Acta Cryst. (1980). B**36**, 650–654

The Structure of Gentiobiose

BY D. C. ROHRER

Medical Foundation of Buffalo, Buffalo, NY 14203, *USA*

AND A. SARKO,* T. L. BLUHM AND Y. N. LEE

Department of Chemistry, State University of New York, College of Environmental Science and Forestry, Syracuse, NY 13210, *USA*

(*Received* 16 *July* 1979; *accepted* 1 *November* 1979)

Abstract

The crystal structure of the disaccharide gentiobiose, β-D-glucopyranosyl-(1→6)-β-D-glucopyranose ($C_{12}H_{22}O_{11}$), has been determined using direct methods. The intensities were measured on a CAD-4 diffractometer, with Ni-filtered Cu $K\alpha$ radiation. The crystals are orthorhombic with $a = 8{\cdot}8693$ (5), $b = 22{\cdot}846$ (1), $c = 7{\cdot}2011$ (4) Å and the space group is $P2_12_12_1$. The unit cell contains four molecules, the measured density $D_m = 1{\cdot}57$ Mg m^{-3} and the calculated density $D_c = 1{\cdot}56$ Mg m^{-3}. The final R value, after full-matrix least-squares refinement, is 0·037 for 1739 reflections. Both glucopyranose residues of the molecule are in the $^4C_1$ chair form with normal bond lengths and angles. The conformations of the linkage between the residues and of the free hydroxymethyl group are both *gg*. The *exo*-anomeric effect appears to be important in determining the conformations of the inter-residue linkage and the free glycosidic hydroxyl. All oxygens, except the OH(2′) hydroxyl, the bridge oxygen and the ring oxygens, are hydrogen-bonded, but there are no intramolecular hydrogen bonds.

Introduction

Many polysaccharides of biological importance contain 1→6 linkages, both in the main chain as well as at branch points. The two homopolymeric glucans, dextran and pustulan, are well known examples of this linkage geometry, with all α-(1→6) linkages in the backbone of the former and only β-(1→6) linkages in the latter (Aspinall, 1970). Little is known about the conformational characteristics of either 1→6 linkage, and this investigation of gentiobiose, the dimer residue of pustulan, was undertaken to obtain structural information on the β-(1→6) link.

* To whom correspondence should be addressed.

0567-7408/80/030650-05$01.00

Experimental

Crystal data

Gentiobiose, β-D-glucopyranosyl-(1→6)-β-D-glucopyranose, $C_{12}H_{22}O_{11}$, m.p. 463 K; orthorhombic, space group $P2_12_12_1$, $a = 8{\cdot}8693$ (5), $b = 22{\cdot}846$ (1), $c = 7{\cdot}2011$ (4) Å, $V = 1459{\cdot}1$ Å^3, $Z = 4$, $D_m = 1{\cdot}57$, $D_c = 1{\cdot}56$ Mg m^{-3}.

Single crystals suitable for X-ray diffraction experiments were grown by dissolving 10 mg of β-gentiobiose (Sigma Chemical Company, St Louis, Missouri) with 0·25 ml of distilled water in one arm of a lambda tube to which 1 ml of methanol was added slowly. The other arm of the lambda tube was filled with methanol, and the tube was stoppered and allowed to stand at room temperature for periods of from one week to one month. The resulting crystals were needle-shaped, with approximate dimensions of 0·05 × 0·1 × 0·5 mm.

Unit-cell parameters and the intensity data were measured on a CAD-4 diffractometer using Ni-filtered Cu $K\alpha$ radiation ($\lambda = 1{\cdot}5418$ Å). Lattice parameters were refined by least-squares fit to measured 2θ values for 56 reflections in the interval $50° < 2\theta < 70°$. The intensities were measured using ω–2θ scans. A total of 1739 reflections with $2\theta \leq 150°$ were recorded, of which 88 were below two standard deviations on the measured intensities. The intensities were corrected by the application of Lorentz and polarization factors, but not absorption.

