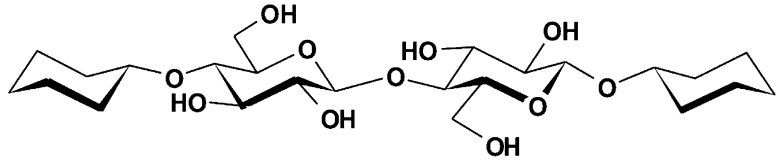
Article

van der Waals versus Hydrogen-Bonding Forces in a Crystalline Analog of Cellotetraose: Cyclohexyl 4#-*O*-Cyclohexyl #-d-Cellobioside Cyclohexane Solvate

Yuko Yoneda, Kurt Mereiter, Christian Jaeger, Lothar Brecker, Paul Kosma, Thomas Rosenau, and Alfred French

J. Am. Chem. Soc., **2008**, 130 (49), 16678-16690 • DOI: 10.1021/ja805147t • Publication Date (Web): 14 November 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES

Published on Web 11/14/2008

van der Waals versus Hydrogen-Bonding Forces in a Crystalline Analog of Cellotetraose: Cyclohexyl 4′-*O*-Cyclohexyl β-D-Cellobioside Cyclohexane Solvate

Yuko Yoneda,[†] Kurt Mereiter,[‡] Christian Jaeger,[§] Lothar Brecker,[⊥] Paul Kosma,[†] Thomas Rosenau,*[,†] and Alfred French*[,¶]

Department of Chemistry, University of Natural Resources and Applied Life Sciences, Vienna (BOKU) Muthgasse 18, A-1190 Vienna, Austria, Faculty of Chemistry, Vienna University of Technology, Getreidemarkt 9/164 SC, A-1060 Vienna, Austria, BAM Federal Institute for Materials Research and Testing, Richard-Willstaetter-Str. 11, D-12489 Berlin, Germany, University of Vienna, Institute of Organic Chemistry, Währinger Str. 38, A-1090 Vienna, Austria, and Southern Regional Research Center, U.S. Department of Agriculture, 1100 Robert E. Lee Boulevard, New Orleans, Louisiana 70124

Received July 3, 2008; E-mail: thomas.rosenau@boku.ac.at; al.french@ars.usda.gov

Abstract: Hydrogen bonding is important in cellulosic and other carbohydrate structures, but the role of interactions between nonpolar groups is less understood. Therefore, we synthesized cyclohexyl 4′-*O*-cyclohexyl β-D-cellobioside (**8**), a molecule that has two glucose rings and two nonpolar cyclohexyl rings. Key to attaching the 4′-*O*-cyclohexyl group was making the 4′-*O*,6′-*O*-cyclohexylidene ketal. After peracetylation, the cyclohexylidene ketal ring was opened regioselectively, providing 65% of **8** after final deacetylation. Comparison of the crystal structure of **8**, as the cyclohexane solvate, with those of cellulose and its fragments, especially cellotetraose with four glucose rings, revealed extensive effects from the cyclohexyl groups. Three conformationally unique molecules (A, B, and C) are in the triclinic unit cell of **8**, along with two solvent cyclohexanes. When viewed down the crystal's *a*-axis, the array of C, A, and B looks like the letter N, with A inclined so that its cyclohexyl groups can stack with those of the reducing ends of the B and C molecules. The lower left and upper right points of the N are stacks of cyclohexyl rings on the nonreducing ends of B and C, interspersed with solvent cyclohexanes. Whereas cellotetraose has antiparallel (up−down) packing, A and B in **8** are oriented "down" in the unit cell while C is "up". "Down-down-up" (or, alternatively, "up-up-down") packing is rare for carbohydrates. Other unusual details include O6 in all three staggered orientations: one is *tg,* two are *gg*, and three are *gt*, confirmed with CP/MAS ^{13}C NMR. The *tg* O6 donates a proton to an intramolecular hydrogen bond to O2′, opposite to the major schemes in native cellulose I. A similar but novel O6B−H···O2′B hydrogen bond is based on a slightly distorted *gg* orientation. The hydrogen bonds between parallel molecules are unique, with linkages between O2′A and O2′B, O3′A and O3′B, and O6A and O6B. Other details, such as the bifurcated O3···O5′ and ···O6′ hydrogen bonds are similar to those of other cellulosic structures. C−H···O hydrogen bonds are extensive along the [110] line of quarter-staggering. The unusual features described here expand the range of structural motifs to be considered for as-yet undetermined cellulose structures.

Introduction

Cellulose, the β-1,4-linked polymer of glucose, is the main constituent of plant cell walls. As the predominant component of cotton, rayon, and lyocell fibers, it is the major renewable molecule in textiles. Likewise, the paper industry is based on cellulose mostly from wood pulp. Consumption of cellulose would expand greatly if its development as a source of biofuels succeeds. Details of the structures of these materials at all levels are needed for improvements in cellulose utilization. The past decade has seen landmark reports of the crystal structures for several important forms of cellulose: Iα,[1] Iβ,[2] mercerized II,[3,4] and III_I.[5] Despite these successes, attempts to better understand cellulose structure continue. Details for other crystalline forms, e.g., cellulose III_{II}, IV_I,[6] and IV_{II} are unknown. Also, the understanding of the transitions among the forms is still in its

[†] University of Natural Resources and Applied Life Sciences Vienna.
[‡] Vienna University of Technology.
[§] BAM Federal Institute for Materials Research and Testing.
[⊥] University of Vienna.
[¶] U.S. Department of Agriculture.

(1) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300–14306.
(2) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074–9082.
(3) Langan, P.; Nishiyama, Y.; Chanzy, H. *J. Am. Chem. Soc.* **1999**, *121*, 9940–9946.
(4) Langan, P.; Nishiyama, Y.; Chanzy, H. *Biomacromolecules* **2001**, *2*, 410–416.
(5) Wada, M.; Chanzy, H.; Nishiyama, Y.; Langan, P. *Macromolecules* **2004**, *37*, 8548–8555.
(6) Wada, M.; Heux, L.; Sugiyama, J. *Biomacromolecules* **2004**, *5*, 1385–1391.

10.1021/ja805147t CCC: $40.75

early stages,[7,8] and relatively little is known about the noncrystalline regions where many reactions and other interactions are thought to occur. A particular issue is the role of polar and nonpolar groups in crystal formation[9] that immediately follows biosynthesis. The crystals of cellulose are thought to make enzymatic degradation difficult, so understanding the forces that hold the crystals together may help in finding ways to reduce "biomass recalcitrance"[10] during processing to make biofuel.

Another approach to understanding cellulose structure is to study the structures of related small molecules. Cellobiose, the dimeric fragment of cellulose, was the second disaccharide structure to be determined by X-ray diffraction.[11,12] Now, results from studies of numerous related small molecules can be combined to synthesize additional information about the possibilities for cellulose in the as-yet unknown structures.[13]

Previous studies of analogs usually depended on the availability of suitable crystalline samples. The present paper reports the second molecule that we synthesized expressly to provide a cellulose analog for structural study. It follows work on the first small relative of cellulose to crystallize in two different forms, 1,4′-dimethyl cellobioside,[14] as well as efforts of Vasella's group.[15−17] Here we report the synthesis and structural study of cyclohexyl 4′-*O*-cyclohexyl-β-D-glucopyranosyl-(1→4)-β-D-glucopyranoside (**8**) or, alternatively, 1,4′-dicyclohexyl cellobioside. (We have used **8** to represent both the molecule itself and the crystalline cyclohexane solvate form.) This molecule was targeted because of its similarity to cellotetraose,[18,19] with the cyclohexyl rings replacing the two terminal glucose rings of the tetraose. It can also be thought of as a derivative of cellobiose to which two nonpolar caps are added. Because cyclohexanes in **8** replace the O1 and O4′ hydroxyl hydrogen atoms of cellobiose, the latter will not be involved in hydrogen bond donation, a situation more like that in midchain cellulose.

