β-D-Cellotetraose Hemihydrate as a Structural Model for Cellulose II. An X-ray Diffraction Study

Katrin Gessler,[†] Norbert Krauss,[†] Thomas Steiner,[†] Christian Betzel,[‡] Anatole Sarko,[§] and Wolfram Saenger*,[†]

Contribution from the Institut für Kristallographie, Freie Universität Berlin, Takustrasse 6, D-14195 Berlin, Germany, European Molecular Biology Laboratory Outstation, Deutsches Elektronen Synchrotron (DESY), Notkestrasse 85, D-22603 Hamburg, Germany, and Chemistry Department, SUNY College of Environmental Science and Forestry, 1 Forestry Drive, Syracuse, New York 13210

Received June 2, 1995[⊗]

Abstract: β-D-Cellotetraose crystallizes in the triclinic space group $P1$ with two independent molecules **A** and **B** and one water molecule **W** in the unit cell with dimensions $a = 8.023(1)$, $b = 8.951(2)$, $c = 22.445(2)$ Å, $\alpha = 89.26(2)$, $\beta = 85.07(1)$, $\gamma = 63.93(1)°$. The crystal structure was determined by Patterson search and direct methods and refined to $R = 6.2\%$ using 3124 reflections $F \geq 1\sigma F$ collected with X-rays from rotating anode generator and synchrotron. Of the expected 28 O−H hydroxyl hydrogens from molecules **A** and **B**, 12 could be located from difference Fourier maps; configurational disorder is observed for anomeric O1 of molecule **A**, 75% β and 25% α. β-D-Cellotetraoses **A** and **B** are arranged antiparallel, the water **W** connects molecules **B** through hydrogen bonds. The orientations of all C6−O6 bonds are *gauche*, *trans*; D-glucopyranoses are more strained in **B** than in **A**, with average Cremer−Pople parameters $Q_2;\Theta$ for **A**, 0.61 Å; 2° and for **B** 0.59 Å; 13°. The conformations of **A** and **B** are stabilized by intramolecular three center hydrogen bonds O3H···O5′,O6′ and the crystal structure shows systematic intermolecular two-center hydrogen bonds involving O2H and O6H, and, most notably, systematic C−H···O interactions to all O4 oxygen atoms, viz. C5b-H···O4a and C3a-H···O4b. If a subcell is constructed of the two central D-glucopyranoses 2 and 3 in molecules **A** and **B**, the obtained cell dimensions are identical to those of polymer cellulose II. Based on this subcell we propose a new model for cellulose II which has all O6 groups in *gt* position and agrees with spectroscopic data.

Introduction

Cellulose, the major structural component of higher plants and the largest biomass on earth, is a linear polymer formed by β-1,4-linked D-glucopyranose residues. The cellulose molecules are arranged in fibers which pack in bundles. Depending on the origin of the sample, the fiber bundles have different degrees of order. Well prepared samples are up to 80% crystalline and can be used for structure analysis by X-ray fiber diffraction methods.[1,2]

The first X-ray studies on cellulose were reported 82 years ago,[3,4] but the diffuse diffraction patterns permitted to derive a molecular structure model only after fiber diffraction techniques were sufficiently advanced and combined with potential energy conformational and packing analysis. It was shown[5−7] for cellulose I and for cellulose II, which is obtained from cellulose I by treatment with alkali (mercerization), that the β-1,4-linked D-glucopyranoses are in the $^4C_1$ chair conformation. Since in both cellulose I and II, the molecules are located on crystallographic 2_1 screw axes in monoclinic unit cells with space group $P2_1$ (Table 1), the D-glucopyranose units are alternatively rotated by 180° so that intramolecular O3···O5′ hydrogen bonds can form and stabilize the molecular conformation of cellulose (the prime indicates an atom in the adjacent unit).

In cellulose I, the two independent molecules in the crystal unit cell are parallel and arranged in sheets which stack parallel to each other. Intermolecular hydrogen bonds are formed only within but not between the sheets.[7−10] In cellulose II, the two independent molecules are antiparallel and pack in the unit cell given in Table 1. The conformations of the D-glucopyranose units and the orientations about the interglucose links are identical for the two molecules but the orientations of the C6−O6 bonds are different, *gauche*, *trans* in one and *trans*, *gauche* in the antiparallel molecule. The molecules are again arranged in sheets but since they are tilted about their long axes, the hydrogen bonding pattern extends also between the sheets to form a three-dimensional network.

The two different orientations of the C6−O6 bonds in cellulose II are under dispute. This is because theoretical studies,[11,12] ^{13}C CPMAS NMR,[13] and polarized infrared spectra[14] suggested the occurrence of only *one* orientation of the C6−O6 bond, but the spectroscopic data cannot differentiate between *gauche*,*trans* and *trans*,*gauche* forms. The dilemma can be solved by a single crystal study of β-D-cellotetraose hemihydrate. It represents the smallest reasonable model for cellulose II

[†] Freie Universität Berlin.
[‡] European Molecular Biology Laboratory Outstation.
[§] SUNY College of Environmental Science and Forestry.
[⊗] Abstract published in *Advance ACS Abstracts*, November 1, 1995.

(1) Atalla, R. H. In *Structure of Cellulose*; ACS Symposium Series; 1987; Vol. 340.
(2) Blackwell, J., et al. *Fiber Diffraction Methods*; ACS Symposium Series; Washington, DC, 1980; p 315.
(3) Nishikawa, S.; Ono, S. *Proc. Tokyo Math. Pys. Soc.* **1913**, *7*, 131.
(4) Polanyi, M. *Naturwissenschaften* **1921**, *18*, 337.
(5) Kolpak, F. J.; Blackwell, J. *Macromolecules* **1976**, *9*, 273.
(6) Stipanovic, A. J.; Sarko, A. *Macromolecules* **1976**, *9*, 851.
(7) Sarko, A.; Muggli, R. *Macromolecules* **1974**, *7*, 486.
(8) Gardner, K. H.; Blackwell, J. *Biopolymers* **1974**, *13*, 1975.
(9) Claffey, W. B.; Blackwell, J. *Biopolymers* **1976**, *15*, 1903.
(10) Takahashi, Y.; Matsunaga, H. *Macromolecules* **1991**, *24*, 3968.
(11) Pertsin, A. J.; Kitaigorodsky, A. I. *The Atom Atom Potential Method: Application to Organic Molecular Solids*; Springer-Verlag: Berlin, 1987; p 204.
(12) Simon, I.; Glasser, L.; Scheraga, H. A.; Manley, R. St. J. *Macromolecules* **1988**, *21*, 991.
(13) Horii, F.; Hirai, A.; Kitamaru, R. In *Polymers for Fibers and Elastomers*; ACS Symposium Series; No. 260, 1984; p 27.
(14) Marchessault, R. H.; Liang, C. J. *J. Polym. Sci.* **1960**, *43*, 71.

0002-7863/95/1517-11397$09.00/0

Table 1. Crystallographic Data for Cellulose Structures and for β-D-Cellotetraose Hemihydrate

cellulose type	a (Å)	b (Å)	c (Å)	α (deg)	β (deg)	γ (deg)	space group	V (Å³)	D_x (g/cm³)
				Cellulose I					
ref 8	16.34	15.72	10.38	90	90	97.0	$P2_1$	2433	1.69
ref 9	8.18	7.86	10.34	90	90	97.0	$P2_1$	659.8	1.63
				Cellulose II					
ref 5	8.01	9.04	10.36	90	90	117.1	$P2_1$	667.8	1.61
ref 6	7.94	9.09	10.31	90	90	117.3	$P2_1$	661.2	1.63
				β-D-Cellotetraose					
setting 1[a]	8.023	8.951	22.445	89.26	85.07	63.93	$P1$	1443	1.56
setting 2[a]	8.023	8.951	22.445	89.26	94.93	116.07	$P1$	1443	1.56
ref 19	8.045	9.003	22.51	89.66	94.93	115.80	$P1$	1461	1.54
				Subcell[a]					
in β-D-cellotetraose	8.02	8.95	10.35	91.25	90.69	117.07	$P1$	667	1.61

[a] This work.

because the two compounds have very similar ^{13}C CPMAS NMR and infrared absorption spectra,[13,14] and their unit cell constants determined by X-ray powder[15] and single crystal work[16,17] are nearly identical. The first diffraction patterns of single crystals of β-D-cellotetraose hemihydrate were recorded in the late 1960s. Because the weak diffraction obtained from the tiny crystals permitted only to calculate a two dimensional electron density map, the structure determination was incomplete and even missed the hemihydrate water.[16,17] In a preliminary publication,[18] we presented the structure of β-D-cellotetraose hemihydrate at 1.1 Å resolution, followed by a report of Chanzy and co-workers[19]; their results are largely identical to ours, which are now based on new data extending to 0.95 Å resolution.

