

The Crystal and Molecular Structure of β -D-Glucoheptose (D-glycero- β -D-gulo-Heptopyranose)

BY J. A. KANTERS, A. J. H. M. KOCK AND G. ROELOFSEN

Structural Chemistry Group, Rijksuniversiteit Utrecht, Transitorium 3, Padualaan 8, Utrecht, The Netherlands

(Received 8 May 1978; accepted 3 July 1978)

β -D-Glucoheptose (D-glycero- β -D-gulo-heptopyranose), $C_7H_{14}O_7$, is orthorhombic, space group $P2_12_12_1$, with $a = 6.959$ (5), $b = 8.526$ (5), $c = 15.280$ (5) Å, $Z = 4$. The structure was solved by direct methods and refined to a final R of 0.029 for 855 counter reflexions. The gulo-pyranose ring is a distorted 4C_1 chair. The conformations of the C(6)–O(6) and C(7)–O(7) bonds of the glycolic side chain are *trans-gauche* and *gauche-gauche* respectively. All six hydroxyl groups are involved as donor and acceptor in intermolecular hydrogen bonding which consists of an infinite coil winding in the a direction. Two runs of this coil are needed for the identity to be shifted one period along a .

Introduction

In connexion with X-ray diffraction studies of selected carbohydrates in this laboratory, the crystal and molecular structure of β -D-glucoheptose (D-glycero- β -D-gulo-heptopyranose) has been determined. β -D-Glucoheptose belongs to the class of aldohexoses and differs from the common aldohexoses by having a glycolic side chain (–CHOH–CH₂OH). The primary objective of this study was to examine the conformation of this side chain and the hydrogen-bonding scheme.

Experimental

Crystals were grown by slow evaporation from an aqueous ethanolic solution of commercially available β -D-glucoheptose (Sigma Corporation). Preliminary unit-cell dimensions and space-group information were

obtained from photographs. Accurate cell dimensions and intensities were measured on an automatic Nonius CAD-4 diffractometer with Zr-filtered Mo $K\alpha$ radiation ($\lambda = 0.71069$ Å) and the $\omega/2\theta$ scan mode. Crystal data are summarized in Table 1.

Integrated intensities of 951 independent reflexions, of which 855 had $I > 2.5\sigma(I)$, were collected up to $\sin \theta/\lambda = 0.60$ Å⁻¹. Lorentz and polarization corrections were applied, but absorption was neglected.

Determination and refinement of the structure

The observed structure factors were scaled and normalized to E values with a Wilson (1942) plot. The structure was solved with the *MULTAN* system (Main, Woolfson, Lessinger, Germain & Declercq, 1974) using 100 $|E|$ values with $|E| \geq 1.55$. The positions of all non-hydrogen atoms could be deduced from the E map calculated with the phase set with the highest combined figure of merit. Block-diagonal least-squares refinement of positional and isotropic thermal parameters followed by anisotropic refinement reduced R to 0.067. A difference synthesis revealed the positions of all H atoms (peak heights ranging from 0.25 to 0.49 e Å⁻³), which were included in the refinement with constant isotropic thermal parameters equal to those of the carrier atoms. Full-matrix refinement resulted in a final $R = \sum ||F_o| - |F_c|| / \sum |F_o| = 0.029$ and $R_w = [\sum w(|F_o| - |F_c|)^2 / \sum wF_o^2]^{1/2} = 0.018$. The quantity minimized was $\sum w(|F_o| - |F_c|)^2$, with $w = \sigma^{-2}(F_o)$ determined by counting statistics. All shifts in the last cycle were less than 0.15σ for O and C and 0.4σ for H; the goodness-of-fit $\{[\sum w(F_o - F_c)^2 / (m - n)]^{1/2}$, where m is the number of observations and n the number of parameters refined} was 3.06. A final difference

Table 1. Crystal data for β -D-glucoheptose

Molecular formula	$C_7H_{14}O_7$
Formula weight	210.19
Crystal system	Orthorhombic
a (Å)	6.959 (5)
b (Å)	8.526 (5)
c (Å)	15.280 (5)
V (Å ³)	906.60
Z	4
Space group	$P2_12_12_1$
Systematic absences	$h00, h = 2n + 1$ $0k0, k = 2n + 1$ $00l, l = 2n + 1$
D_x (g cm ⁻³)	1.540
λ (Mo $K\alpha$) (Å)	0.71069
μ (Mo $K\alpha$) (cm ⁻¹)	1.50
Crystal dimensions (mm)	$0.5 \times 0.5 \times 0.06$

synthesis showed no peak above $0.20 \text{ e } \text{\AA}^{-3}$. The scattering factors of Cromer & Mann (1968) were used for O and C, and of Stewart, Davidson & Simpson (1965) for H. The final positional parameters are listed in Table 2.* Refinement and subsequent calculations were performed with the XRAY system (Stewart, 1976).