Added cyclohexyl groups, along with the existing C−H groups of the glucose rings, should increase the influence of nonpolar group attractions on the overall structure. Looking ahead to the results, this "self-assembly of sub-nano particles", also called crystallization, has resulted in a molecular array that is influenced to a surprising extent by the cyclohexyl moieties. O−H···O and C−H···O hydrogen bonding are also extensive, but the crystal packing, O6 orientation, and hydrogen bonding scheme all have unusual features in response to the presence of cyclohexane.

Experimental Section

General. Commercial chemicals were of the highest grade available and were used without further purification. Reagent-grade solvents were used for all extractions and workup procedures. Distilled water was used for all aqueous extractions and for all aqueous solutions. *n*-Hexane, diethyl ether, ethyl acetate, and petroleum ether used in chromatography were distilled before use. All reactions involving nonaqueous conditions were conducted in oven-dried (140 °C, overnight) or flame-dried glassware under an argon or nitrogen atmosphere. Thin layer chromatography (TLC) was performed using Merck silica gel 60 F254 precoated plates. Flash chromatography was performed using Baker silica gel (40 μm particle size). The use of brine refers to saturated aqueous NaCl.

Melting points, determined on a Kofler-type micro hot stage with Reichert-Biovar microscope, are uncorrected. Elemental analyses were performed at the Microanalytical Laboratory of the Institute of Physical Chemistry at the University of Vienna.

Solution NMR. ^{1}H NMR spectra were recorded at 600.13, 400.13 or 300.13 MHz for ^{1}H and at 150.86, 100.03 or 75.47 MHz for ^{13}C NMR in $CDCl_3$ as the solvent if not otherwise stated. Chemical shifts, relative to TMS as internal standard, are given in δ values, and coupling constants in Hz. ^{13}C peaks were assigned by means of APT, HMQC, and HMBC spectra.

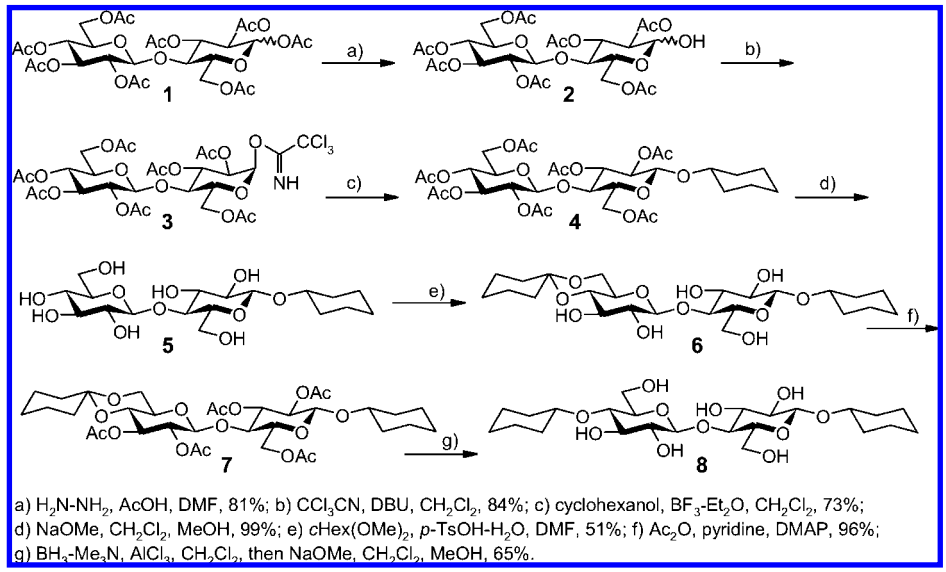
Solid-State NMR. ^{13}C solid-state NMR measurements were performed using an AVANCE 600 (14.1 T, Bruker Biospin GmbH, Rheinstetten, Germany). Magic angle sample spinning (MAS) was applied using 4 mm zirconia rotors at a rotation frequency of 15.5 kHz. ^{13}C NMR spectra were recorded using cross-polarization[20] with magic angle sample spinning (CPMAS) at a Larmor frequency of 150.9 MHz. The ^{1}H 90° pulse length was 2.5 μs, and a CP contact time of 2 ms was used with repetition times of 3 s. During the contact time, the carbon spin lock field strength was held constant, while the proton spin-lock field was ramped linearly (ramped-CP)[21] down to 50% of the initial value. Proton decoupling was carried out with a 15° two pulse phase modulation (TPPM) sequence.[22] ^{13}C chemical shifts (δ) have been calibrated using the glycine COOH signal set to $\delta = 176.4$ ppm as a secondary standard.

Crystallography. X-ray crystallography utilized a Bruker Smart APEX CCD 3-axis diffractometer with sealed X-ray tube, Mo Kα radiation, a graphite monochromator, and a crystal-to-detector distance of 50 mm. 0.3° ω-scan frames covered a complete sphere of the reciprocal space (5 × 600 frames). After data integration, corrections for absorption and $\lambda/2$-effects were applied to the data using program SADABS.[23] Structure solution was by direct methods using SHELXS97,[24] and the structure was refined using full-matrix least-squares on F^2 (SHELXL97).[25] Carbon-bonded hydrogen atoms were inserted in idealized positions, oxygen-bonded hydrogen atoms were at first located in a difference Fourier map and were then refined with AFIX 147 constraints. The orientation disorder of two out of the six independent cyclohexyl termini was taken into account with stabilizing 1-2 and 1-3 distance restraints. Moreover, a DELU 0.003 restraint was applied to the U_{ij} of all non-hydrogen atoms.

Chemical Synthesis. All products were purified to homogeneity by TLC/GCMS analysis. All given yields refer to isolated, pure

(7) Shibazaki, H.; Kuga, S.; Okano, T. *Cellulose* **1997**, *4*, 75–87.
(8) Dinand, E.; Vignon, M.; Chanzy, H.; Heux, L. *Cellulose* **2002**, *9*, 7–18.
(9) Cousins, S. K.; Brown, R. M. *Polymer* **1995**, *36*, 3885–3888.
(10) Himmel, M. E.; Ding, S.-Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. *Science* **2007**, *315*, 804–807.
(11) Jacobson, R. A.; Wunderlich, J. A.; Lipscomb, W. N. *Acta Crystallogr.* **1961**, *14*, 598–607.
(12) Chu, S. S. C.; Jeffrey, G. A. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1968**, *24*, 830–838.
(13) French, A. D.; Johnson, G. P. *Cellulose* **2004**, *11*, 5–22.
(14) Rencurosi, A.; Röhrling, J.; Pauli, J.; Potthast, A.; Jäger, C.; Pérez, S.; Kosma, P.; Imberty, A. *Angew. Chem., Int. Ed.* **2002**, *41*, 4277–4281.
(15) Murty, K.V.S.N.; Xie, T.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **2006**, *89*, 675–730.
(16) Murty, K. V. S. N.; Vasella, A. *Helv. Chim. Acta* **2001**, *84*, 939–963.
(17) Bernet, B.; Xu, J.; Vasella, A. *Helv. Chim. Acta* **2000**, *83*, 2072–2114.
(18) Gessler, K.; Krauss, N.; Steiner, T.; Betzel, C.; Sandmann, C.; Saenger, W. *Science* **1994**, *266*, 1027–1029.
(19) Gessler, K.; Krauss, N.; Steiner, T.; Betzel, C.; Sarko, A.; Saenger, W. *J. Am. Chem. Soc.* **1995**, *117*, 11397–11406.
(20) Hartmann, S. R.; Hahn, E. L. *Phys. Rev.* **1962**, *128*, 2042–2053.
(21) Metz, G.; Wu, X. L.; Smith, S. O. *J. Magn. Reson., Ser. A* **1994**, *110*, 219–227.
(22) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. *J. Chem. Phys.* **1995**, *103*, 6951–6958.
(23) Sheldrick, G. M. *SADABS*, Program for Empirical Absorption Correction of Area Detector Data; University of Gottingen: Germany, 1996.
(24) Sheldrick, G. M. *SHELXS97*, Program for the Solution of Crystal Structures; University of Gottingen: Germany, 1996.
(25) Sheldrick, G. M. *SHELXL97*, Program for Crystal Structure Refinement; University of Gottingen: Germany, 1997.