Structure determination and refinement

The locations of the nonhydrogen atoms were derived using *MULTAN* (Germain, Main & Woolfson, 1971) in conjunction with the negative quartet figure of merit (DeTitta, Edmonds, Langs & Hauptman, 1975). The H atoms were located on a difference map calculated at an intermediate stage of least-squares refinement. In the

final cycles of full-matrix least-squares refinement, positional parameters for all the atoms, anisotropic thermal-vibration parameters for the nonhydrogen atoms, and isotropic thermal-vibration parameters for the H atoms were varied. The quantities $(1/\sigma_F^2)$, where σ_F was as defined by Stout & Jensen (1968), but with an instability factor of 0·06, were used to weight the least-squares differences for the observed data; differences for the unobserved data were given zero weight. The largest shift to the e.s.d. in the last cycle of refinement was 1·1. The final values of the residual, $R = \sum||F_o| - |F_c||/\sum|F_o|$, were 0·039 for the observed data and 0·037 for all measured data. The scattering factors were generated from the coefficients given by Cromer & Waber (1974). Final positional parameters are listed in Table 1.*

Table 1. *Fractional atomic coordinates (for nonhydrogen atoms* $\times 10^5$, *for hydrogen atoms* $\times 10^4$)

Numbers in parentheses are the estimated standard deviations.

	x	y	z
C(1)	36218 (20)	26112 (6)	19597 (23)
C(2)	43856 (19)	21779 (7)	32577 (22)
C(3)	55205 (18)	18136 (6)	21904 (24)
C(4)	48267 (18)	15351 (7)	4846 (23)
C(5)	39748 (18)	19946 (6)	−6461 (22)
C(6)	31445 (20)	17245 (7)	−22641 (24)
O(1)	25098 (15)	29177 (6)	29185 (19)
O(2)	50983 (15)	24987 (6)	47008 (18)
O(3)	61077 (14)	13740 (5)	34115 (19)
O(4)	60100 (14)	12947 (5)	−6464 (19)
O(5)	29039 (13)	22856 (5)	5159 (16)
O(6)	20955 (14)	13045 (5)	−15780 (17)
C(1′)	17116 (20)	8805 (7)	−28784 (24)
C(2′)	5225 (21)	4845 (7)	−20401 (26)
C(3′)	−726 (19)	606 (7)	−35044 (29)
C(4′)	−5611 (20)	3775 (7)	−52616 (28)
C(5′)	6590 (20)	8040 (8)	−59267 (25)
C(6′)	1145 (23)	12118 (9)	−74422 (27)
O(2′)	11691 (18)	1676 (7)	−5295 (23)
O(3′)	−13227 (17)	−2491 (6)	−27530 (29)
O(4′)	−8298 (16)	−482 (6)	−66847 (24)
O(5′)	10994 (14)	11752 (5)	−44328 (17)
O(6′)	−10175 (15)	10657 (5)	−67982 (20)
H(1)	4290 (18)	2937 (7)	1488 (23)
H(2)	3493 (26)	1900 (8)	3825 (33)
H(3)	6379 (25)	2057 (8)	1865 (27)
H(4)	4124 (21)	1243 (8)	814 (30)
H(5)	4676 (20)	2294 (7)	−1044 (25)
H(6*A*)	2549 (21)	2010 (8)	−2911 (27)
H(6*B*)	4035 (22)	1546 (9)	−3102 (31)
H(O1)	1890 (30)	3184 (11)	1985 (39)
H(O2)	5785 (30)	2249 (10)	5146 (34)
H(O3)	7126 (30)	1444 (12)	3474 (49)
H(O4)	5861 (26)	950 (11)	−836 (34)
H(1′)	2640 (22)	647 (8)	−3215 (31)
H(2′)	−300 (28)	710 (9)	−1671 (30)
H(3′)	796 (27)	−246 (9)	−3778 (31)
H(4′)	−1507 (23)	592 (8)	−5032 (26)
H(5′)	1553 (32)	579 (9)	−6353 (35)
H(6′*A*)	−378 (25)	968 (10)	−8513 (31)
H(6′*B*)	869 (25)	1428 (10)	−7802 (38)
H(O3′)	−1042 (28)	−586 (10)	−2558 (35)
H(O4′)	−1794 (47)	41 (15)	−7022 (48)
H(O6′)	−651 (27)	1822 (10)	−5961 (40)
H(O2′)	309 (63)	131 (16)	427 (46)