Experimental Section

X-ray Experiments. Crystals were grown by the method of vapour diffusion in "sitting drop" crystallization experiments. The precipitation buffer (10 μL) containing 20% 2-methyl-2,4-pentanediol and 0.2 M sodium acetate was added to the same volume of an aqueous solution of 150 mg/mL β-D-cellotetraose (Seigaku Coop., Japan) at 18 °C. Thin platelets of the size 0.4 × 0.1 × 0.01 mm³ were obtained after 12 weeks.

Although the crystals are stable at atmosphere, X-ray work was performed on a crystal mounted in a glass capillary with some mother liquor. The unit cell dimensions were determined from the diffraction angles of 81 strong reflections measured with an Enraf-Nonius FAST area detector mounted on an FR 571 X-ray generator with rotating anode, Ni-filtered CuK_α radiation, λ = 1.542 Å (Table 2). X-ray data (4328) were collected with an Enraf-Nonius Turbo-CAD4 diffractometer mounted on the same generator yielding 2165 unique reflections $F_o \geq 1\sigma(F_o)$ to a resolution of 1.2 Å. In addition, 8384 X-ray diffraction data were collected with synchrotron radiation at beamline X31 (DESY, EMBL-Outstation, Hamburg), λ = 0.75 Å, with a Mar Research image plate detector (radius 90 mm), crystal to detector distance 82.5 mm, $\delta\phi = 5°$, resolution 7.0–0.95 Å and merged (R_{merge} = 0.095) to 2487 unique data with $F_o \geq 1\sigma(F_o)$. Diffractometer data from 8.0 to 1.5 Å and synchrotron data from 2.0 to 0.95 Å resolution, respectively, were merged to yield 3124 unique reflections $F_o \geq 1\sigma(F_o)$, R_{merge} = 0.085. This is higher than the final R-factor of 0.062 (see below) probably because we merged data sets from two different, very small crystals collected with different instruments and wavelengths, and because errors in measurements are partially compensated by temperature factors. After transformation to obtuse angle γ, the unit cell is comparable to that of cellulose II[5] (except for the c-axis) and to that derived by the group of Chanzy, see Table 1.

Table 2. Crystal Data and Refinement Parameters for β-D-Cellotetraose Hemihydrate

chemical formula	$(C_{24}H_{42}O_{22})\cdot 0.5H_2O$
formula weight	691
space group	triclinic, $P1$, Z = 2
unit cell constants	a = 8.023(1), b = 8.951(2), c = 22.445(2) Å
	α = 89.26(2), β = 85.07(1), γ = 63.93(2)°
volume	1443.4(4) Å³
calculated density	1.56g/cm³
CAD4 data (λ = 1.542 Å)	4328 reflections
	unique reflections: 2165 ≥ 1σ(F_o)
synchrotron (λ = 0.75 Å)	8384 reflections R_{symm} = 0.095
	unique reflections: 2487 ≥ 1σ(F_o)
data set in refinement	3124 data (resolution 8.0–0.95 Å)
	R_{merge} = 0.082 for 884 data
refinement in two blocks	3124 reflections ≥ 1σ and 818 variables
final $R(F_o \geq 1\sigma(F_o))$	0.062
wR_2 (all data)	0.062
$(\Delta/\sigma)_{max}$	0.03 for hydrogens, 0.01 for non-hydrogens
$\Delta(\rho)_{max}$, $\Delta(\rho)_{min}$	0.42, −0.49e/Å³

Structure Determination and Refinement. The crystal structure was solved by a combination of Patterson search and direct methods using programs PATSEE[20] and SIR92).[21] A β-D-cellotetraose molecule was constructed with the published coordinates of β-D-cellobiose[22] and served as a model in the Patterson search routine. The best solution with one molecule per unit cell was used to calculate an initial set of phases for direct methods.

Consecutive cycles of least-squares refinement and difference Fourier analyses gradually revealed the positions of the atoms of the second β-D-cellotetraose in the unit cell and of the water molecule (program SHELX76[23]), guided by inspection of the electron density maps with the computer graphics program FRODO version E4.4.[24] Twelve of the 28 hydroxyl H-atoms could be located from the difference electron density map; their positions were refined without constraints and have been normalized to an O–H distance of 0.98 Å. Hydrogen atoms bonded to carbon were placed in their calculated positions, at a C–H distance of 1.08 Å. Configurational disorder was observed for the anomeric hydroxyl group of the β-D-cellotetraose molecule **A**, whereas in molecule **B** and otherwise all atoms

(15) Wellard, H. J. *J. Polymer Sci.* **1954**, *13*, 471.
(16) Poppleton, B. J.; Mathieson, A. McL. *Nature* **1968**, *219*, 1046.
(17) Poppleton, B. J.; Gatehouse, B. M. *J. Polym. Sci. Phys. Ed.* **1972**, *10*, 375.
(18) Gessler, K.; Krauss, N.; Steiner, Th.; Sandmann, C.; Betzel, Ch.; Saenger, W. *Science* **1994**, *266*, 1027.
(19) Raymond, S.; Heyraud, A.; Tran Qui, D.; Kvick, Å.; Chanzy, H. *Macromolecules* **1995**, *28*, 2096.
(20) Egert, E.; Sheldrick G, M. *Acta Crystallogr. Sect.* A **1985**, *41*, 262.
(21) Altomare, A.; Casparano, G.; Giacovazzo, C.; Guagliardi, A.; Burla, M. C.; Polidori, G.; Camalli, M. *J. Appl. Cryst.* **1994**, *27*, 435.
(22) Jacobson, R. A.; Wunderlich, J. A.; Lipscomb, N. *Acta Crystallogr.* **1961**, *14*, 598.
(23) Sheldrick, G. M. SHELX76, Program for Crystal Struture Determination, University of Cambridge: UK, 1976.

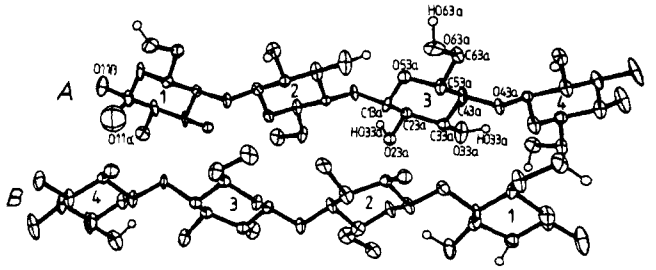

Figure 1. Molecules **A** (x,y,z) and **B** ($x-1,y,z$) projected on the (120) plane. The labeling of the non-H atoms and of the glucose residues are shown. H-atoms were omitted for clarity exept for the O−H atoms located in difference Fourier maps; they are labeled in the same way as the C- or O-atoms to which they are bonded. There is configurational disorder at the anomerinc C1a, with O11a(α) at 25% and O11a(β) at 75% occupation. ORTEP plot with thermal ellipsoids at the 50% probability level.

are well ordered. Refinement converged with anisotropic displacement parameters except for 10 non-hydrogen atoms for which the anisotropic refinement was not possible, at an R-value of 0.062 for 3124 unique reflections $F_o \geq 1\sigma(F_o)$.