Scheme 1. Synthesis Summary. Seven-Step Sequence Leading to Cyclohexyl 4′-*O*-Cyclohexyl-β-D-glucopyranosyl-(1→4)-β-D-glucopyranoside (Cyclohexyl 4′-*O*-Cyclohexyl β-D-Cellobioside, **8**) in 16% Overall Yield; Note the Novel Regioselective Opening of the 4′-*O*,6′-*O*-Cyclohexylidene Ketal



products. Seven intermediate compounds were synthesized (Scheme 1), of which the first three compounds (**1**−**3**) are standard in oligosaccharide chemistry. Therefore, their preparation and confirmatory data are not listed in the following.

Cyclohexyl 2′,3′,4′,6′-Tetra-*O*-acetyl-β-D-glucopyranosyl-(1→4)-2,3,6-tri-*O*-acetyl-β-D-glucopyranoside (4). To a solution of 2′,3′,4′,6′-tetra-*O*-acetyl-β-D-glucopyranosyl-(1→4)-2,3,6-tri-*O*-acetyl- α-D-glucopyranosyl trichloroacetimidate (**3**, 4.947 g, 6.335 mmol) in anhydrous CH_2Cl_2 in the presence of molecular sieve (4Å, powder) was added cyclohexanol (3.3 mL, 32 mmol) at room temperature. $BF_3 \cdot Et_2O$ (80 μL, 0.633 mmol) was added at −18 °C, and the solution was stirred for 1 h. Stirring was continued for 3 h, and a new portion of $BF_3 \cdot Et_2O$ (80 μL, 0.633 mmol) was added every hour at −18 °C. The reaction mixture was neutralized with trimethylamine (500 μL) at −18 °C, and the molecular sieve was filtered off. The filtrate was evaporated, diluted with EtOAc (200 mL), washed with 0.01 N aqueous HCl, neutralized with saturated aqueous $NaHCO_3$, washed with water and brine, and dried over $MgSO_4$. After evaporation to dryness, the residue was purified by flash column chromatography (EtOAc/toluene, v/v = 1:1) to give a colorless solid, which was crystallized from CH_2Cl_2/EtOH as colorless plates (1.959 g, 43%). Additional material was contained in the filtrate (1.356 g, 30%). R_f = 0.63 (EtOAc/toluene, v/v = 2:1); mp = 207−208 °C; $[\alpha]^{20}_D$ = −25.5 (*c* 1.00, $CHCl_3$). ^{1}H NMR: δ 1.18−1.87 (m, 10H, cyclohexyl), 2.12, 2.09, 2.03, 2.02, 2.014, 2.008, 1.98 (7*s*, 7 × 3H, CH_3 in acetate), 3.57 (ddd, $J_{4,5}$ = 9.9 Hz, $J_{5,6a}$ = 5.1 Hz, $J_{5,6b}$ = 2.0 Hz, 1H, H-5), 3.60 (m, 1H, O-CH in cyclohexyl), 3.66 (ddd, $J_{4',5'}$ = 9.6 Hz, $J_{5',6'a}$ = 2.3 Hz, $J_{5',6'b}$ = 4.5 Hz, 1H, H-5′), 3.75 (dd, $J_{3,4}$ = 9.1 Hz, $J_{4,5}$ = 9.9 Hz, 1H, H-4), 4.04 (dd, $J_{5',6'a}$ = 2.3 Hz, $J_{6'a,6'b}$ = 12.5 Hz, 1H, H-6′a), 4.09 (dd, $J_{5,6a}$ = 5.1 Hz, $J_{6a,6b}$ = 11.9 Hz, 1H, H-6a), 4.36 (dd, $J_{5',6'b}$ = 4.5 Hz, $J_{6'a,6'b}$ = 12.5 Hz, 1H, H-6′b), 4.48 (dd, $J_{5,6b}$ = 2.0 Hz, $J_{6a,6b}$ = 11.9 Hz, 1H, H-6b), 4.50 (d, $J_{1',2'}$ = 8.1 Hz, 1H, H-1′), 4.53 (d, $J_{1,2}$ = 8.1 Hz, 1H, H-1), 4.87 (dd, $J_{1,2}$ = 8.1 Hz, $J_{2,3}$ = 9.7 Hz, 1H, H-2), 4.92 (dd, $J_{1',2'}$ = 8.1 Hz, $J_{2',3'}$ = 9.1 Hz, 1H, H-2′), 5.05 (dd, $J_{3',4'}$ = 9.4 Hz, $J_{4',5'}$ = 9.6 Hz, 1H, H-4′), 5.14 (dd, $J_{2',3'}$ = 9.1 Hz, $J_{3',4'}$ = 9.4 Hz, 1H, H-3′), 5.17 (dd, $J_{2,3}$ = 9.7 Hz, $J_{3,4}$ = 9.1 Hz, 1H, H-3). ^{13}C NMR: δ 20.61, 20.61, 20.61, 20.66, 20.71, 20.76, 20.92 (7 × CH_3 in acetate), 23.62, 36.75, 25.52, 31.67, 33.25 (5 × CH_2 in cycohexyl), 61.61 (C-6′), 62.03 (C-6), 67.85 (C-4′), 71.64 (C-2′), 71.76 (C-2), 71.93 (C-5′), 72.54 (C-5), 72.61 (C-3), 72.94 (C-3′), 76.63 (C-4), 78.02 (O-CH in cyclohexyl), 99.18 (C-1), 100.68 (C-1′), 168.84, 169.10, 169.31, 169.66, 170.01, 170.12, 170.29 (7 × CO in acetate). Anal. Calcd for $C_{32}H_{46}O_{18}$: C, 53.48; H, 6.45. Found: C, 53.37; H, 6.39.

Cyclohexyl β-D-Glucopyranosyl-(1→4)-β-D-glucopyranoside (5). To a solution of **4** (782 mg, 1.09 mmol) in CH_2Cl_2/MeOH (v/v = 1:1, 21 mL) was added NaOMe in MeOH (0.1 M, 1.1 mL) at 0 °C. The solution was stirred at room temperature for 4 h, and then NaOMe was neutralized with acidic ion-exchange resin DOWEX 50W X8. The resin was filtered off, and the filtrate was evaporated. The residue was purified by flash column chromatography (MeOH/CH_2Cl_2, v/v = 1:4) to give compound **5** (460 mg, 99%) as colorless glass-like solid. R_f = 0.45 (MeOH/CH_2Cl_2, v/v = 3:7); mp = 209−211 °C; $[\alpha]^{20}_D$ = −22.5 (*c* 1.00, water); ^{1}H NMR (D_2O): δ 1.19−2.03 (m, 10H, cyclohexyl), 3.31 (dd, $J_{1,2}$ = 8.0 Hz, $J_{2,3}$ = 9.2 Hz, 1H, H-2), 3.36 (dd, $J_{1',2'}$ = 7.9 Hz, $J_{2',3'}$ = 9.0 Hz, 1H, H-2′), 3.46 (dd, $J_{3',4'}$ = 8.9 Hz, $J_{4',5'}$ = 9.7 Hz, 1H, H-4′), 3.54 (ddd, $J_{4',5'}$ = 9.7 Hz, $J_{5',6'a}$ = 5.6 Hz, $J_{5',6'b}$ = 2.1 Hz, 1H, H-5′), 3.56 (br, dd, $J_{2',3'}$ = 9.0 Hz, $J_{3',4'}$ = 8.9 Hz, 1H, H-3′), 3.60−3.69 (m, 3H, H-5, H-3, H-4), 3.77 (dd, $J_{5',6'a}$ = 5.6 Hz, $J_{6'a,6'b}$ = 12.0 Hz, 1H, H-6′a), 3.82 (m, 1H, O-CH in cyclohexyl), 3.84 (dd, $J_{5,6a}$ = 4.6 Hz, $J_{6a,6b}$ = 12.2 Hz, 1H, H-6a), 3.96 (dd, $J_{5',6'b}$ = 2.1 Hz, $J_{6'a,6'b}$ = 12.0 Hz, 1H, H-6′b), 4.00 (dd, $J_{5,6b}$ = 2.1 Hz, $J_{6a,6b}$ = 12.2 Hz, 1H, H-6b), 4.55 (d, $J_{1',2'}$ = 7.9 Hz, 1H, H-1′), 4.63 (d, $J_{1,2}$ = 8.0 Hz, 1H, H-1); ^{13}C NMR (D_2O): δ 24.13, 24.31, 25.50, 31.92, 33.47 (5 × CH_2 in cyclohexyl), 60.66 (C-6), 61.09 (C-6′), 69.95 (C-4′), 73.39 (C-2), 73.64 (C-2′), 74.89 (C-3), 75.11 (C-5), 76.01 (C-3′), 76.42 (C-5′), 79.26 (O-CH in cyclohexyl), 79.31 (C-4), 100.50 (C-1), 102.94 (C-1′). Anal. Calcd for $C_{18}H_{32}O_{11}$: C, 50.94; H, 7.60. Found: C, 50.90; H, 7.55.