Molecular geometry

The conformation of the molecule and the numbering of the atoms are shown in Fig. 1. Bond distances and angles involving C and O atoms are shown in Table 3. The C—C bonds range from 1.509 (4) to 1.529 (4) Å (mean 1.518 Å). The exocyclic C—O distances range from 1.401 (3) to 1.437 (3) Å (mean 1.418 Å), where the shortest distance relates to the anomeric C(1)—O(1) bond. The endocyclic C(1)—O(5) and C(5)—O(5) bonds are 1.429 (3) and 1.440 (3) Å respectively. The

* Lists of structure factors and thermal parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 33783 (8 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Table 2. Fractional atomic coordinates ($\times 10^4$ for C and O; $\times 10^3$ for H)

The estimated standard deviations are given in parentheses and refer to the last decimal positions of respective values.

	x	y	z
C(1)	-155 (4)	3012 (3)	5632 (2)
C(2)	995 (4)	3737 (3)	6374 (2)
C(3)	918 (4)	5518 (3)	6341 (2)
C(4)	1397 (4)	6121 (3)	5433 (2)
C(5)	177 (4)	5302 (3)	4757 (2)
C(6)	671 (4)	5740 (3)	3813 (2)
C(7)	-742 (5)	5069 (4)	3157 (2)
O(1)	278 (3)	1407 (2)	5612 (1)
O(2)	372 (3)	3174 (2)	7197 (1)
O(3)	-982 (3)	6009 (2)	6580 (1)
O(4)	3344 (3)	5838 (2)	5232 (1)
O(5)	479 (3)	3635 (2)	4813 (1)
O(6)	623 (3)	7418 (2)	3730 (1)
O(7)	-2646 (3)	5584 (2)	3304 (1)
H(O1)	-78 (4)	91 (4)	526 (2)
H(O2)	-56 (5)	360 (4)	739 (2)
H(O3)	-100 (4)	654 (3)	695 (2)
H(O4)	397 (4)	626 (3)	556 (2)
H(O6)	181 (4)	775 (3)	373 (2)
H(O7)	-310 (5)	500 (4)	359 (2)
H(C1)	-183 (4)	328 (3)	576 (2)
H(C2)	229 (4)	345 (3)	631 (2)
H(C3)	192 (4)	586 (3)	675 (2)
H(C4)	120 (4)	721 (3)	542 (2)
H(C5)	-114 (4)	549 (3)	486 (2)
H(C6)	207 (4)	541 (3)	370 (2)
H(C7)	-44 (4)	540 (3)	260 (2)
H(C7)'	-63 (4)	394 (3)	315 (2)

C—O distances in the hemiacetal sequence C(5)—O(5)—C(1)—O(1)—H agree well with observations on α - and β -pyranosides (Jeffrey & Takagi, 1977; Arnott & Scott, 1972) and with theory (Jeffrey, Pople & Radom, 1972, 1974), in that the anomeric C(1)—O(1) is shortened and the difference in endocyclic C—O bond lengths in β -pyranosides is smaller than in α -pyranosides. Though the bond angles have a wide range (105.4 to 114.0° , mean 109.9°), the majority are within 2° of the ideal tetrahedral value. Notable exceptions are C(4)—C(5)—C(6) 114.0 , O(5)—C(5)—C(6) 105.4 , C(1)—O(5)—C(5) 112.0 and O(5)—C(1)—O(1) 106.1° . As in many pyranosides, C(4)—C(5)—C(6) is enlarged at the expense of O(5)—C(5)—C(6) (Arnott & Scott, 1972; Gress & Jeffrey, 1977). The bond angles in the hemiacetal group, C(1)—O(5)—C(5) and O(1)—C(1)—O(5), agree with the average values of eight β -pyranosides: 111.6 and 107.3° respectively (Arnott & Scott, 1972). The C—H and O—H distances have mean values of 0.99 and 0.82 Å respectively. The bond angles involving H atoms have mean values distributed as follows over the different classes: H—C—C 109 , H—C—O 109 , H—O—C 109 and H—C—H 106° .