New Refinement of Cellulose II with X-ray Fiber Diffraction Data. A model of the cellulose II structure was constructed using the two central residues of both molecules **A** and **B** (see the subcell described in Table 1). The residues were placed in the unit cell of cellulose II, and their positions were adjusted to establish polymer chain continuity and to conform to the 10.31 Å fiber repeat of cellulose II. As is evident from the characteristics of the subcell (see Figure 2), only slight adjustments were necessary to reach a reasonable starting model. The unit cell dimensions and X-ray intensity data were those previously determined.[6] The same refinement method (program PS79[25]) was used, but with one important change in procedure: all structural parameters of both chains of the unit cell were now independently variable, and the space group was assumed to be $P1$.

The refinement was carried out in two stages. During the first stage, the positions of the four glucopyranose residues on the two polymer chains remained invariant, only the chain positions (i.e., rotations of the chains about their axes and the translational position of the center chain) and the rotational orientations of the four O6 atoms were allowed to vary. This was followed by the second refinement stage in which the residues were allowed to relax by letting all torsion angles of the chains vary within narrow limits (about ±5°). Within each stage the refinement was, additionally, carried out in two steps. In the first step, no stereochemical constraints were imposed in order to establish whether a reasonable structure was possible based on X-ray data alone. This refinement was carried out by minimizing the usual weighted crystallographic residual R''. It was then followed by the second step in which stereochemical constraints were imposed to remove any unreasonable structural features that may have developed during the first refinement, such as short nonbonded contacts. The refinement with stereochemical constraints was carried out by minimizing the function $\Phi = fR'' + (1-f)E$, where R'' was expressed in percent, E was a measure of the potential energy of the model, and f was a variable fraction (kept at 0.95 in this instance). All refinements were terminated when the applicable refinement criterion reached a minimum value which subsequently changed by less than 0.05%.

Results

Crystal Structure Analysis of β-D-Cellotetraose. The crystal unit cell contains two symmetry independent β-D-cellotetraose molecules **A** and **B** and one water molecule **W**. Fractional atomic coordinates, anisotropic displacement parameters, bond lengths and angles, observed and calculated structure factor amplitudes, and atomic coordinates of a new model for cellulose II are deposited as supplementary material.

The atom labeling is shown in Figure 1 for molecules **A** and **B**, with C23a meaning carbon atom 2 of glucose residue 3 of molecule **A**. Hydrogen bonded sheets containing only molecules **A** are labeled **AA**; those with alternating molecules **A** and **B** are labeled **AB**.

Crystal Packing. In c-direction, the lath-shaped molecules are aligned head-to-tail to form pseudopolymeric strands (Figure 2). Only slight rearrangement would be required to actually form the covalent bond between O1 and C4 to obtain the typical β-1,4-linkage of cellulose II. Neighboring stands in α-direction are hydrogen bonded to form sheets **AA** and **BB** parallel to the a-c-plane, Figure 3. These sheets stack along the b-axis in antiparallel orientation, **AA** in (010) and **BB** in (020), and are displaced relative to each other by 2.5 Å in c-direction. The water molecule is located in sheet **BB** but is also hydrogen bonded to sheet **AA**. The β-D-cellotetraose laths are tilted with respect to the a,c plane (Figure 3) so that, in addition, hydrogen bonds connect the sheets **A** and **B** to form a third set of sheets **AB** parallel to (120). This implies that the crystal structure is stabilized by a three-dimensional network of hydrogen bonds.

Molecular Structure. The bond angles and distances within the D-glucopyranose units do not deviate significantly from standard values. The interglucose β-1,4 links are defined by C1−O4′ and C4−O4 with average distances of 1.40 Å, 1.45 Å, and the angle C1−O4′−C4′ is 116.5°, see Table 3.

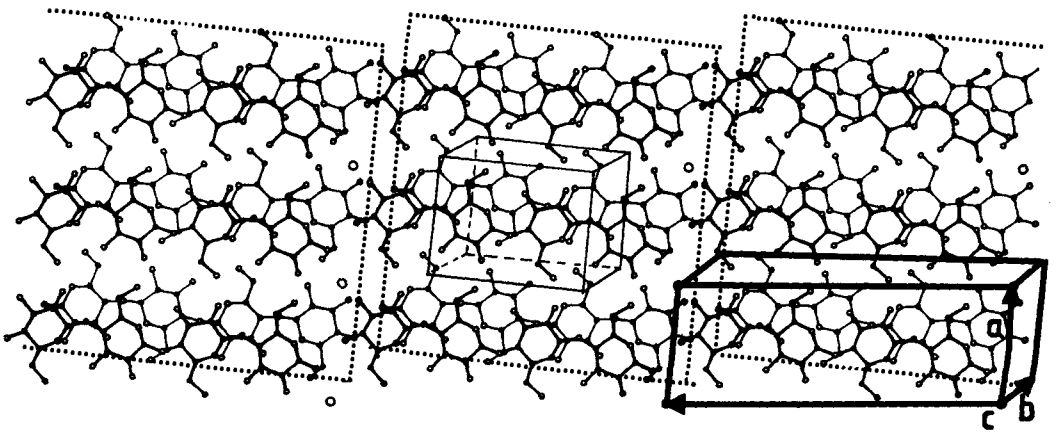

Figure 2. Molecular packing in β-D-cellotetraose hemihydrate viewed in direction b^*, molecules **A** drawn thick, **B** thin. Small circles, C; larger circles, O; largest circles, H_2O. The crystal unit cell is indicated by thick lines (bottom right), the layers of β-D-cellotetraose molecules parallel to a,b by dotted lines, the subcell which is nearly identical to the unit cell of cellulose II (Table 1) by thin lines (center).

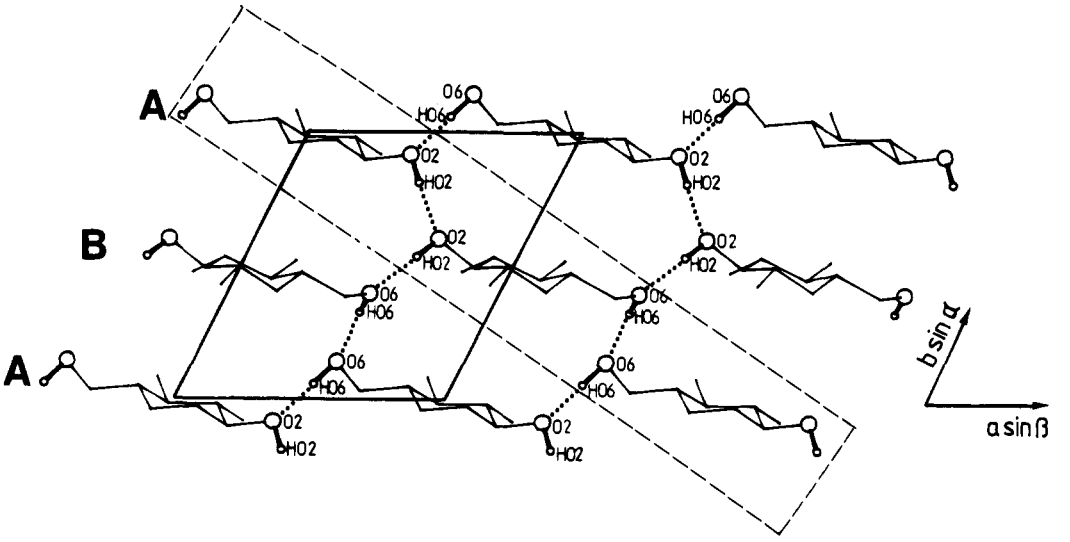

Figure 3. View of the crystal structure along the *c*-axis. Only one glucose residue per molecule is shown. In the direction of the *a*-axis, neighboring molecules **A** and **B** form hydrogen bonded sheets **AA** and **BB**. The hydroxyl gruops O2−H and O6−H are hydrogen bonded. They connect these sheets and form infinite cooperative homodromic chains, O6a−H···O2a−H···O2b−H···O6b−H···O6a−H···. Dashed lines designate the (120) section, which is shown in Figure 5b in more detail.