Cyclohexyl 4′,6′-*O*-Cyclohexylidene-β-D-glucopyranosyl-(1→4)-β-D-glucopyranoside (6). To a solution of **5** (425 mg, 1.00 mmol) in DMF (10 mL) were added cyclohexanone dimethyl ketal (305 μL, 2.00 mmol) and *p*-toluenesulfonic acid monohydrate (10 mg, 0.050 mmol) at room temperature. The solution was stirred under decreased pressure (4 kPa) at 30 °C for 4 h. Another portion of cyclohexanone dimethyl ketal (152 μL, 1.00 mmol) and *p*-toluenesulfonic acid monohydrate (10 mg, 0.050 mmol) was added, and stirring was continued for another 4 h. After addition of a third portion of cyclohexanone dimethyl ketal (152 μL, 1.00 mmol) and *p*-toluenesulfonic acid monohydrate (10 mg, 0.050 mmol) and stirring of the solution for 1 h, the reaction mixture was neutralized with solid $NaHCO_3$ by stirring overnight. Solids were removed by filtration, and the filtrate was coevaporated with toluene to dryness. The residue was purified by flash column chromatography (MeOH/

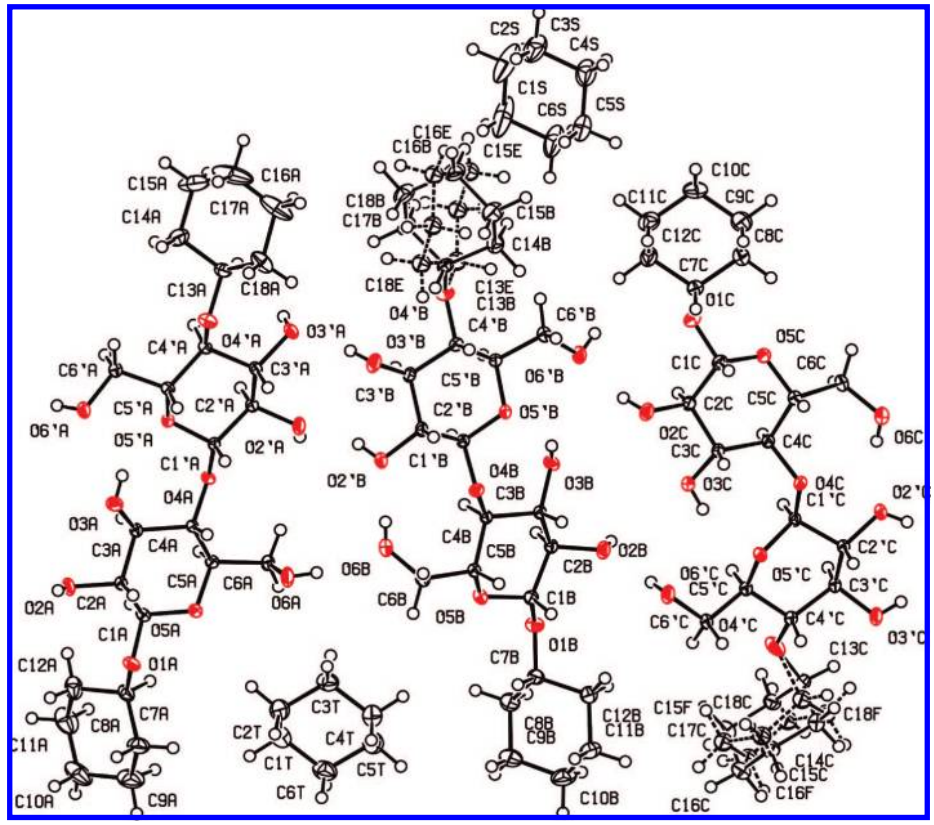


Figure 1. Thermal ellipsoid plot (50% probability) for **8** showing, from left to right, the three molecules of **8** (A, B, and C), and two cyclohexane molecules (S and T). Carbon and oxygen atoms are labeled. The molecules are projected onto the (210) plane. The minor components of the disordered B and C molecules are labeled E and F, respectively.

CH_2Cl_2, v/v = 1:19) to give a colorless solid, which was crystallized from EtOAc to afford **6** as colorless crystals (106 mg 21%), the filtrate containing an additional portion of the product (150 mg, 30%). R_f = 0.32 (MeOH/CH_2Cl_2, v/v = 1:9); mp = 219−221 °C, $[\alpha]^{20}_D$ = −41.7 (*c* 1.0, MeOH). ^{1}H NMR (CD_3OD): δ 1.19−2.05 (m, 20H, cyclohexyl and cyclohexylidene), 3.19 (br, dd, $J_{1,2}$ = 7.9 Hz, $J_{2,3}$ = 8.7 Hz, 1H, H-2 or H-2′), 3.28 (br, dd, $J_{1,2}$ = 7.8 Hz, $J_{2,3}$ = 8.9 Hz, 1H, H-2 or H-2′), 3.31−3.38 (m, 2H, H-5, H-5′), 3.47−3.48 (m, 2H, H-3, H-3′), 3.54−3.55 (m, 2H, H-4, H-4′), 3.68 (m, 1H, cyclohexyl), 3.74−3.90 (m, 4H, H-6a, H-6b, H-6′a, H-6′b), 4.39 (d, $J_{1,2}$ = 7.9 Hz, 1H, H-1 or H-1′), 4.48 (d, $J_{1,2}$ = 7.8 Hz, 1H, H-1 or H-1′). ^{13}C NMR (CD_3OD) δ 23.56, 23.78, 25.08, 25.27, 26.78, 26.83, 28.78, 32.83, 34.72, 39.01 (5 × CH_2 in cyclohexyl, 5 × CH_2 in cyclohexylidene), 61.86, 62.18 (C-6, C-6′), 68.86 (C-5 or C-5′), 73.67 (C-4′), 74.83, 74.96, 75.81, 76.09 (C-2, C-2′, C-3, C-3′), 76.19 (C-5 or C-5′), 78.38 (O-CH in cyclohexyl), 80.85 (C-4), 100.81 (O-C-O in cyclohexylidene), 102.15, 105.08 (C-1, C-1′). Anal. Calcd for $C_{24}H_{40}O_{11}$: C, 55.16; H, 8.10. Found: C, 55.28; H, 8.07.