Molecular conformation

The conformation of the gentiobiose molecule is shown in Fig. 1. As expected, both glucose residues are in the $^4C_1$ chair form. The bond lengths and valence angles, shown in Fig. 1, are normal. The average C—C and C—O bond lengths are 1·520 and 1·424 Å, respectively. The anomeric C(1)—OH bond is shortened to 1·393 Å, as in other β-D-glucose residues, and the same is true of the C(1′)—O(6) bond of the glycosidic linkage which is 1·390 Å (Berman, Chu & Jeffrey, 1967). The bonds to O(5) and O(5′) are not longer than average; in fact, the C(1′)—O(5) bond is somewhat shorter at 1·415 Å.

* Lists of structure factors and thermal parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 34951 (11 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

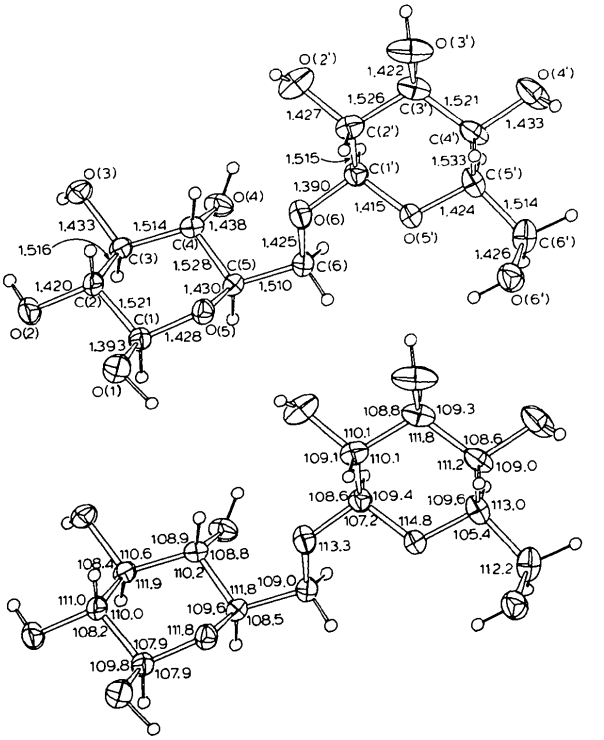

Fig. 1. The molecular conformation of gentiobiose, showing atom numbering, bond lengths (Å) and bond angles (°). The thermal ellipsoids are drawn at the 60% probability level (Johnson, 1965). The estimated standard deviations in the bond lengths range from 0·002 to 0·003 Å and in the bond angles from 0·2 to 0·5°.

The valence-bond angles fall in the expected range. The O—C—O angles at both anomeric carbons, C(1) and C(1′), are 107·9 and 107·2°, respectively, and the angles at the ring oxygens, O(5) and O(5′), are 111·8 and 114·8°, respectively. With the exception of the angle at O(5′), these angles are as expected for β sugars; the angle at O(5′) is somewhat larger (Jeffrey, 1979).

A useful procedure in polysaccharide crystal-structure determinations is molecular modeling based on an *average* monomer residue. A survey of 19 crystal structures of mono- and oligosaccharides containing a total of 27 pyranose rings (over half of which were glucopyranose) showed that the values of bond lengths, valence-bond angles and ring torsion angles fall in fairly narrow ranges (Arnott & Scott, 1972). This results in a well defined *average* sugar residue which has been successfully used in polysaccharide crystal-structure studies (see, for example, Stipanovic & Sarko, 1976). Since most of the gentiobiose bond lengths and angles are very close to the average values, the Arnott–Scott *average* pyranose residue should be suitable for modeling the β-(1→6)-linked polysaccharides. The same is also indicated by the torsion angles, listed in Table 2. The ring torsion angles average to a value of 56·5°, which is within 0·3° of the value for the *average* residue. The ring of the nonreducing residue is slightly flatter than that of the reducing residue, as the corresponding average ring torsion angles of 55·4 and 57·7° indicate.