Table 3. Selection of Average Structural Parameters

(1) Interglucosic Links

	C4−O4 (Å)	O4−C1′ (Å)	C4−O4−C1′ (deg)
A	1.44(1)	1.40(1)	116.3
B	1.45(1)	1.40(1)	116.7

(2) Cremer and Pople puckering parameters[26]

glucoses 1−3	Q_2 (Å)	Θ (deg)	glucose 4	Q_2 (Å)	Θ (deg)
A	0.61(1)	2(1)	**A**	0.50(1)	5(1)
B	0.58(1)	13(1)	**B**	0.62(1)	4(1)

(3) D-Glucopyranose Torsion Angles for Glucoses 1−3[a]

angle	**A**	**B**
C1−C2−C3−C4	−57(1)	−47(1)
C2−C3−C4−C5	58(1)	48(1)
C3−C4−C5−O5	−58(1)	−59(1)
C4−C5−C5−C1	63(1)	69(1)
C5−O5−C1−C2	−63(1)	−67(1)
O5−C1−C2−C3	58(1)	54(1)
O5−C5−C6−O6(χ)	65(1)	54(1)
C4−C5−C6−O6(χ′)	175(1)	170(1)
O5−C1−O4−C4(Φ)	−95(1)	−94(1)
C1−O4−C4−C3(Ψ)	97(1)	88(1)
H1−C1−O4−C4(Φ′)	27(1)	27(1)
C1−O4−C4−H4(Ψ′)	−21(1)	−30(1)

[a] Data for glucose 4 omitted because of end effects.

The reducing end of molecule **B** is in the β-form, but for molecule **A** there is a configurational disorder with 75% β and 25% α, *vide infra*.

All eight D-glucopyranoses are in the typical $^4C_1$ chair conformation and rotated alternately by ~180° about the molecular axis. The C6−O6 bonds are in *gauche,trans* orientation. The cross conformational features of molecules **A** and **B** are comparable, but there are differences in detail. The puckering parameters of the D-glucopyranoses as defined by Cremer and Pople[26] indicate that in both molecules, glucose 4 differs from the other three which we associate with "end" and/or packing effects (Table 3). For residues 1, 2, and 3 the puckering amplitudes Q_2 are comparable (~0.60 Å for molecules **A** and **B**, but the Θ values differ significantly, 2(1)° for **A** and 13(1)° for **B**, Table 3). Since the ideal value for unstrained α-D-glucopyranose is Θ = 2.7°,[27] this indicates that the D-glucopyranoses in **B** are conformationally more strained than those in **A**.

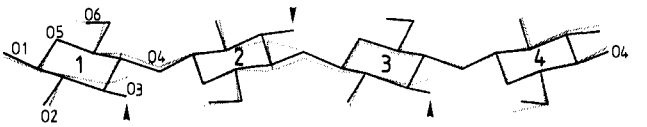

Figure 4. Superposition of β-D-cellotetraose molecules. **A** black, **B** grey, arrows mark the positional differences of the O3H hydroxyl groups.

In detail, the distortions are best described by the endocyclic torsion angles, see Table 3. They indicate again that conformations of the terminal glucopyranoses 4 in molecules **A** and **B** are different from the conformations of the other three residues 1, 2, and 3. In the latter, torsion angles involving C3, i.e., C1−C2−C3−C4 and C2−C3−C4−C5, differ on average by 12° and 10°, respectively, and are smaller in molecule **B** than in molecule **A**. As a consequence, the C3−O3 bonds of D-glucopyranoses 1, 2, and 3 in molecule **B** are in a more "axial" orientation compared with molecule **A** which is clearly illustrated in the superposition of the two β-D-cellotetraose molecules, Figure 4.

The relative orientations of the D-glucopyranose residues in the two molecules are described by torsion angles Φ (O5−C1−O4′−C4′) and Ψ (C1−O4′−C4′−C3′); other definitions involving hydrogen atoms HC1 and HC4 are in Table 3. For molecules **A** and **B**, the average Φ is comparable, −95(1)° for **A** and −94(1)° for **B**. However, Ψ differs by 9°, with average values of 97(1)° for **A** and 88(1)° for **B**. Similarly, the exocyclic torsion angles χ (O5−C5−C6−O6), which are in the *gauche* range, differ significantly by ~11°, with averages of 65(1)° for **A** and 54(1)° for **B**.

Configurational Disorder Due to Mutarotation. If glucopyranose is dissolved in water, mutarotation occurs with an equilibrium concentration of 38% α and 62% β.[28] For β-D-cellotetraose, we confirmed this ratio by ^{1}H-NMR spectroscopy (not shown).

In the crystal structure, molecule **B** occurs only in the β-form. In contrast, the reducing end of molecule **A** shows configurational disorder with the anomeric oxygen O1 25% in α and 75% in β, Figures 1 and 5a. This ratio was initially assessed on the basis of difference electron density and then refined without constraints. We assume that the disorder of only molecule **A** is associated with packing effects, see discussion.

Such heterogeneous crystals of saccharides which contain

Table 4. Hydrogen Bonds of the type O–H···O in the Crystal Structure of β-D-Cellotetraose Hemihydrate

1. Intramolecular Hydrogen Bonds

	O···O (Å)		H···O (Å)		O–Ĥ···O (deg)	
hydrogen bond	A	B	A	B	A	B
O31···O52	2.83	2.87		1.93		161
O31···O62	3.33	3.06		2.47		118
O32···O53	2.84	2.90	2.04		138	
O32···O63	3.31	3.12	2.73		118	
O33···O54	2.85	2.95	1.92		160	
O33···O64	3.67	3.39	3.02		126	
av						
O3···O5′	2.84	2.91				
O3···O6′	3.32[a]	3.09[a]				

2. Intermolecular in the Sheets **AA** and **BB**

O···O (Å)	A	B	O6–H···O2 (Å)	A	O6–Ĥ···O2 (deg)	O2–H···O6 (Å)	B	O2–Ĥ···O6 (deg)
O61···O21	2.76	2.62	O61–H···O21	1.95	139	O21–H···O61	1.80	138
O62···O22	2.68	2.63	O62–H···O22			O22–H···O62		
O63···O23	2.69	2.63	O63–H···O23	1.79	156	O23–H···O63		
O64···O24	2.76	2.73	O64–H···O24	1.79	172	O24–H···O64		
av								
O6···O2	2.72	2.65						

3. Intermolecular in Sheets **AB**

bond	O···O Å	O–H···O	O–Ĥ···O (deg)
O61b–H···O64a	2.74	1.79	163
O62b–H···O63a	2.72		
O63b–H···O62a	2.68		
O64b–H···O61a	2.86	1.91	161
av			
O6b···O6a	2.75		
O21a–H···O24b	2.81		
O22a–H···O23b	2.73	1.77	170
O23a–H···O22b	2.74	1.79	163
O24a–H···O21b	2.77	1.82	163
av			
O2a···O2b	2.76		

[a] The distances O33···O64 are not included in the averaging.

both α and β forms have observed earlier.[30] The ratio α/β may vary depending on the crystallization conditions.

Intramolecular Hydrogen Bonds Are of the Three-Center Type, O3H···O5 (Major) and O6 (Minor). Because the D-glucopyranoses in the β-D-cellotetraose molecules are rotated alternatively by ~180°, a systematic intramolecular hydrogen bond pattern is formed. It is of the three center type[30] and involves $O3_nH$ as donor with major hydrogen bond component to $O5_{n+1}$ and minor component to $O6_{n+1}$, Figure 5. Distances, for molecules **A** and **B** were averaged except for O33···O64, which is wider due to "end" effects, 3.67 Å for **A** and 3.39 Å for **B**. The average $O3_n\cdots O5_{n+1}$;$O3_n\cdots O6_{n+1}$ for **A** are 2.84;3.32 Å, and for **B** 2.91;3.09 Å, Table 4. The average differences between major and minor components are 0.48 Å in **A** but only 0.18 Å in **B**; these data indicate again that molecules **A** and **B** are not structurally identical.

Intramolecular Hydrogen Bonds within and between Layers AA and BB Involve O2H and O6H. In the crystal structure of β-D-cellotetraose hemihydrate, intermolecular O–H···O hydrogen bonds are formed by O2H and O6H hydroxyl groups in different combinations. They define the organization of molecules **A** and **B** in three intersecting layers.