Cyclohexyl 2′,3′-Di-*O*-acetyl-4′,6′-*O*-cyclohexylidene-*β*-D-glucopyranosyl-(1→4)-2,3,6-tri-*O*-acetyl-*β*-D-glucopyranoside (7). Ketal **6** (80 mg, 0.159 mmol) was stirred in pyridine (2 mL) and acetic acid anhydride (8 mL) in the presence of 4-(dimethylamino)pyridine (1 mg) at room temperature overnight. The reaction mixture was coevaporated with EtOH to dryness. The residue was purified by flash column chromatography (EtOAc/toluene, v/v = 1:4) to afford **7** (109 mg, 96%) as colorless syrup. The syrup was crystallized from CH_2Cl_2/*n*-hexane. R_f = 0.48 (EtOAc/toluene, v/v = 1:2); mp 189−191 °C; $[\alpha]^{20}_D$ = −30.9 (*c* 1.00, $CHCl_3$). ^{1}H NMR: δ 1.20−1.87 (m, 20H, cyclohexyl, cyclohexylidene), 2.00, 2.02, 2.03, 2.04, 2.11 (5*s*, 5 × 3H, CH_3 in acetate), 3.29 (dt, $J_{4',5'}$ = $J_{5',6'a}$ = 9.9 Hz, $J_{5',6'b}$ = 5.3 Hz, 1H, H-5′), 3.54 (ddd, $J_{4,5}$ = 9.6 Hz, $J_{5,6a}$ = 5.0 Hz, $J_{5,6b}$ = 2.0 Hz, 1H, H-5), 3.57 (m, 1H, O-CH in cyclohexyl), 3.68 (br, dd, $J_{3',4'}$ = 9.2 Hz, $J_{4',5'}$ = 9.9 Hz, 1H, H-4′), 3.73 (br, dd, $J_{5',6'a}$ = 9.9 Hz, $J_{6'a,6'b}$ = 10.6 Hz, 1H, H-6′a), 3.75 (br, dd, $J_{3,4}$ = 8.9 Hz, $J_{4,5}$ = 9.6 Hz, 1H, H-4), 3.92 (dd, $J_{5',6'b}$ = 5.3 Hz, $J_{6'a,6'b}$ = 10.6 Hz, 1H, H-6′b), 4.07 (dd, $J_{5,6a}$ = 5.0 Hz, $J_{6a,6b}$ = 11.9 Hz, 1H, H-6a), 4.45 (dd, $J_{5,6b}$ = 2.0 Hz, $J_{6a,6b}$ = 11.9 Hz, 1H, H-6b), 4.50 (d, $J_{1',2'}$ = 7.6 Hz, 1H, H-1′), 4.53 (d, $J_{1,2}$ = 7.9 Hz, 1H, H-1), 4.84 (dd, $J_{1,2}$ = 7.9 Hz, $J_{2,3}$ = 9.6 Hz, 1H, H-2), 4.88 (dd, $J_{1',2'}$ = 7.6 Hz, $J_{2',3'}$ = 9.2 Hz, 1H, H-2′), 5.07 (t, $J_{2',3'}$ = $J_{3',4'}$ = 9.2 Hz, 1H, H-3′), 5.15 (dd, $J_{2,3}$ = 9.6 Hz, $J_{3,4}$ = 8.9 Hz, 1H, H-3). ^{13}C NMR: δ 20.68, 20.76, 20.76, 20.93, 21.13 (5 × CH_3 in acetate), 22.53, 22.78, 23.64, 23.77, 25.52, 25.52, 27.52, 31.67, 33.24, 37.67 (5 × CH_2 in cyclohexyl, 5 × CH_2 cyclohexylidene), 61.32 (C-6′), 62.08 (C-6), 67.47 (C-5′), 70.15 (C-4′), 71.79 (C-2), 72.39 (C-5), 72.44 (C-3′), 72.53 (C-2′), 73.23 (C-3), 77.21 (C-4), 78.02 (O-CH in cyclohexyl), 99.05 (C-1), 99.84 (O−C−O in cyclohexylidene), 101.55 (C-1′), 169.26, 169.36, 169.42, 169.95, 170.19 (5 × CO in acetate). Anal. Calcd for $C_{34}H_{50}O_{16}$: C, 57.13; H, 7.05. Found: C, 56.89; H, 7.20.

Cyclohexyl 4′-*O*-Cyclohexyl-*β*-D-glucopyranosyl-(1→4)-*β*-D-glucopyranoside (8). To a solution of **7** (109 mg, 0.153 mmol), in anhydrous CH_2Cl_2 (10 mL) were added borane trimethylamine complex (56 mg, 0.763 mmol), molecular sieve (4 Å, powder), and $AlCl_3$ (102 mg, 0.763 mmol) at −72 °C. The solution was stirred for 1 h at this temperature, and then a second portion of borane trimethylamine complex (56 mg, 0.763 mmol) and $AlCl_3$ (102 mg, 0.763 mmol) was added. The solution was stirred for another h, and ion-exchange resin DOWEX 50W X8 and MeOH were added at −72 °C. Solids were filtered off, and the filtrate was evaporated. The residue was purified by flash column chromatography (EtOAc/toluene (v/v = 1:2), then EtOH/EtOAc/toluene, v/v/v = 1:2:2). To a solution of the product in MeOH/CH_2Cl_2 (v/v = 1:1, 10 mL) NaOMe in MeOH (0.1 M, 152 μL) was added. The solution was stirred at room temperature overnight. NaOMe was neutralized with DOWEX 50W X8 ion-exchange resin. The solid was filtered off, and the filtrate was evaporated. The residue was purified by flash

column chromatography (EtOH/EtOAc/toluene, v/v/v = 3:10:10) to give target **8** (50 mg, 65%) as colorless solid. Recrystallization was performed from cyclohexane/MeOH (v/v = 3:1). $R_f = 0.30$ (EtOH/EtOAc/toluene, v/v/v = 3:10:10), mp 220−280 °C, $[\alpha]^{20}_D = -14.9$ (*c* 0.490, MeOH). ^{1}H NMR (CD_3OD): δ 1.15−2.04 (m, 20H, 2 × cyclohexyl), 3.21 (dd, $J_{1,2} = 7.9$ Hz, $J_{2,3} = 8.9$ Hz, 1H, H-2), 3.21 (dd, $J_{1',2'} = 7.9$ Hz, $J_{2',3'} = 9.0$ Hz, 1H, H-2′), 3.29 (ddd, $J_{4',5'} = 8.5$ Hz, $J_{5',6'a} = 5.0$ Hz, $J_{5',6'b} = 2.3$ Hz, 1H, H-5′), 3.33 (t, $J_{3',4'} = J_{4',5'} = 8.5$ Hz, 1H, H-4′), 3.38 (ddd, $J_{4,5} = 9.1$ Hz, $J_{5,6a} = 4.0$ Hz, $J_{5,6b} = 2.9$ Hz, 1H, H-5), 3.40 (dd, $J_{2',3'} = 9.0$ Hz, $J_{3',4'} = 8.5$ Hz, 1H, H-3′), 3.50 (t, $J_{2,3} = J_{3,4} = 8.9$ Hz, 1H, H-3), 3.54 (br, dd, $J_{3,4} = 8.9$ Hz, $J_{4,5} = 9.1$ Hz, 1H, H-4), 3.61 (tt, $J_{aa} = 9.4$ Hz, $J_{ae} = 3.9$ Hz, 1H, 4′-*O*-CH in cyclohexyl), 3.65 (dd, $J_{5',6'a} = 5.0$ Hz, $J_{6'a,6'b} = 11.8$ Hz, 1H, H-6′a), 3.69 (tt, $J_{aa} = 9.4$ Hz, $J_{ae} = 3.8$ Hz, 1H, 1-*O*-CH in cyclohexyl), 3.84 (dd, $J_{5',6'b} = 2.3$ Hz, $J_{6'a,6'b} = 11.8$ Hz, 1H, H-6′b), 3.85 (dd, $J_{5,6a} = 4.0$ Hz, $J_{6a,6b} = 12.1$ Hz, 1H, H-6a), 3.86 (dd, $J_{5,6b} = 2.9$ Hz, $J_{6a,6b} = 12.1$ Hz, 1H, H-6b), 4.39 (d, $J_{1',2'} = 7.9$ Hz, 1H, H-1′), 4.40 (d, $J_{1,2} = 7.9$ Hz, 1H, H-1). ^{13}C NMR: δ 25.40, 25.52, 25.83, 25.85, 27.13, 27.20, 33.19, 34.27, 34.73, 35.06 (5 × CH_2 in 1-*O*-cyclohexyl, 5 × CH_2 in 4′-*O*-cyclohexyl), 62.32 (C-6), 62.38 (C-6′), 75.20 (C-2), 75.45 (C-2′), 76.74, 76.83, 76.94 (C-3, C-5, C-4′), 77.83 (C-5′), 78.40 (C-3′), 78.78 (O-CH in 1-*O*-cyclohexyl), 80.56 (O-CH in 4′-*O*-cyclohexyl), 81.14 (C-4), 102.61 (C-1), 104.93 (C-1′). Anal. Calcd for $C_{24}H_{42}O_{11}$: C, 56.90; H, 8.36. Found: C, 56.56; H, 8.36.