Table 2. *Ring and selected exocyclic torsion angles* (°)

Torsion angle $A(1)$–$A(2)$–$A(3)$–$A(4)$ viewed from $A(2)$ to $A(3)$ is positive when $A(4)$ is rotated clockwise relative to $A(1)$. The estimated standard deviations in torsion angles not involving hydrogens are 0·2°, those involving hydrogens are 2°.

(*a*) Ring angles for the reducing residue

O(5)–C(1)–C(2)–C(3)	58·7
C(1)–C(2)–C(3)–C(4)	−52·3
C(2)–C(3)–C(4)–C(5)	49·7
C(3)–C(4)–C(5)–O(5)	−54·3
C(4)–C(5)–O(5)–C(1)	64·6
C(5)–O(5)–C(1)–C(2)	−66·4
Average	57·7

(*b*) Ring angles for the non-reducing residue

O(5′)–C(1′)–C(2′)–C(3′)	56·3
C(1′)–C(2′)–C(3′)–C(4′)	−51·3
C(2′)–C(3′)–C(4′)–C(5′)	49·2
C(3′)–C(4′)–C(5′)–O(5′)	−51·7
C(4′)–C(5′)–O(5′)–C(1′)	60·6
C(5′)–O(5′)–C(1′)–C(2′)	−63·3
Average	55·4
Average over both residues	56·5

(*c*) Exocyclic angles

		Observed	Theoretical*
$\omega_{O(5)}$	O(5)–C(5)–C(6)–O(6)	−61·5	−59·6
$\omega_{C(4)}$	C(4)–C(5)–C(6)–O(6)	59·5	61·4
ψ	C(5)–C(6)–O(6)–C(1′)	−156·3	−155·9
$\varphi_{O(5')}$	C(6)–O(6)–C(1′)–O(5′)	−58·3	−140·2
$\varphi_{C(2')}$	C(6)–O(6)–C(1′)–C(2′)	−176·4	101·7
$\omega_{O(5')}$	O(5′)–C(5′)–C(6′)–O(6′)	−53·6	
$\omega_{C(4')}$	C(4′)–C(5′)–C(6′)–O(6′)	66·1	
$\varphi_{O(5)}$	O(5)–C(1)–O(1)–HO(1)	−56	
$\varphi_{C(2)}$	C(2)–C(1)–O(1)–HO(1)	−173	

* From energy minimization (see text).

The exocyclic torsion angles, particularly those for the inter-residue bonds, for the anomeric hydroxyl OH(1), and for the primary hydroxyl OH(6), are of more interest to this study. As shown in Table 2, the angles $\omega_{O(5)}$ and $\omega_{C(4)}$ of the linkage are −61·5 and 59·5°, respectively, indicating that the C(6)—O(6) bond is in an almost perfect *gg* position, *i.e. gauche* to both the O(5)—C(5) and the C(4)—C(5) bonds. The hydroxymethyl group of the nonreducing residue is also near the *gg* position, as indicated by the angles $\omega_{O(5')} = -53{\cdot}6°$ and $\omega_{C(4')} = 66{\cdot}1°$. The *gg* position is regarded as the most stable rotational position for the hydroxymethyl group in a glucose residue (Tvaroska, Pérez & Marchessault, 1978; Pérez & Marchessault, 1979).

The ψ torsion angle of the glycosidic linkage, C(5)—C(6)—O(6)—C(1′), −156·3°, is 24° from the most stable *trans* conformation. This deviation may be caused by nonbonded interactions between the atoms O(5′) and one of the H(6), which, in turn, appear to be the result of the *exo*-anomeric effect. The latter arises because of the interaction of the lone-pair electrons of the glycosidic oxygen with those of the ring oxygen. As the angles $\varphi_{O(5')} = -58{\cdot}3°$ and $\varphi_{C(2')} = -176{\cdot}4°$ show, the C(6)—O(6) bond is *gauche* to the C(1′)—O(5′) bond and *trans* to the C(1′)—C(2′) bond, which is the expected position when the *exo*-anomeric effect is present (Lemieux, 1971; Pérez & Marchessault, 1978). The same is true of the glycosidic OH(1) hydroxyl group, as its angles $\varphi_{O(5)} = -56{\cdot}1°$ and $\varphi_{C(2)} = -173{\cdot}4°$ indicate. Apparently, the rotational positions of both the free glycosidic hydroxyl and the O(6)—C(6) bond of the glycosidic linkage are, at least in part, determined by the *exo*-anomeric effect.