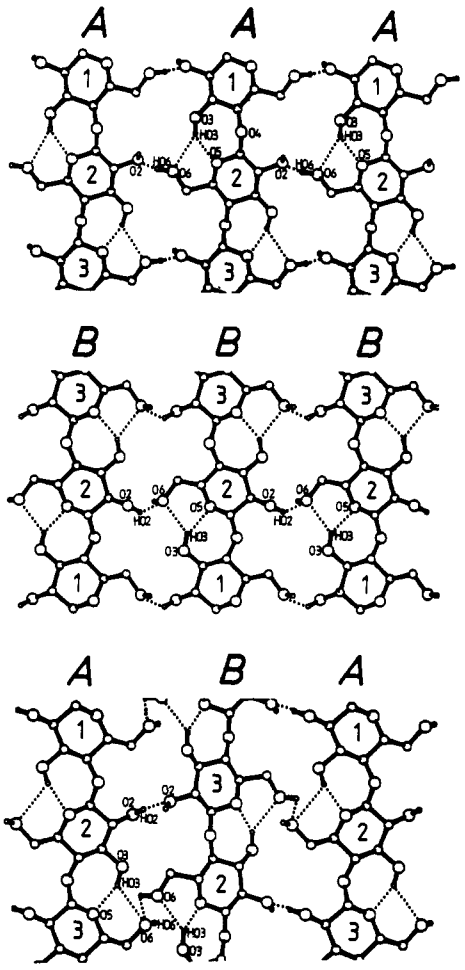

Figure 5. Section of sheets **AA** and **BB**, view is on the *a*,*c*-plane (010). Systematic intra- and intermolecular hydrogen bonds, as derived and extrapolated on the basis of the located H-atoms, are shown with dotted lines. Intermolecular hydrogen bonds are O6–H···O2 in **AA**, but O2–H···O6 in **BB**. The intramolecular three-center hydrogen bonds are more symmetric in layer **BB** than in layer **AA**, see Table 4. Section of sheet **AB** which is parallel to (120), see Figure 3, with systematic hydrogen bonds indicated by dotted lines.

Because molecules **A** are oriented along *c* and **B** along $-c$, translation in *a*-direction gives rise to antiparallel sheets **AA** and **BB** parallel to the *a*,*c* (010) plane, Figure 5, which are displaced by ~2.5 Å along *c*. A third layer **AB** parallel to (120), (Figure 3), is formed by alternating molecules **A** and **B**, Figure 5.

Although only 12 of the 28 OH-hydrogen atoms could be located from difference electron density maps, the following systematic hydrogen bonding scheme emerges: In layer **A**, O6H (x,y,z) donates a hydrogen bond to O2 $(x-1,y,z)$ of an equivalent molecule **A** translated in *a*-direction, with average O···O separation of 2.72 Å, Table 4. A comparable interaction is observed in layer **B**, but now O2H (x,y,z) is the donor and O6 $(x+1,y,z)$ the acceptor, with an average O···O distance of 2.65 Å. Since neighboring molecules **A** and **B** in layer **AB** and are shifted in the *c*-direction by ~2.5 Å and approach each other alternately with their O2H and O6H groups, hydrogen bonds between like hydroxyl groups can form. These are of type O2a–H···O2b and O6b–H···O6a with average O···O distances of 2.76 and 2.75 Å, respectively.

In direction of the *b*-axis, these three different hydrogen bonds form a homodromic chain O6a–H···O2a–H···O2b–H···O6b–H···O6a–H..., Figure 3. The complete course of these homodromic chains is observed only for glucopyranoses 1a and 4b and as fragments for the other glucopyranoses. They suggest

(24) Hubbard, R.; Evans, P. MRC Laboratory of Molecular Biology, Cambridge, UK, 1985.
(25) Zugenmaier, P.; Sarko, A. *Biopolymers* **1976**, *15*, 2121.
(26) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354.
(27) Dowd, M. K.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1994**, *264*, 1.
(28) Dubrunfaut, K. *C.r.* **1846**, *23*, 38.
(29) Jeffrey, G. A. *Acta Crystallogr.* **1990**, *B46*, 89.
(30) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer Verlag: Berlin, 1991.

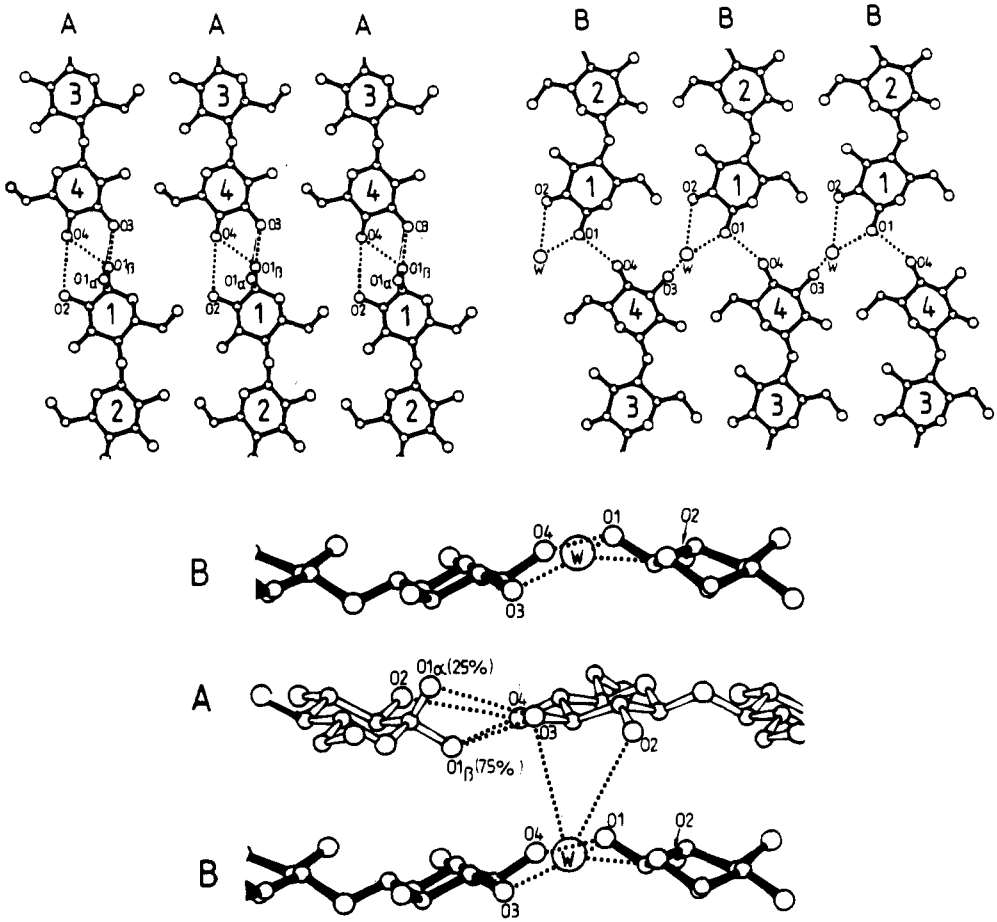

Figure 6. (a) Section of the layer **AA** (left) and **BB** (right) in the crystal structure, view is on 010. The head-to-tail hydrogen bonding pattern is indicated by dotted lines. The water molecule **W** is located in layer **BB**. The hydrogen bond formed by the α-position of O11a···O34a is shown by a broken line. (b) Hydrogen bonding pattern involving the water molecule *W* and the α- and β-positions of O11a. Interactions between adjacent molecules along the *c*-axis are given by dotted lines. Hydrogen atoms are not shown.

that the chains run alternatively in direction +b and −b, i.e., −*b* for glucoses 1a,4b; +*b* for 2a;3b; −*b* for 3a,2b; +*b* for 4a,1b.

Connection of Adjacent Molecules Along the *c*-Axis: the Water Molecule. Along the *c*-axis, adjacent β-D-cellotetraose molecules are linked by hydrogen bonds O11···O44 which may be taken as substitutes for the covalent C1−O4−C4 bond (Figure 6) in the formation of pseudopolymers.

