(1) → C(1)]–O(5)	−88·9	+12·6
C(4′)–[O(1) → C(1)]–C(2)	+152·0	+14·6
C(1)–[O(1) → C(4′)]–C(3′)	+80·3	−6·7
C(1)–[O(1) → C(4′)]–C(5′)	−160·7	+28·4
C(7′)–[O(1′) → C(1′)]–O(5′)	−76·2	—
C(7′)–[O(1′) → C(1′)]–C(2′)	+166·8	—

* δ is the deviation in degrees from the comparable data o cellobiose (Chu & Jeffrey, 1968).

The rotational freedom of the primary alcohol groups is shown by the difference in conformation angles of the C(5)–C(6) bonds. The C(5)–C(6) bond has the +*syn*-clinal conformation (Reeves, 1950), whereas the C(5′)–C(6′) bond has the −*syn*-clinal conformation which allows O(6′) to hydrogen bond with the methanol molecule.

The conformational twists about the O(1)–C(1) and O(1)–C(4′) bonds are of interest for comparison of models for cellulose, which vary from 90° in the bent-chain model to 180° in a straight-chain model. The angle between the least-squares planes formed by the two-membered pyranose rings is 169·3°, and the mid planes, O(4)–C(4)–C(1) and C(4′)–C(1′)–O(1′), through the glucose units make angles of 150·1 and 143·5° with the plane of the C(1)–O(1)–C(4′) link. The corresponding angles in cellobiose (Chu & Jeffrey, 1968) are different, 148·6, 135·9 and 161·6° respectively. Rees & Skerrett (1968) describe ring-to-ring conformation in carbohydrates with the dihedral angles, φ and ψ, between C(1)–H(1) and O(1)–C(4′) in the H(1)–C(1)–O(1)–C(4′) system, and between O(1)–C(1) and C(4′)–H(4′) in the O(1)–C(1)–C(4′)–H(4′) system respectively, as shown in Fig. 3. In methyl β-cellobioside and cellobiose, these angles are (−25, +142°) and (−42, +162°) (Rees & Skerrett, 1968). In both compounds the intramolecular hydrogen bonding would not be possible if the molecules had the (+0, +180°) straight-chain, or the (+0, +90°) bent-chain conformation.

The changes in the conformation in the methyl β-cellobioside structure with respect to the cellobiose structure can be attributed to the differences between their two hydrogen bonding schemes, especially the intramolecular hydrogen bonding, and to the steric conditions introduced by the extension of the cellulose monomer unit by a methyl group.

Table 11 gives the thermal ellipsoid parameters. From the ratio of the average ellipsoid volume calculated from the cellobiose data (Chu & Jeffrey, 1968) to the corresponding volume calculated for the methyl β-cellobioside molecule (Table 11), it is inferred that cooling the crystal to liquid nitrogen temperature reduced the overall atomic thermal motion by a factor of 3·8. The largest reduction in thermal motion was that of the primary carbon and oxygen atoms. Interestingly, the thermal motion of the hydroxyl oxygen atoms is inversely proportional to the number of hydrogen bonds in which they take part. The fact that O(3′) acts both as a donor and an acceptor in the methyl β-cellobioside-methanol structure eliminated the unusually anisotropic thermal motion found for O(3′) in the cellobioside structure in which it acts only as a donor (Chu & Jeffrey, 1968).

The authors acknowledge advice given by Drs Jerome and Isabella Karle concerning the application of their noncentrosymmetric symbolic addition method. They thank Dr George A. Jeffrey of the University of Pittsburgh for several data reduction and data analysis computer programs and for a preprint describing his refinement of the cellobiose crystal structure. They also thank Dr Farid R. Ahmed of the Canadian National Research Council for the least-squares program used in the refinement process. Thanks are due to Dr Gerald F. Richards and Dr Richard Nelson of The Institute of Paper Chemistry for their many helpful discussions. One of the authors (J.T.H.) is especially indebted to his fellow student, William E. Scott, for assistance in writing the computer programs required to apply noncentrosymmetric symbolic addition with an IBM 40K 1620 computer.

References

BERMAN, H. M., CHU, S. C. & JEFFREY, G. A. (1967). *Science*, **157** (3796), 1576.