Theoretical values for the linkage conformation angles were obtained with *CRYSP* 79 (Sarko, 1979), an empirical procedure for minimizing conformation energies using Lennard-Jones potential-energy functions. The results are also shown in Table 2. The good agreement between the observed and theoretical ω and ψ angles indicates that the conformation about the C(5)—C(6)—O(6) bonds is mainly determined by nonbonded interactions. On the other hand, the large differences between the observed and theoretical φ angles show that the rotational position about the O(6)—C(1′) bond is determined by other effects. Thus, it is likely that the *exo*-anomeric effect is important here. The conformational energy of the theoretical structure was 3·2 kJ mol^{-1} lower than that of the actual structure, by this calculation. This points out that

Table 3. *Hydrogen-bond distances* (Å) *and angles* (°)

The estimated standard deviations in the last digits are in parentheses.

D–H→A	d_{DA}	d_{DH}	d_{AH}	∠DHA	Symmetry operation on A*
O(6′)–H→O(2)	2·729 (2)	0·84 (3)	1·92 (2)	161 (2)	3,1̄00
O(2)–H→O(1)	2·902 (2)	0·89 (2)	2·10 (3)	148 (2)	3,001
O(1)–H→O(4)	2·772 (2)	1·06 (3)	1·72 (3)	171 (2)	3,1̄00
O(4)–H→O(4′)	2·949 (2)	0·81 (3)	2·15 (3)	169 (2)	2,000
O(4′)–H→O(3′)	2·726 (2)	0·91 (4)	1·82 (4)	177 (3)	2,10̄1̄
O(3′)–H→O(3)	2·710 (2)	0·82 (2)	1·93 (2)	158 (2)	2,00̄1̄
O(3)–H→O(6′)	2·609 (2)	0·92 (3)	1·70 (3)	170 (2)	1,101

* The first number denotes the symmetry operator while the second group of numbers indicates lattice translation. Symmetry operators are: (1) x,y,z; (2) $\frac{1}{2}-x,-y,\frac{1}{2}+z$; (3) $\frac{1}{2}+x,\frac{1}{2}-y,-z$; (4) $-x,\frac{1}{2}+y,\frac{1}{2}-z$.

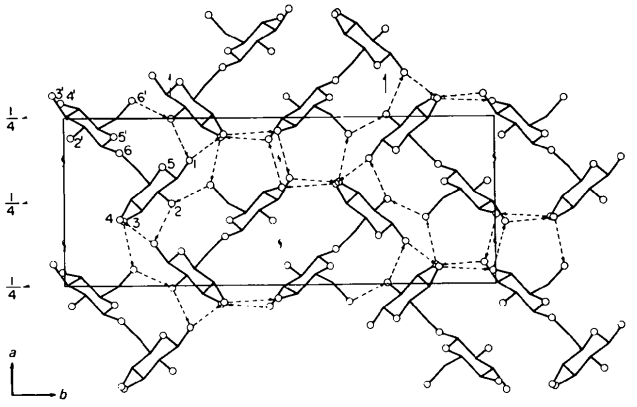

Fig. 2. Projection of the gentiobiose packing and hydrogen-bonding scheme on to the *ab* plane of the unit cell. All hydrogen atoms have been omitted.

determining minimum-energy conformations of di- or higher saccharides and, probably, of polysaccharides by empirical procedures based only on nonbonded interactions, may not be reliable. At least, the *exo*-anomeric effect should be included in some fashion. The quantum-mechanical calculations appear to be better in this regard (Jeffrey, 1979).

Intermolecular packing and hydrogen bonding

The extensive system of hydrogen bonds in the crystal structure of gentiobiose is described in Table 3 and illustrated in Fig. 2. There are no intramolecular hydrogen bonds and neither the glycosidic bridge oxygen nor the ring oxygens are hydrogen-bonded. The absence of intramolecular hydrogen bonding sets gentiobiose apart in this respect from the (1→4)-linked sugars (Quigley, Sarko & Marchessault, 1970; Chu & Jeffrey, 1968). The absence of hydrogen bonding to the bridge oxygen, however, appears to be common to all di- and higher saccharides thus far studied.