In layer **AA**, molecules **A** which are adjacent in *c*-direction, i.e., the positions (*x*,*y*,*z*) and (*x*,*y*,*z*+*1*), are slightly displaced laterally in direction of *a* so that the disordered O11 positions, O11α(75%) and O11β(25%), at (*x*,*y*,*z*) hydrogen bond simultaneously to O34 of molecule **A** in (*x*,*y*,*z*+*1*), 2.65 and 3.04 Å, respectively; O11β(75%) is in addition engaged in a bond to O44, 2.82 Å. A weaker interaction connects O21(*x*,*y*,*z*) and O44(*x*,*y*,*z*+*1*), 3.33 Å.

For molecules **B**, the lateral displacement in direction of *a* is more severe so that molecules in (*x*,*y*,*z*) are almost halfway between those in (*x*,*y*,*z*+*1*). Hydrogen bonding interactions between these molecules are direct, O11(*x*,*y*,*z*)···O44(*x*,*y*,*z*+*1*), 2.79 Å, and also mediated by the water molecule **W**. The water forms three hydrogen bonds within layer **BB**, two with O11 and O21 of molecule **B** in (*x*,*y*,*z*), 2.94 and 3.05 Å, and one with O34 in (*x*,*y*,*z*+*1*), 2.72 Å, Figure 6a. Besides these hydrogen bonds in layer **BB**, the water molecule is in contact with layer **A** through hydrogen bonds to O24a and O34a, 3.28 and 2.59 Å (Figure 6b). It appears that the different end-to-end contacts in layers **AA** and **BB** are due to the insertion of the water molecule in **BB** which might be associated with the slight rotation of glucose 4 in molecule **B**. This rotation does not occur about the glycosidic link as the Φ, Ψ angles are similar to those between the other D-glucopyranoses (Table 3), but it is due to a different sugar pucker which is more "normal" for this residue compared with the other three D-glucopyranoses in molecule **B**.

C−H···O Hydrogen Bonds and Short H···H Contacts Occur Systematically. C−H···O hydrogen bonds are weaker interactions than O−H···O bonds.[30−33] They occur in many crystal structures of organic and biological molecules[30] and are abundant in the carbohydrates.[33,34] For all C−H groups in β-D-cellotetraose, the H-atom positions are defined by the molecular skeleton and were calculated on the basis of a theoretical C−H distance of 1.08 Å; C−H···O and H···H contacts are indicated in Figure 7 and their geometry is described in Table 5.

Intramolecular short intramolecular $C6_n-H\cdots O2_{n+1}$ interactions occur in glucopyranoses if the C6−O6 bond is in *gauche,trans* orientation. They are observed systematically in the cyclodextrins where they contribute to the stabilization of this conformation.[30,31] In β-D-cellotetraose, they are found for all the glucopyranoses, with H···O distances in the range 2.68−2.77 Å for molecule **A** and 3.04−3.14 Å for molecule **B**. The difference in these distances reflects again the conformational differences discussed above. Because alternating D-glucopyranoses are located ~180° about the glycosidic link, short intramolecular H···H contacts are found for $C4_n-H\cdots H-C1_{n+1}$, 2.03−2.09 Å for molecule **A** and 2.07−2.25 Å for molecule **B**. These distances are all shorter than expected for van der Waals separations, ~2.4 Å. This might be due to the location of H-atoms in "ideal" calculated positions which does not take into account "bending" of C−H bonds to avoid steric clashes.

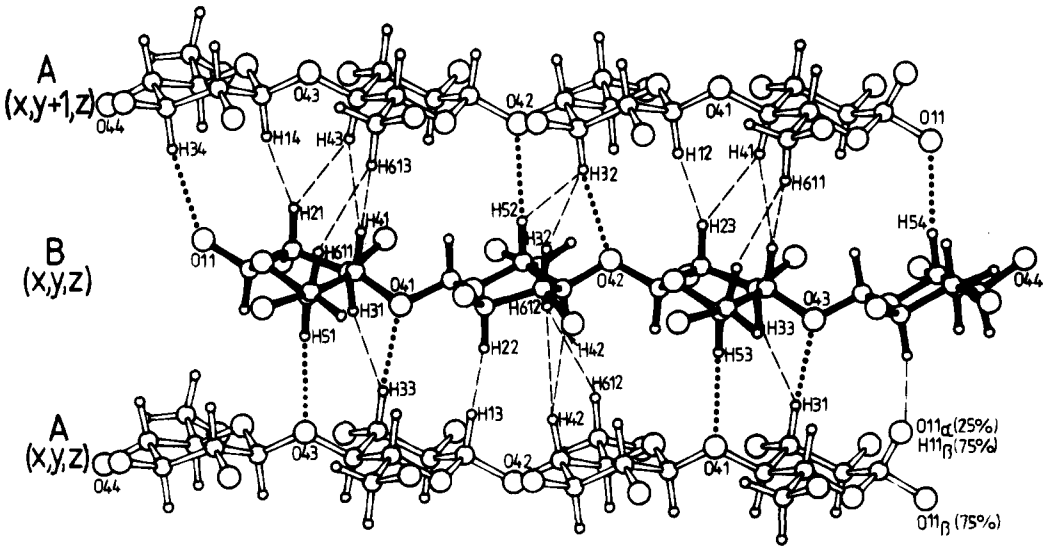

Figure 7. View along *a*-axis, with calculated C−H hydrogens included. Dotted lines: C−H···O contacts, dashed lines: H···H contacts. The intramolecular contacts are not shown for clarity. See also Table 5.

Table 5. Intermolecular C−H···O4 Hydrogen Bonds

contact	H···O (Å)	C···O (Å)	C−Ĥ···O (deg)	symmetry
O41a···H53b	2.37	3.44	172	*(x,y,z)*
O42a···H52b	2.38	3.45	172	*(x,y−1,z)*
O43a···H51b	2.41	3.49	173	*(x,y,z)*
O11a···H54b	2.30	3.38	180	*(x,y−1,z)*
O11b···H34a	2.71	3.73	158	*(x,y+1,z)*
O41b···H33a	2.67	3.70	160	*(x,y,z)*
O42b···H32a	2.64	3.69	163	*(x,y+1,z)*
O43b···H31a	2.58	3.63	164	*(x,y,z)*

Intermolecular within layers **AA** or **BB**, no obvious short C−H···O or H···H contacts are observed. Between the layers, however, they occur systematically, see Figure 7 and Table 5. The acceptor atoms are O4 in both molecules **A** and **B**. They protude from the molecular planes and are able to form contacts with C3−H and C5−H of adjacent layers **AA** and **BB**, respectively. The interactions are C5b−H···O4a, 2.30 to 2.41 Å, and C3a−H···O4b, 2.58 to 2.71 Å, and are accompanied by short H3a···H3b contacts, 2.21 to 2.23 Å, see Figure 7.

The pattern of H···H contacts is more complex; it involves H1, H3, H4 of **A** and H2, H3, H4 of **B**. Because adjacent layers are displaced by 2.5 Å along *c*, H1 and H3 of each glucopyranose **A** interact with a disaccharide unit of **B** (H1a···H2b(*n*) and H3a···H3b(*n*+1)) and of each D-glucopyranose **B**, H2 and H4 interact with a disaccharide of **A** (H2b···H1a(*n*) and H4b···H4a(*n*+1)). These contacts are short and in the range of 2.21 to 2.36 Å. There are additional H···H interactions between 2.4 and 2.6 Å which are indicated in Figure 7. These intermolecular contacts suggest that besides O−H···O hydrogen bonds the more numerous C−H···O and H···H contacts contribute significantly to the stability of the crystal structure. They are of importance in the fine-tuning of the packing arrangement of the flat, lath-shaped molecules.