BERMAN, H. M. & KIM, S. H. (1968). *Acta Cryst.* B**24**, 897.

BEST, E. V. & GREEN, J. W. (1968). *Tappi*, **52**, 1321.

BROWN, C. J. (1966). *J. Chem. Soc.* (*A*). p. 922, 927.

BROWN, G. M. & LEVY, H. A. (1965). *Science* 147, 1038.

CASU, B., REGGIANI, M., GALLO, G. G. & VIGEVANI, A. (1966). *Tetrahedron*, **22**, 3061.

CHU, S. C. & JEFFREY, G. A. (1967). *Acta Cryst.* **23**, 1038.

CHU, S. C. & JEFFREY, G. A. (1968). *Acta Cryst.* B**24**, 830.

CRUICKSHANK, D. W. J. (1961). In *Computing Methods and the Phase Problem in X-ray Crystal Analysis*, Ed. R. PEPINSKY, J. M. ROBERTSON & J. C. SPEAKMAN. p. 45. New York: Pergamon Press.

Table 11. *Thermal ellipsoid parameters*

The root-mean-square atomic displacement, U_i, corresponds to the ith principal axis of the thermal ellipsoid, and θ_{ia}, θ_{ib}, θ_{ic} are the angles between the ith axis and the crystallographic axes, a, b and c respectively. The U_i were used to calculate the ellipsoid volume.

	i	B_i	U_i	θ_{ia}	θ_{ib}	θ_{ic}	Ellipsoid volume (Å^3 × 10^3)
C(1)	1	0·716 Å^2	0·0952 Å	175·1°	87·6°	73·8°	
	2	1·065	0·1161	95·8	167·9	78·5	
	3	1·199	0·1232	108·1	80·1	11·8	5·71
C(2)	1	0·582	0·0858	22·6	97·9	80·6	
	2	0·935	0·1088	121·9	68·4	31·0	
	3	1·493	0·1375	111·9	149·1	65·1	5·38
C(3)	1	0·482	0·0781	23·5	69·6	89·7	
	2	0·975	0·1111	120·2	94·2	18·9	
	3	1·520	0·1387	110·0	20·1	84·4	5·05
C(4)	1	0·348	0·0664	52·4	83·8	49·9	
	2	1·266	0·1266	136·3	109·2	42·1	
	3	1·350	0·1307	113·9	28·2	71·2	4·61
C(5)	1	0·256	0·0569	35·2	105·0	70·3	
	2	1·192	0·1228	83·6	10·7	82·9	
	3	1·281	0·1274	132·9	92·5	31·2	3·73
C(6)	1	0·385	0·0698	169·4	80·5	73·6	
	2	1·088	0·1174	100·7	168·8	84·8	
	3	1·540	0·1396	108·0	87·9	6·5	4·80
O(1)	1	0·431	0·0738	27·1	83·2	75·5	
	2	1·148	0·1206	92·8	168·7	78·7	
	3	1·377	0·1321	129·8	82·2	29·4	4·93
O(2)	1	0·549	0·0834	170·6	82·8	72·2	
	2	1·026	0·1140	95·7	40·8	49·2	
	3	2·073	0·1620	107·7	128·2	39·5	6·45
O(3)	1	0·516	0·0809	168·1	99·5	71·2	
	2	1·319	0·1292	107·2	81·5	10·1	
	3	1·741	0·1485	82·1	169·0	84·1	6·50
O(4)	1	0·797	0·1005	52·1	109·8	55·0	
	2	1·344	0·1304	128·2	136·9	65·8	
	3	1·954	0·1573	126·6	61·4	41·0	8·64
O(5)	1	0·801	0·1007	25·1	109·6	86·1	
	2	0·980	0·1114	98·6	38·6	51·3	
	3	1·073	0·1165	126·5	117·7	40·2	5·48
O(6)	1	0·631	0·0894	173·9	92·0	72·4	
	2	1·139	0·1201	101·7	130·9	41·1	
	3	1·624	0·1434	103·3	43·5	46·9	6·45
C(1′)	1	0·499	0·0795	9·9	89·7	91·8	
	2	0·974	0·1110	109·2	40·4	552·0	
	3	1·676	0·1457	114·8	123·1	37·3	5·39
C(2′)	1	0·512	0·0805	171·1	91·9	69·5	
	2	0·978	0·1113	103·5	60·9	29·3	