All of the hydroxyl groups except OH(2′) are hydrogen-bonded, forming an infinite chain $O(6')_{1,001}\rightarrow O(2)_{3,\bar{1}01}\rightarrow O(1)_{1,000}\rightarrow O(4)_{3,\bar{1}00}\rightarrow O(4')_{4,00\bar{1}}\rightarrow O(3')_{3,000}\rightarrow O(3')_{4,100}\rightarrow O(6')_{4,00\bar{1}}$ that runs principally along the *b* axis of the unit cell but contains hydrogen bonds directed in all three dimensions (*cf.* Fig. 2). The bonds range in length from 2·609 to 2·949 Å, with an average of 2·771 Å. The shortest O···O contact distance to the OH(2′) hydroxyl is 3·324 Å, effectively leaving this hydroxyl completely non-hydrogen-bonded.

The packing of the molecules exhibits no unusual features or short nonbonded contacts, using presently accepted contact-distance criteria (Ramachandran, Ramakrishnan & Sasisekharan, 1963). However, a packing feature may be of potential significance, *viz* a zigzag arrangement of molecules along the screw axis in the **a** direction that could form a polymer chain with very little change in the packing. The resulting chain would have twofold screw symmetry with the dimer as an asymmetric residue and would repeat in approximately 9 Å. It will be interesting to see, once the crystal structure of a β-(1→6)-glucan has been determined, whether this packing feature of gentiobiose and the absence of intramolecular hydrogen bonding actually have any bearing on the polymer conformation.

This work was supported, in part, by the National Science Foundation, grant No. CHE7727749.

References

ARNOTT, S. & SCOTT, W. E. (1972). *J. Chem. Soc. Perkin Trans.* 2, pp. 324–335.

ASPINALL, G. O. (1970). *Polysaccharides*. New York: Pergamon Press.

BERMAN, H. M., CHU, S. C. & JEFFREY, G. A. (1967). *Science*, **157**, 1576–1577.

CHU, S. C. & JEFFREY, G. A. (1968). *Acta Cryst.* B**24**, 830–838.

CROMER, D. T. & WABER, J. T. (1974). *International Tables for X-ray Crystallography*, Vol. IV, Table 2.2B, edited by J. A. IBERS & W. C. HAMILTON. Birmingham: Kynoch Press.

DETITTA, G. T., EDMONDS, J. W., LANGS, D. A. & HAUPTMAN, H. (1975). *Acta Cryst.* A**31**, 472–479.

GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* A**27**, 368–376.

JEFFREY, G. A. (1979). *Am. Chem. Soc. Symp. Ser.* **87**, 50–62.

JOHNSON, C. K. (1965). *ORTEP*. Report ORNL-3794. Oak Ridge National Laboratory, Tennessee.

LEMIEUX, R. U. (1971). *Pure Appl. Chem.* **25**, 527–548.

PÉREZ, S. & MARCHESSAULT, R. H. (1978). *Carbohydr. Res.* **65**, 114–120.

PÉREZ, S. & MARCHESSAULT, R. H. (1979). *Biopolymers*. In the press.

QUIGLEY, G. J., SARKO, A. & MARCHESSAULT, R. H. (1970). *J. Am. Chem. Soc.* **92**, 5834–5839.

RAMACHANDRAN, G. N., RAMAKRISHNAN, C. & SASISEKHARAN, V. (1963). *J. Mol. Biol.* **7**, 95–99.

SARKO, A. (1979). *CRYSP* 79. A Fortran program for molecular refinement. State Univ. of New York.

STIPANOVIC, A. J. & SARKO, A. (1976). *Macromolecules*, **9**, 851–857.

STOUT, G. H. & JENSEN, L. H. (1968). *X-ray Structure Determination*. New York: Macmillan.

TVAROSKA, I., PÉREZ, S. & MARCHESSAULT, R. H. (1978). *Carbohydr. Res.* **61**, 97–106.