Comparison with the β-D-Cellotetraose of Chanzy et al. The least squares superposition of the C and O-atoms of the crystal structure of β-D-cellotetraose hemihydrate published by Chanzy and co-workers and of this structure has an RMS-deviation of 0.05 Å for the non-hydrogen atoms. There are several discrepancies in the positions of O−H hydrogen atoms and consequently in hydrogen bonding scheme. This is because Chanzy et al. calculated H-atoms positions with the AFIX routine of SHELX93 whereas the H-atoms in the present study were located from difference Fourier analyses. The disorder of the anomeric hydroxyl group O11a was not observed by Chanzy et al.[19] which might be associated with different crystallization conditions.[29]

The Structure of β-D-Cellotetraose as a Model for Cellulose II

Definition of a Subcell with Dimensions as in Cellulose II. A closer look at the packing scheme (Figure 2) shows that the crystal structure is composed of layers of β-D-cellotetraose molecules which are parallel to the *a,b* plane. If we draw a subcell containing only glucopyranose residues 2 and 3 (to avoid "end" effects, *vide infra*) of molecules **A** and **B**, it has unit cell constants (Table 1) that are practically identical to those of cellulose II. We have shifted the origin of this subcell to satisfy the symmetry conditions of space group $P2_1$ and find (1) that the two glucopyranoses 2 and 3 for both molecules are related by a 2_1 srew axis and (2) that superposition with the structure of cellulose II (omitting O6 atoms because they are in different orientation, *vide infra*) is nearly ideal, with an rms deviation of only 0.17 Å. This suggests that the subcell of β-D-cellotetraose hemihydrate is, in fact, the basic unit cell of cellulose II. Atomic coordinates of the cellulose II model are given in the supporting information.

Modeling of Cellulose II Structure. As described in the experimental section, modeling of the cellulose II crystal structure was carried out in two stages, each consisting of two steps. The first refinement stage was designed to determine the applicability of an essentially unaltered cellotetraose structure as a viable model for cellulose II. In the second stage the model was allowed to relax (minimally) to permit removal of any unreasonable stereochemical features that may have developed in the initial refinement. Two starting models were used: the first had all four O6 positions *gt*—as in the cellotetraose structure—and the second had two O6 of the corner chain *tg* and the two O6 of the center chain *gt*, as in the original cellulose II structure.[6] The total number of refinement runs was thus eight: each of the two starting models was refined in two stages—with fixed and variable residues, respectively—and within each stage the structures were refined in two steps—both with and without added stereochemical constraints. The results of all eight runs are summarized in Table 6. The principal structural and refinement parameters shown in this table should

(31) Steiner, Th.; Saenger, W. *Carbohydr. Res.* **1994**, 259. Addentum **1995**, *1*, 266.
(32) Desiraju, G. R. *Acc. Chem. Res.* **1991**, *24*, 290.
(33) Steiner, Th.; Saenger, W. *J. Am. Chem. Soc.* **1992**, *114*, 10146.
(34) Steiner, Th.; Saenger, W. *J. Am. Chem. Soc.* **1993**, *115*, 4514.

Table 6. Characteristics of Variously Refined, Cellotetraose-Based Cellulose II Models

I. Model: All O(6)*gt*							II. Model: O(6)*tg*+*gt*						
R″	chain rotations[a] (deg)	center chain translation (Å)	Φ′, Ψ′ torsions[b] (deg)	bridge angle (deg)	O(6) torsions[c] (deg)	H bond lengths (Å)	R″	chain rotations[a] (deg)	center chain translation (Å)	Φ′, Ψ′ torsions[b] (deg)	bridge angle (deg)	O(6) torsions[c] (deg)	H bond lengths (Å)
						1. Chains: Fixed Geometry							
						a. Refinement: with X-ray Data Only							
0.183	33.4	−3.036	23.2, −23.0	115.7	72.9	O2−O6 2.26, 2.45, 2.75, 2.76	0.182	32.1	−3.142	23.2, −23.0	115.7	160.5	O2−O6 2.84, 2.54
	123.5		23.8, −20.6	112.4	66.8	O3−O6 2.84, 2.84, 2.79, 2.87		123.2		23.8, −20.6	112.4	156.0	O3−O6 2.67, 2.60, 3.09, 2.84,
			29.7, −33.3	113.2	57.1	O2−O2 2.75, 2.66				29.7, −33.3	113.2	61.0	O3−O6 3.04, 2.88
			28.8, −31.9	118.3	56.2	O6−O6 2.84, 2.53				28.8, −31.9	118.3	41.0	O2−O2 2.59, 2.65
						O3−O5′ 2.81, 2.72, 2.86, 2.81							O3−O5′ 2.81, 2.72, 2.86, 2.81
													O6−O2′ 3.10, 2.86
						b. Refinement: with Stereochemical Constraints							
0.189	33.3	−2.907	23.2, −23.0	115.7	60.9	O2−O6 2.45, 2.53, 2.65, 2.67	0.194	32.4	−3.129	23.2, −23.0	115.7	170.6	O2−O6 2.70, 2.67
	123.5		23.8, −20.6	112.4	62.4	O3−O6 2.72, 2.80, 2.68, 2.77		124.5		23.8, −20.6	112.4	169.5	O3−O6 2.80, 2.74, 2.83, 2.87,
			29.7, −33.3	113.2	59.5	O2−O2 2.80, 2.77				29.7, −33.3	113.2	60.7	O3−O6 3.03, 2.92
			28.8, −31.9	118.3	59.9	O6−O6 2.65, 2.53				28.8, −31.9	118.3	58.1	O2−O2 2.70, 2.69
						O3−O5′ 2.81, 2.72, 2.86, 2.81							O3−O5′ 2.81, 2.72, 2.86, 2.81
													O6−O2′ 2.80, 2.64
						2. Chains: Variable Geometry (Torsions)							
						a. Refinement: with X-ray Data Only							
0.168	33.1	−2.997	35.4, −30.3	115.7	69.3	O2−O6 2.13, 2.62, 2.55, 2.80	0.170	31.5	−3.173	34.3, −27.7	117.8	160.6	O2−O6 2.70, 2.67
	121.1		20.6, −21.2	112.4	68.7	O3−O6 2.82, 2.88, 2.80, 2.77		122.1		20.0, −20.6	112.4	152.1	O3−O6 2.74, 2.62, 2.83, 3.09,
			18.2, −29.6	108.2	54.3	O2−O2 2.84, 2.52				30.8, −32.5	109.5	51.7	O3−O6 2.70, 2.89
			29.6, −28.0	118.3	45.5	O6−O6 3.16, 2.62				28.1, −29.2	118.3	49.9	O2−O2 2.58, 2.50
						O3−O5′ 2.80, 2.71, 2.85, 2.89							O3−O5′ 2.72, 2.67, 2.98, 2.81
													O6−O2′ 3.31, 3.04
						b. Refinement: with Stereochemical Constraints							
0.185	34.6	−2.860	28.1, −26.6	116.0	63.7	O2−O6 2.37, 2.61, 2.65, 2.75	0.187	33.3	−3.003	32.3, −31.1	116.8	172.8	O2−O6 2.67, 2.60
	124.5		22.7, −18.4	112.4	58.9	O3−O6 2.59, 2.81, 2.67, 2.64		123.8		23.8, −18.9	112.4	169.9	O3−O6 2.90, 2.67, 2.75, 2.94,
			30.4, −34.4	111.1	60.6	O2−O2 2.76, 2.65				37.5, −38.8	111.3	58.7	O3−O6 2.89, 2.85
			29.9, −31.7	118.3	57.5	O6−O6 2.76, 2.76				29.3, −33.3	118.3	60.0	O2−O2 2.75, 2.75
						O3−O5′ 2.89, 2.72, 2.85, 2.86							O3−O5′ 2.84, 2.64, 2.91, 2.85
													O6−O2′ 2.87, 2.55

[a] In ref 6 the rotation for the center chain was reported as the difference between the angle given here and 180°. [b] Torsion angle Φ′: H(1)−C(1)−O(1)−C(4′); angle Ψ′: C(1)−O(4′)−C(4′)−H(4′). A different convention for Φ′ and Ψ′ was used in ref 6; according to the current convention these angles for cellulose II are Φ′ = 23.0°, Ψ′ = −24.8°. [c] The O(6) orientations are reported as torsion angle O(5)−C(5)−C(6)−O(6).