Molecular conformation

The pyranose ring is distorted from the ideal 4C_1 chair as follows from the ring torsion angles (range 49.9 to 64.3° , Table 4) and the displacements of the atoms from the least-squares planes through opposite bonds. The distortion is in the direction of a twist-boat form as follows from the Cremer & Pople (1975) puckering parameters: $q_2 = 0.061$ Å, $q_3 = 0.564$ Å, $\varphi_2 = 39.2^\circ$, $\theta = 6.2^\circ$, $Q = 0.568$ Å. The exocyclic torsion angles (Table 4) are close to the angles for the ideal *gauche* (60°) or *trans* (180°) arrangements. The configuration of the O atoms at C atoms in the chain C(3)—C(4)—C(5)—C(6)—C(7) is identical to that of D-arabinitol. The conformation of this pentitol fragment and the

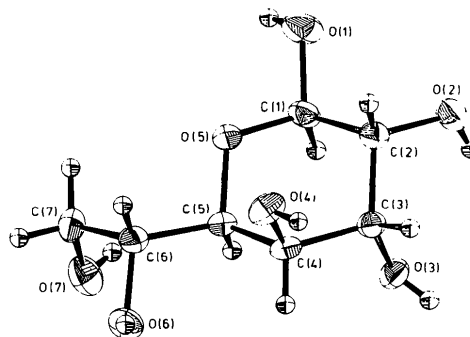


Fig. 1. Molecular conformation and atomic numbering of β -D-glucoheptose. Non-hydrogen thermal ellipsoids enclose 50% probability.

planarity of the C atom chain compare well with the conformation and planarity of D-arabinitol in the structure of DL-arabinitol (Hunter & Rosenstein, 1968).

The conformations of C(6)–O(6) and C(7)–O(7) of the side chain are *trans-gauche* and *gauche-gauche* respectively. The *tg* conformation of C(6)–O(6) seldom occurs in pyranoses (Longchambon, Ohanesian, Avenel & Neuman, 1975), where the *gg* and *gt* conformations are predominant in nearly equal proportions. However, in galactosides, with axially oriented O(4), the *gg* conformation is unfavoured because of the *peri* interaction between synaxial O atoms. Accordingly, the distribution over the possible *gg*, *gt* and *tg* conformations in the structures of 15 known galactosides is 2, 8 and 5, indicating that there is a slight preference for the *gt* conformation. The two *gg* forms are exceptional because of the presence of an intramolecular O(4)···O(6) bond. In D-glucoheptose the *gt* conformation is ruled out because it would result in a short O(4)···C(7) contact. The conformation around C(6)–C(7) is *gg* with torsion angles of 59.4 and 60.5°.

Table 3. Bond distances (Å) and bond angles (°)

The estimated standard deviations are given in parentheses and refer to the last decimal positions of respective values.

C(1)–C(2)	1.519 (4)	C(2)–O(2)	1.414 (3)
C(2)–C(3)	1.520 (4)	C(3)–O(3)	1.434 (3)
C(3)–C(4)	1.517 (4)	C(4)–O(4)	1.410 (3)
C(4)–C(5)	1.509 (4)	C(5)–O(5)	1.440 (3)
C(5)–C(6)	1.529 (4)	C(6)–O(6)	1.437 (3)
C(6)–C(7)	1.516 (4)	C(7)–O(7)	1.414 (4)
C(1)–O(1)	1.401 (3)	C(1)–O(5)	1.429 (3)
O(1)–C(1)–C(2)	107.5 (2)	C(4)–C(5)–C(6)	114.0 (2)
C(1)–C(2)–C(3)	111.3 (2)	C(4)–C(5)–O(5)	109.5 (2)
C(1)–C(2)–O(2)	111.3 (2)	C(6)–C(5)–O(5)	105.4 (2)
C(3)–C(2)–O(2)	110.9 (2)	C(2)–C(1)–O(5)	109.8 (2)
C(2)–C(3)–C(4)	111.2 (2)	C(5)–O(5)–C(1)	112.0 (2)
C(2)–C(3)–O(3)	108.4 (2)	O(5)–C(1)–O(1)	106.1 (2)
C(4)–C(3)–O(3)	109.7 (2)	C(5)–C(6)–C(7)	112.7 (2)
C(3)–C(4)–C(5)	110.2 (2)	C(5)–C(6)–O(6)	108.7 (2)
C(3)–C(4)–O(4)	110.6 (2)	C(7)–C(6)–O(6)	107.6 (2)
C(5)–C(4)–O(4)	108.2 (2)	C(6)–C(7)–O(7)	112.7 (2)

Table 4. Endocyclic and exocyclic torsion angles (°)

The torsion angle A(1)–A(2)–A(3)–A(4) is viewed along A(2)–A(3), with a clockwise rotation of A(1) to A(4) taken to be positive.