Results and Discussion

Synthesis. Target compound **8** was synthesized in 16% overall yield, starting from cellobiose octaacetate (**1**) according to the seven-step sequence shown above in Scheme 1. After selective deprotection of the reducing end to give cellobiose heptaacetate (**2**), glycosidation with cyclohexanol was performed under "Schmidt conditions",[26] i.e., via the corresponding trichloroacetimidate (**3**), which afforded 73% of the β-configured product (**4**) when working in anhydrous CH_2Cl_2 in the presence of powdered molecular sieve (4Å) and $BF_3 \cdot Et_2O$ catalyst at −18 °C. The resulting cyclohexyl β-D-cellobioside peracetate (**4**) was deprotected into cyclohexyl β-D-cellobioside (**5**), and a 4′-*O*,6′-*O*-cyclohexylidene ketal[27] was introduced (**6**). This approach was chosen after all alternative attempts to introduce the 4′-O-cyclohexyl group by etherification failed. After peracetylation (**7**), this cyclohexylidene ketal ring was opened regioselectively, which was the key step in the sequence, providing 65% of the final compound (**8**) after deacetylation. The regioselective opening of the cyclohexylidene ketal was comprehensively studied and optimized, and more than 20 reductant/Lewis acid combinations under different solvent and temperature conditions were tested, starting from conditions used for the opening of the common benzylidene acetals. The reagent pair boran trimethylamine complex/aluminum chloride in dichloromethane at −72 °C gave the best results, providing 75% of the desired 4′-*O*-cyclohexyl derivative, with only 10% of the 6′-*O*-cyclohexyl side product. Crystals for diffraction were grown from cyclohexane/MeOH (v/v = 3:1) by slow evaporation.

Basic Crystal Structure. Table 1 presents the key crystallographic information. The triclinic $P\bar{1}$ unit cell contains three molecules (A, B, and C) of **8**, each with its own geometric features, as well as two separate cyclohexane solvent molecules (S and T). The finding of three crystallographically unique substrate molecules in the unit cell is rare, and novel for small cellulose analogs. Most have one or two (symmetry-related) molecules in the unit cell. Cellotetraose has two similar molecules,[19] and methyl cellotrioside has four crystallographically unique but structurally similar molecules per unit cell.[28] For triclinic crystal structures, a reduced cell is conventionally used. Our "preferred" cell differs from the reduced cell, also reported in Table 1, that was automatically derived during analysis of the X-ray data. Both cells have the same volumes and *a*- and *b*-dimensions, but their *c*-axes and α- and β-angles differ.

Table 1. Crystal Data and Structure Refinement for Cyclohexyl 4′-*O*-Cyclohexyl β-Cellobioside

empirical formula	$C_{84}H_{150}O_{33}$	
moiety formula	$3(C_{24}H_{42}O_{11})$, $2(C_6H_{12})$	
formula weight	1688.04	
temp (K)	100(2)	
wavelength (Å)	0.71073	
crystal system, space group	triclinic, $P\bar{1}$	
unit cell dimensions	preferred	reduced
a (Å)	10.7024(7)	10.7024(7)
b (Å)	10.9502(7)	10.9502(7)
c (Å)	23.4669(15)	19.0194(12)
α (deg)	80.043(1)	78.077(1)
β (deg)	52.828(1)	79.467(1)
γ (deg)	89.314(1)	89.314(1)
volume (Å^3)	2143.4(2)	
Z, calcd density (g/cm^3)	1, 1.308	
absorption coefficient (mm^{-1})	0.099	
F(000)	918	
crystal size (mm^3)	0.60 × 0.60 × 0.20	
θ range for data collection (deg)	2.40 to 30.07	
index ranges	$-15 \le h \le 14$, $-15 \le k \le 15$, $-26 \le l \le 26$	
reflections collected/ unique	38659/12443 [R_{int} = 0.0257]	
completeness to θ = 30.07°	98.9%	
absorption correction	multiscan (program SADABS)	
max and min transmission	0.98 and 0.92	
refinement method	full-matrix least-squares on F^2	
data/restraints/ parameters	12443/336/1112	
goodness-of-fit on F^2	1.041	
final *R* indices [$I > 2\sigma(I)$]	$R_1 = 0.0451$, $wR_2 = 0.1121$	
R indices (all data)	$R_1 = 0.0492$, $wR_2 = 0.1161$	
absolute structure parameter	0.1(4)	
largest diff peak and hole (e Å^{-3})	0.77 and −0.36	

We chose the less-conventional cell because the molecular alignment deviates less from its *c*-axis. In cellulose the *c*-axis and the molecular axes coincide. In the tetraose[19] structure, the alignment is closer than in **8** but not quite coincident.

Figure 1 shows the thermal ellipsoids from the structure determination as well as the disordered cyclohexyl moieties on the nonreducing ends of the B and C molecules. Most ellipsoids are satisfactorily small except for atoms of one cyclohexyl residue (nonreducing end of A) and one cyclohexane solvent molecule (S). A closer view of the disordered ends is given in Figure S1, Supporting Information. The B and C labels also apply to the major components of the disordered cyclohexyl rings (62.4% and 61.0%, respectively). The minor components

(26) Hoffmann, M. G.; Schmidt, R. R. *Liebigs Ann. Chem.* **1985**, *12*, 2403.
(27) Bissett, F. H.; Evans, M. E.; Parrish, F. W. *Carbohydr. Res.* **1967**, *5*, 184.
(28) Raymond, S.; Henrissat, B.; Qui, D. T.; Kvick, Å.; Chanzy, H. *Carbohydr. Res.* **1995**, *277*, 209–229.

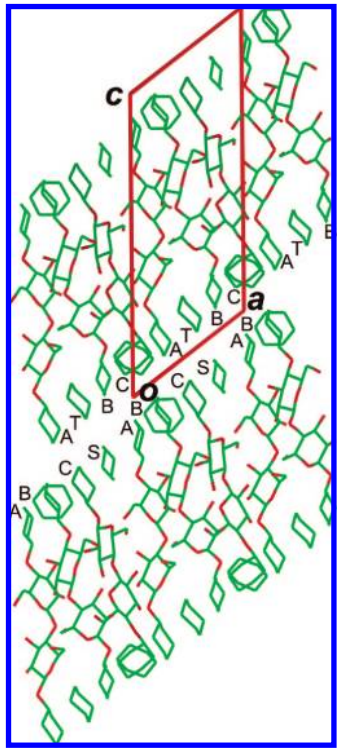


Figure 2. Projection down the *b*-axis of the unit cell, showing the *a*- and *c*-axes and two of the layers of **8**. Carbon atoms are green and oxygen atoms are red; hydrogen atoms are not shown. Letters denote the five different molecules shown in Figure 1, and the disordered ends of B and C are shown.

of the disordered cyclohexyl groups are given letters E (37.6% occupancy, alternative part of the B molecule) and F (39.0%, part of the C molecule). Bonds from the disordered C and F cyclohexyl rings to O4′ are axial, while linkages from all other cyclohexyl groups are equatorial.