be compared with the corresponding data given in Tables 4, 7, and 8 (supporting information) of the original report on cellulose II.[6]

In most ways the results were consistent with expectations based on the cellotetraose structure, although there were some surprises. First, in all refinements the main skeletal features and placements of the chains in the unit cell were closely grouped, regardless of the refinement protocol or the starting model. All of these characteristics were also reasonably close to those of the original cellulose II structure. For example, the chain rotations differed less than ±2° among all models, on the average, and were within 4° to 7° in comparison with cellulose II. The vertical offset of the center chain relative to the corner chain was virtually the same in all cases, including the original cellulose II structure. Likewise, reasonable values for the glycosidic bond angles ("bridge angles") were evident, as were the O6 rotational positions and the Φ, Ψ torsion angles. The O6 rotations were very close to true *gt* (and *tg*, where applicable) positions; the Φ, Ψ angles varied within a range of approximately 19° for Φ and approximately 13° for Ψ. The corresponding values for the original cellulose II structure also fell within the latter ranges. The hydrogen bond lengths were surprisingly reasonable in all models, with only very few showing somewhat short O···O distances. No excessively short nonbonded contacts developed in any of the models. The changes that occurred in models upon the addition of stereochemical constraints and/or letting the structures relax were small relative to the fixed-residue models refined against X-ray data only. Taken together, these observations show that the cellotetraose structure can be an excellent model for cellulose II.

More surprisingly, however, a distinction between the two main competing models—the all-*gt* O6 model of cellotetraose and the mixed *tg*+*gt* O6 model for the original cellulose II structure—did not develop during this modeling. As shown in Table 6 the *R*″-factors were virtually identical for both models when compared within the same refinement protocol. On this basis, a single most probable model could not be determined. Other possible criteria—such as bridge angles, the Φ, Ψ torsions, O6 positions and short contacts—also did not clearly point to one single model. Only the hydrogen bonds favor—albeit slightly—the *tg*+*gt* model in that a short O2···O6 hydrogen bond found in the all-*gt* model disappeared in the *tg*+*gt* model. However, in the absence of any other clear-cut preferences it would be difficult to pick a favored model based on these refinements.

The most likely reason for the lack of distinction between the two models probably stems from the differences in the refining procedures employed here and during the analysis of the original cellulose II structure. As the data in Table 6 illustrate, significant differences between the stereochemical features of both adjacent glucopyranose residues along a chain and those of the two chains of the unit cell were found in these refinements. In contrast, it should be recalled that in the original cellulose II structure the main conformations of all four residues of the unit cell remained identical, within the assumed $P2_1$ symmetry.

Discussion

With the crystal structure of β-D-cellotetraose, we have information on molecular conformation, intra- and intermolecular interactions, and the packing arrangement of the molecules in the crystal lattice. The two molecules **A** and **B** are comparable in their overall structure which resembles a lath, with alternating D-glucopyranoses rotated ~180°. In detail, the conformations of the D-glucopyranoses in molecule **A** are "normal" and those in molecule **B** are strained. These structural differences are not only observed with the Cremer and Pople parameter Θ but also with the glycosidic torsion angles Φ and Ψ, with the torsion angles χ(O5−C5−C6−O6) and with the intramolecular three-center hydrogen bonds O3−H···O5′/O6′.

Why are the molecules different?

Molecules **A** are in standard conformation and arranged in layers **AA** parallel to the crystallographic *a*,*c*-plane. Between these layers **AA**, layers **BB** are sandwiched in antiparallel orientation and shifted ~2.5 Å in the direction of the *c*-axis. It appears that a snug fitting of layers **AA** and **BB** is only possible if the molecules **BB** adopt a slightly unfavorable conformation to permit the tight interlayer contacts which are not only through "soft" O−H···O and C−H···O hydrogen bonds but also through "hard" H···H contacts. Since these occur systematically between the D-glucopyranose units, we infer a kind of zipper-like "molecular complementarity". This complementarity appears to be possible only if one molecule (**A**) is "normal" and the other (**B**) is "strained" rather than to have both molecules somewhat strained and otherwise identical in their conformation, which would probably interfere with the tight packing in the crystal lattice. The tight fit between the two molecules is also indicated by the relatively high density of the crystals (D_x = 1.56 g/cm^3), whereas crystals of the comparable cyclodextrins have a typical D_x of around 1.45 g/cm^3. Since the molecules in cellulose II are packed without "end" effects, the density is even somewhat higher, 1.61 to 1.63 g/cm^3 (Table 1).

In Table 1, we have described a subcell of the β-D-cellotetraose crystal structure for the central two D-glucopyranoses 2 and 3 of molecules **A** and **B**. Since the cell dimensions and space group symmetry of the subcell are identical to those of cellulose II, we conclude that the subcell in fact reflects the crystal structure of cellulose II. This is in agreement with X-ray powder diffraction[15] which indicated comparable unit cell and intensity distribution and also with ^{13}C CPMAS NMR and polarized infrared spectra which suggested only one conformation for the C6−O6 bond, as observed in this crystal structure. We have constructed a new model of cellulose II based on the subcell of β-D-cellotetraose hemihydrate. The new model agrees with the spectroscopic data and shows all the intra- and intermolecular contacts described for β-D-cellotetraose.

There are several discrepancies with the published X-ray fiber structures of cellulose II. In agreement with the new model, the two cellulose II molecules in the unit cell are arranged antiparallel and shifted ~2.5 Å relative to each other along the *c*-axis. In contrast, however, the two molecules have identical conformation; the D-glucopyranoses are in $^4C_1$ conformation with nearly the same Cremer−Pople parameters Q 0.58 Å and θ 4°, and the orientations about the glycosidic bonds are given by Φ, Ψ angles of 94°, −97°. The most salient difference in comparison with the crystal structure of β-D-cellotetraose and the new model of cellulose II is that the orientations of the C6−O6 bonds differ, being *gauche*,*trans* in one and *trans*, *gauche* in the other molecule. This implies significant differences in the intermolecular hydrogen bonding scheme. The only interactions that are consistent in the old and new models of cellulose II are intramolecular O3···O5′,O6′ and intermolecular O2a···O6a and O2a···O2b, all others are different.

The refinement of the new model with the original cellulose II X-ray intensities indeed produced a new cellulose II structure in which all O6 are in *gt* orientations. This structure is marked by close similarities both with the cellotetraose structure—including the latter's hydrogen bonding network—and the main features of the original cellulose II structure. In terms of the agreement with X-ray data, i.e., the *R*″-factors, the new structure and the O6 *tg*+*gt* structure are virtually identical. Thus, a single

most probable cellulose II structure cannot be picked based on this criterion alone. Other criteria, including NMR findings, do, however, favor the all-*gt* model and on this basis, the latter model may now be considered with some legitimacy as a favored structure for cellulose II. (The atomic coordinates of this structure have been deposited.) It is likely, though, given the kinds of disorder that are common in partially crystalline fibrous morphologies such as cellulose, that both all-*gt* and *tg*+*gt* O6 substructures may coexist in fibers of cellulose II to varying extents, with all-*gt* form predominating as shown by NMR and infrared spectroscopy. After all, both may be equally likely to form based on such energy-minimizing features as the stability of the structure and the maximizing of hydrogen bonding.

Acknowledgment. The authors thank Dietmar Leitner for measuring ^{1}H NMR spectra on β-D-cellotetraose solutions and Claus Sandmann for carrying out energy minimization and molecular modeling calculations on the β-D-cellotetraose molecules and on the structural models for the polymer cellulose II. This work was supported by Deutsche Forschungsgemeinschaft (Sa 196/12-4), by Bundesministerium für Forschung und Technologie (03-SA3FUB), and by Fonds der Chemischen Industrie.

Supporting Information Available: Tables of fractional atomic coordinates of β-D-cellotetraose hemihydrate, fractional atomic coordinates for the new model of cellulose II, anisotropic displacement parameters, bond distances and bond angles, selection of D-glucopyranose torsion angles, and D-glucopyranose exozyclic torsion angles (13 pages); table of F_o-F_c (37 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA951795Q