Endocyclic			
O(5)–C(1)–C(2)–C(3)	54.2 (4)	C(3)–C(4)–C(5)–O(5)	–57.6 (4)
C(1)–C(2)–C(3)–C(4)	–49.9	C(4)–C(5)–O(5)–C(1)	64.3
C(2)–C(3)–C(4)–C(5)	51.5	C(5)–O(5)–C(1)–C(2)	–62.2
Exocyclic			
O(1)–C(1)–C(2)–C(3)	169.2	O(4)–C(4)–C(5)–C(6)	–54.3
O(1)–C(1)–C(2)–O(2)	–66.4	O(4)–C(4)–C(5)–O(5)	63.5
C(1)–C(2)–C(3)–O(3)	70.8	C(4)–C(5)–C(6)–O(6)	–53.2
O(2)–C(2)–C(3)–C(4)	–174.5	C(5)–O(5)–C(1)–O(1)	–178.1
O(2)–C(2)–C(3)–O(3)	–53.8	C(5)–C(6)–C(7)–O(7)	59.4
C(2)–C(3)–C(4)–O(4)	–68.1	O(5)–C(5)–C(6)–O(6)	–173.3
O(3)–C(3)–C(4)–C(5)	–68.4	O(6)–C(6)–C(7)–O(7)	–60.5
O(3)–C(3)–C(4)–O(4)	171.9		

The H atoms on adjacent C atoms of the pyranose ring have the common *trans* or *gauche* orientation (Table 5). As is usual in saccharide structures (Jeffrey, McMullan & Takagi, 1977; Kanters, Gaykema & Roelofsen, 1978; Fries, Rao & Sundaralingam, 1971; Kanters, Batenburg, Gaykema & Roelofsen, 1978) the H–C–O–H torsion angles of the ring (Table 5) show large deviations from the ideal value of 60°, which would be favoured in the isolated molecule (Sundaralingam, 1968; Takagi & Jeffrey, 1977). The resulting increase in repulsion by the eclipsing of vicinal H atoms is balanced by the gain of hydrogen-bond energy through participation of the hydroxyl groups in intermolecular hydrogen bonding.

Hydrogen bonding

All hydrogen bonds are intermolecular (Table 6). The six O–H groups each function once as donor and acceptor; the ring O atom does not accept a hydrogen bond. The H···O distances, adjusted so as to normalize the covalent O–H distance to 0.97 Å, established by neutron diffraction, are in the narrow range 1.73–1.85 Å. Although generally the anomeric O atom

Table 5. Torsion angles (°) involving the hydrogen atoms

H(C1)···H(C2)*	174	H(O1)···H(C1)†	43
H(C2)···H(C3)	–48	H(O2)···H(C2)	159
H(C3)···H(C4)	–71	H(O3)···H(C3)	7
H(C4)···H(C5)	–60	H(O4)···H(C4)	58
H(C5)···H(C6)	–174	H(O6)···H(C6)	–13
H(C6)···H(C7)	–60	H(O7)···H(C7)	148
H(C6)···H(C7)′	56	H(O7)···H(C7)′	32

* Refers to the torsion angle H(C1)–C(1)–C(2)–H(C2).

† Refers to the torsion angle H(O1)–O(1)–C(1)–H(C1).

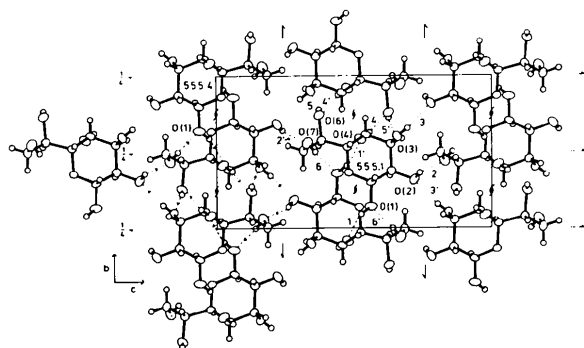


Fig. 2. A view of the molecular packing and hydrogen bonds of β -D-glucoheptose seen along *a*. Hydrogen bonds, related to molecule 555.1, are numbered according to Table 6. The repeating sequence, originating from molecule 555.4, is indicated at the left-hand side.