Crystal Packing; Chain Orientation. Aspects of the crystal packing are shown in Figures 2−6. Figure 2, a view down the *b*-axis, shows that the constituents of the crystal structure are arranged in layers that are tilted with respect to the *c*-axis. The bonded and solvent cyclohexanes comprise the top and bottom, relatively nonpolar surfaces of each layer, with the central cellobiose residues being able to form both O−H···O and C−H···O hydrogen bonds and less polar van der Waals interactions. Unlike a lipid bilayer, single molecules of **8** span the entire thickness of the layer, and the "head groups" are nonpolar. The disordered cyclohexyl group on the B molecule is on the top of each layer, and the disordered group on the C molecule is on the bottom. As also shown in Figure 1, the A and B molecules have "down" orientations in the unit cell, that is, the z-coordinates of their O5 atoms are greater than the z-coordinates of the O1 atoms. On the other hand, the C molecule is oriented "up." Therefore, the unit cell contains twice as many down-pointing molecules as up, giving a down-down-up packing. Such packing is rare but has been found for the 3-fold double helices of welan[29] and calcium ι-carrageenan,[30] structures that have trigonal unit cells. Another molecule, poly-L-lactide, can pack in a north-south-south (up-down-down)

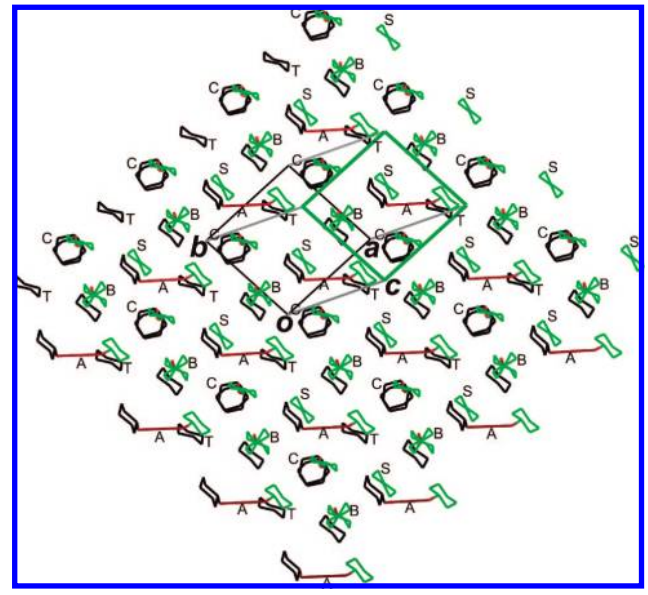


Figure 3. View from the top of the crystal structure with the carbohydrate portion replaced by straight, red lines. The molecules are aligned so that the carbohydrate lines of the B and C molecules are almost perpendicular to the page (red dots), whereas the A molecule is tilted, as indicated by its "carbohydrate line" that stretches from one row of molecules parallel to the *a*-axis to its adjacent row. Green lines are at the top of the layer, and black are at the bottom.

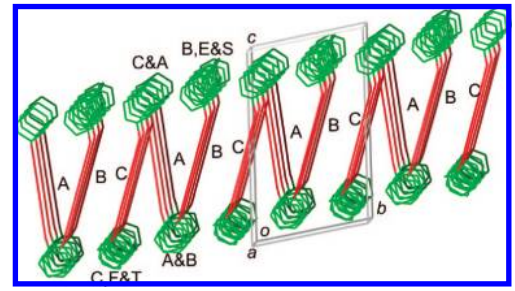


Figure 4. View of stacking of the cyclohexyl moieties, projected down the *a*-axis of a single layer, with the carbohydrate replaced by a straight line. The view is slightly rotated so that four repeats of each molecule can be seen. Three "N"-shaped groupings are shown, along with the A and B partners on the left side and C molecules without the A and B partners on the right.

scheme.[31] Although not a double helix, it is a 3-fold helix that also packs in a trigonal unit cell. More relevant to cellulose, Lotz has speculated[32] that the correct packing of 3-fold triethyl cellulose helices could be similar to poly-L-lactide instead of the six-chain cell proposed by Zugenmaier.[33] The molecules of **8** do not correspond to 3-fold helices. Finally, γ-chitin, which has a cellulosic backbone, has been proposed to contain three chains per cell based on early work, with two up and the other down or vice versa.[34]

In Figure 2, counterpart cyclohexane molecules interact at the interface between the top of the lower layer and the bottom of the upper layer. Thus, the S cyclohexane molecule is closest to the T solvent molecule. Also, the reducing-end groups of

(29) Chandrasekaran, R.; Radha, A.; Lee, E. J. *Carbohydr. Res.* **1994**, *252*, 183–207.
(30) Janaswamy, S.; Chandrasekaran, R. *Carbohydr. Res.* **2001**, *335*, 181–194.
(31) Puiggali, J.; Ikada, Y.; Tsuji, H.; Cartier, L.; Okihara, T.; Lotz, B. *Polymer* **2000**, *41*, 8921–8930.
(32) Cartier, L.; Spassky, N.; Lotz, B. *C. R. Acad. Sci. Paris* **1996**, *322*, 429–435.
(33) Zugenmaier, P. *J. Appl. Polym. Sci, Appl. Polym. Symp.* **1983**, *37*, 223–238.
(34) Rudall, K. M. *Adv. Insect Physiol.* **1963**, *1*, 257–313.

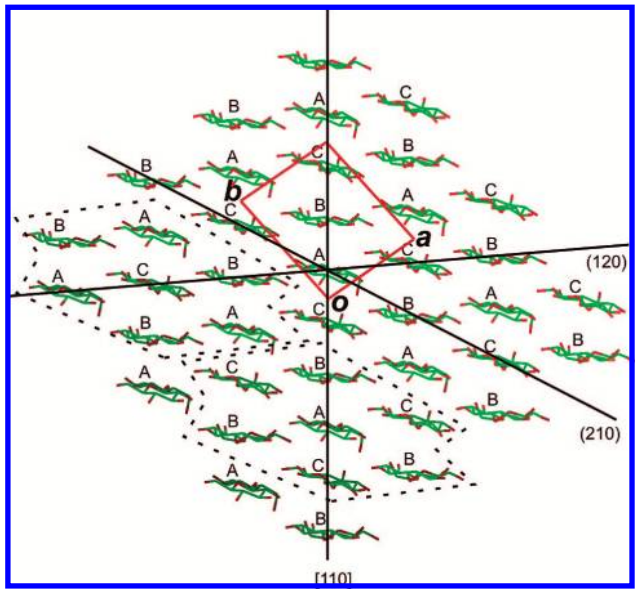


Figure 5. Packing of the carbohydrate moieties of the A, B, and C molecules, projected down the *c*-axis. The unit cell, the [110] line, and the (210) and (120) planes are indicated. Dashed lines surround the six molecules nearest to the A and C molecules.

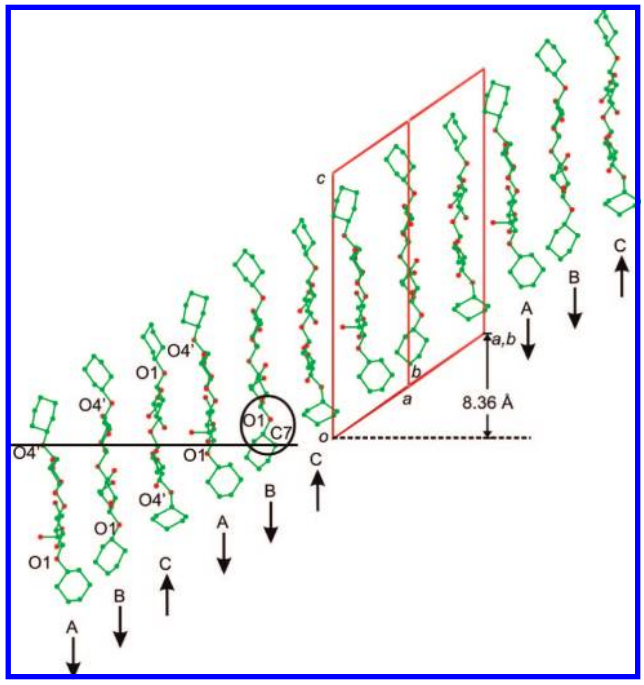


Figure 6. A, B, and C molecules, arranged along the [110] line (see Figure 5) that connects the *o* and *a*,*b* points of the base of the unit cell. Arrows indicate the "up" or "down" orientation of the chains. The minor fractions of the disordered cyclohexyl substituents are not shown. The heavy horizontal line passes through the O4′ of the left-most A molecule and underneath the C7 atom (circled) of the fifth molecule from the left. If the line passed through O1 instead, that would show perfect "quarter stagger" of the chains. The dashed line and 8.36 Å distance show the advance along the *c*-axis for three chains, or 2.79 Å for each.

the C and A molecules are close, as are the reducing B and nonreducing A ends.