Table 11 (*cont.*)

	3	1·342	0·1303	95·3	150·3	60·3	4·89
C(3′)	1	0·688	0·0933	51·3	39·7	89·7	
	2	1·137	0·1200	67·7	132·7	57·4	
	3	1·403	0·1333	133·8	75·9	36·2	6·25
C(4′)	1	0·477	0·0777	21·0	69·2	97·2	
	2	1·109	0·1185	119·8	43·1	55·8	
	3	1·459	0·1359	103·3	119·3	29·5	5·24
C(5′)	1	0·389	0·0702	52·3	46·6	74·9	
	2	0·991	0·1120	76·4	144·7	62·0	
	3	1·573	0·1411	137·9	80·8	37·9	4·65
C(6′)	1	0·368	0·0683	52·9	96·2	49·4	
	2	1·337	0·1301	86·8	4·2	87·9	
	3	1·653	0·1447	141·6	89·4	39·8	5·39
O(1′)	1	0·573	0·0852	24·1	68·0	91·2	
	2	1·078	0·1168	70·1	160·0	92·3	
	3	1·903	0·1552	120·0	91·4	18·3	6·47
O(2′)	1	0·755	0·0977	156·5	87·7	54·8	
	2	1·051	0·1154	84·8	24·3	67·9	
	3	2·110	0·1634	86·2	120·1	33·8	7·72
O(3′)	1	0·618	0·0885	28·3	118·3	100·1	
	2	0·986	0·1117	64·1	26·6	88·9	
	3	2·063	0·1616	115·2	94·8	14·3	6·69
O(5′)	1	0·559	0·0871	168·4	100·0	72·8	
	2	1·045	0·1150	107·3	32·5	60·0	
	3	1·636	0·1439	101·2	116·2	26·2	6·04
O(6′)	1	0·669	0·0920	54·8	88·9	46·9	
	2	1·282	0·1274	142·1	83·0	41·3	
	3	1·437	0·1349	97·9	169·4	81·6	6·63
C(7′)	1	1·016	0·1134	3·2	86·7	101·4	
	2	1·517	0·1386	90·4	171·4	81·5	
	3	2·569	0·1804	113·2	82·2	13·9	11·88
C(M)	1	1·003	0·1127	29·2	69·4	81·0	
	2	1·682	0·1459	84·4	161·3	73·8	
	3	2·938	0·1929	129·0	80·2	29·4	13·29
O(M)	1	1·209	0·1237	32·2	119·1	87·7	
	2	1·432	0·1342	67·6	24·0	86·3	
	3	3·257	0·2031	123·8	91·7	22·1	14·14

FERRIER, W. G. (1963). *Acta Cryst.* **16**, 1023.

HAM, J. T. (1969*a*). Unpublished work.

HAM, J. T. (1969*b*). Ph. D. Thesis, The Institute of Paper Chemistry, Appleton, Wisconsin.

HANSON, H. P., HERMAN, F., LEA, J. D. & SKILLMAN, S. (1964). *Acta Cryst.* **17**, 1040.

HAUPTMAN, H. & KARLE, J. (1956). *Acta Cryst.* **9**, 45.

JACOBSON, R. A., WUNDERLICH, J. A. & LIPSCOMB, W. N. (1961). *Acta Cryst.* **14**, 598.

KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.

KILLEAN, R. C. G. (1966). *Acta Cryst.* **20**, 547.

MAIR, G. A. (1963). *Structure Factor and Least Squares Programs for the* IBM 1620. Ottawa, Canada, Pure Chemistry Division, National Research Council, October, 1963.

REES, D. A. & SKERRETT, R. J. (1968). *Carbohydrate Res.* In the press.

REEVES, R. E. (1950). *J. Amer. Chem. Soc.* **72**, 1499.

RICHARDS, G. F. (1964). Ph. D. Thesis, University of Iowa, Iowa City, Iowa.

ROLLETT, J. S. & SPARKS, R. A. (1960). *Acta Cryst.* **13**, 273.

STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175.