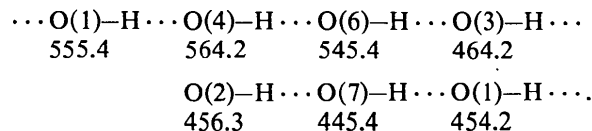
Table 6. *Hydrogen-bond geometry*

		O—H	H...O	H...O*	O...O	O—H...O	Symmetry operation†
1	O(1)—H(O1)...O(4')	1.01 (3) Å	1.78 (3) Å	1.82 Å	2.672 (3) Å	146 (3)°	456.3
2	O(2)—H(O2)...O(7')	0.80	1.99	1.85	2.754	158	465.2
3	O(3)—H(O3)...O(2')	0.73	1.96	1.73	2.662	164	556.4
4	O(4)—H(O4)...O(6')	0.76	1.94	1.74	2.691	171	566.3
5	O(6)—H(O6)...O(3')	0.87	1.93	1.84	2.758	159	566.3
6	O(7)—H(O7)...O(1')	0.73	2.05	1.83	2.776	171	456.3

* Corrected by expanding the covalent O—H bond distances to the neutron diffraction value of 0.97 Å in the direction of the bond.

† The symmetry operation is performed on the O' atoms. The first three digits specify the lattice translations. The last digit indicates one of the following symmetry operations. (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) $-x, \frac{1}{2} + y, \frac{1}{2} - z$.

should be a relatively weak acceptor as concluded from theoretical calculations on methanediol (Tse & Newton, 1977) and from available experimental data (Jeffrey & Lewis, 1978) the relevant H...O distance in β -D-glucoheptose is 1.83 Å, which is in the range normally found for ...O—H...O—H chains. Each molecule is bonded to eight neighbours by twelve hydrogen bonds to form a closely knit three-dimensional network (Fig. 2). The hydrogen bonds form an uninterrupted coiled chain with the following repeating sequence:



When two sequences are completed the chain continues at the same equivalent position, translated one period in the a direction.

References

- ARNOTT, S. & SCOTT, W. E. (1972). *J. Chem. Soc. Perkin Trans.* 2, pp. 324–335.
- CREMER, D. & POPLE, J. A. (1975). *J. Am. Chem. Soc.* **97**, 1354–1358.
- CROMER, D. T. & MANN, J. B. (1968). *Acta Cryst.* **A24**, 321–324.
- FRIES, D. C., RAO, S. T. & SUNDARALINGAM, M. (1971). *Acta Cryst.* **B27**, 994–1005.
- GRESS, M. E. & JEFFREY, G. A. (1977). *Acta Cryst.* **B33**, 2490–2495.
- HUNTER, F. D. & ROSENSTEIN, R. D. (1968). *Acta Cryst.* **B24**, 1652–1660.
- JEFFREY, G. A. & LEWIS, L. (1978). *Carbohydr. Res.* **60**, 179–182.
- JEFFREY, G. A., McMULLAN, R. K. & TAKAGI, S. (1977). *Acta Cryst.* **B33**, 728–737.
- JEFFREY, G. A., POPLE, J. A. & RADOM, L. (1972). *Carbohydr. Res.* **25**, 117–131.
- JEFFREY, G. A., POPLE, J. A. & RADOM, L. (1974). *Carbohydr. Res.* **38**, 81–95.
- JEFFREY, G. A. & TAKAGI, S. (1977). *Acta Cryst.* **B33**, 738–742.
- KANTERS, J. A., BATENBURG, L. M. J., GAYKEMA, W. P. J. & ROELOFSEN, G. (1978). *Acta Cryst.* **B34**, 3049–3053.
- KANTERS, J. A., GAYKEMA, W. P. J. & ROELOFSEN, G. (1978). *Acta Cryst.* **B34**, 1873–1881.
- LONGCHAMON, F., OHANESSIAN, J., AVENEL, D. & NEUMAN, A. (1975). *Acta Cryst.* **B31**, 2623–2627.
- MAIN, P., WOLFSON, M. M., LESSINGER, L., GERMAIN, G. & DECLERCQ, J. P. (1974). *MULTAN 74. A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data*. Univ. of York, England, and Louvain-la-Neuve, Belgium.
- STEWART, J. M. (1976). The XRAY 76 system. Tech. Rep. TR-446. Computer Science Center, Univ. of Maryland, College Park, Maryland.
- STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.
- SUNDARALINGAM, M. (1968). *Biopolymers*, **6**, 189–213.
- TAKAGI, S. & JEFFREY, G. A. (1977). *Acta Cryst.* **B33**, 3033–3040.
- TSE, Y.-C. & NEWTON, M. D. (1977). *J. Am. Chem. Soc.* **99**, 611–613.
- WILSON, A. J. C. (1942). *Nature (London)*, **150**, 151–152.