The disordered B and E cyclohexyl groups interact with the disordered C and F groups. B-C and E-C combinations had no contacts between layers that were shorter than 2.6 Å, whereas B-F and E-F had contacts of 2.21 and 2.28 Å, respectively (Figure S1)

Cyclohexane. The solvent molecules fill in gaps on the upper and lower surfaces of the individual layers, presumably compensating for the lesser bulk of a cyclohexyl group compared to a glucopyranosyl residue. Further understanding of the role of the cyclohexyl groups was attained by drawing the carbohydrate portion of each molecule of **8** as a straight line in Figures 3 and 4. In Figure 3, tilted slightly away from a view down the *c*-axis, the projected carbohydrate line for the A structures is fairly long, indicating that A is tilted relative to B and C, which have very short projected lines. The S molecule is above the lower end of the A molecule and the T molecule is below the upper end of the A molecule. The disordered end of each B molecule appears in this view to have a propeller shape arising from the major and minor components, and the two components of the disordered ends of the C molecules are seen as hexagons that are rotationally offset from each other but otherwise superimposed.

In Figure 4, a view along the *a*-axis, the disordered cyclohexyl parts of the C molecules are stacked with the T solvent molecules. The ordered cyclohexyl moieties on the reducing ends of the C molecules are interspersed with the ordered cyclohexyl groups on the nonreducing ends of the A molecules, and the ordered cyclohexyl groups on the reducing ends of A are interspersed with ordered cyclohexyls on the reducing ends of the B molecules. The N-shaped sequence of C-A-B ends with the disordered cyclohexyl moieties on the nonreducing end of B interspersed with S molecules, an arrangement very similar to the nonreducing end of the C molecules at the beginning of the sequence. The importance of the interactions of the nonpolar groups is underscored by the alignments of the cyclohexyl moieties in this *a*-axis projection and the differing tilts for the lines from the different cellobiose portions. Additional views of the cyclohexane stacking are in Figure S2, Supporting Information.

Carbohydrate. To best view the features of the carbohydrate portion, the cyclohexyl moieties have been eliminated from Figure 5, which shows the packing of the cellobiose residues. The view is down the *c*-axis and is similar to a view down the chain axis of cellulose II or cellotetraose. The six neighbors around the A and C residues are each enclosed by dashed lines. To reduce clutter, an enclosure for B residues is not shown. Each kind (A, B, or C) is surrounded entirely by the two other kinds of residue. A consequence is that the "up" C molecules are surrounded only by "down" A and B molecules, whereas the As and Bs are surrounded by both up and down molecules. A molecules are not in contact with other As, and Bs are not in contact with other Bs. Despite the visual similarity with the packing in cellulose II and cellotetraose, the details for **8** are different because some of the contacts in those other structures are between identical molecules that are related by translational symmetry.

Also shown in Figure 5 are the *a* and *b* unit cell axes and the three alignments of molecules that make intermolecular hydrogen-bonding contacts. The [110] lines pass through a stack of residues with their relatively nonpolar sides facing each other. On the other hand, the (210) planes contain sheets of molecules that are all linked by hydrogen bonds, each residue to both of its in-sheet neighbors. The (120) planes also contain sheets, but as will be seen below, only three molecules are linked with hydrogen bonds before there is an interruption.

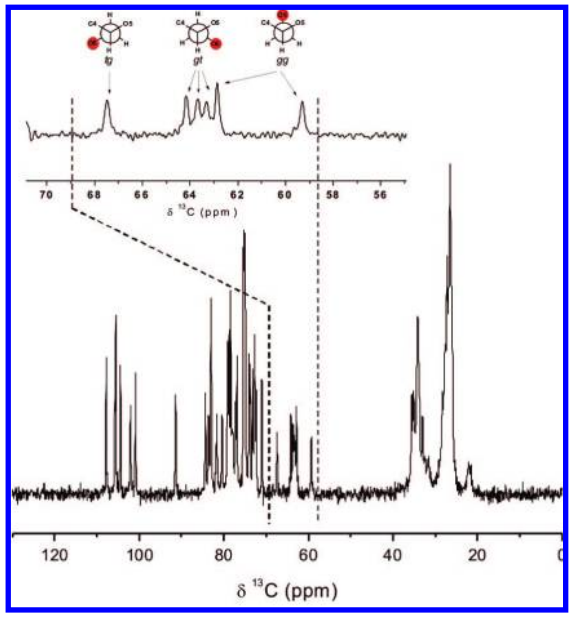


Figure 7. Solid-state (^{13}C CPMAS, CP = 1 ms, 150.9 MHz) NMR of compound **8**. Inset: C-6 region showing all three preferred conformations, *gt*/*gg*/*tg* = 3:2:1.

Approximate Quarter Staggering in the [110] Stack. An important feature of the [110] stack is indicated in Figure 6 by a horizontal line. This line intersects the O4′ of the first molecule on the left, and is just under the C7 on the second B molecule to its right. That B molecule is the fourth molecule after the left-most A molecule. If the upward shift, or stagger, for each molecule were an advance of one-fourth of the length of a cellobiose residue, the horizontal line would intersect O1 on the fourth molecule, but the advance is slightly greater. Working with the unit cell geometry, the *a*,*b* point is 8.36 Å higher than the origin point (*o*) for a distance that spans three molecules. This corresponds to a shift of 2.79 Å for one molecule or 11.15 Å for four molecules, a number that can be compared with $d_{O1-O4'}$ values of 10.21 Å, 10.35 Å and 10.13 Å for the A, B and C residues. Similarly continuous "quarter stagger" (2.81 Å) is found in parallel-packed, triclinic cellulose Iα. In parallel cellulose Iβ,[2] antiparallel II,[4] and cellotetraose,[19] alternating chains of the unit cell are also elevated by about one-fourth (2.72, 2.40, and 2.44 Å, respectively) of the cellobiose unit, compared to the other chains. In cellulose III_I,[5] there is no stagger of the chains.

Hydrogen Bonding and O6 Orientation. All 18 are donors, 15 are hydroxyl acceptors, and three are C−H acceptors. Altogether, there are 28 different O−H···O hydrogen bonds, all networked. We have also identified 14 different C−H···O hydrogen bonds. Because C−H groups cannot be acceptors, they only participate in networks as branches. Three of the C−H···O hydrogen bonds are not networked because their acceptors are ether oxygens that do not happen to also accept other protons. Of the 15 ether oxygens, nine are either O−H or C−H acceptors. No reducing ring O5 atoms are involved, nor are O1B, O4A or O4′C.

These hydrogen bonding systems stabilize the O6 groups in all three of the staggered orientations, namely *gt*, *gg* and *tg*. ^{13}C cross-polarization magic angle spinning (CP-MAS) NMR (Figure 7) confirmed the presence of all three C6-OH orientations in one crystal structure. The six distinct peaks from the different O6 groups are indicated in the inset along with their assignments.[35] In **8**, all O6 orientations (Table 5) are within 23° of the ideal O5−C5−C6-O6 torsion angles of +60°, −60° and 180°, respectively. It is unusual to find examples with the *tg* orientation for molecules having the *gluco* configuration at O4, although the *tg* orientation is found in both native forms of cellulose, Iα[1] and Iβ.[2] Structures having all three forms in the same crystal are truly rare. Only two crystals in the Cambridge Structural Database[36] have CH_2OH orientations that are within 30° of the ideal values for each of the three orientations. Both are β-cyclodextrins.[37,38] The O6 *gg*, O6′ *gt* motif was found in both crystal structures of the 1,4′-dimethyl cellobioside, while cellotetraose has all O6 in the *gt* orientation.

Table 2. Intramolecular O−H···O Hydrogen Bonding Geometries[a]

donor	acceptor	H···O (Å)	O···O (Å)	O−H···O (deg)
O3A−H3A	O5′A	1.82	2.719(2)	153
O3A−H3A	O6′A	2.37	2.984(2)	120
O3B−H3B	O5′B	2.03	2.847(2)	141
O3B−H3B	O6′B	2.53	3.276(2)	134
O3C−H3C	O5′C	1.84	2.755(2)	157
O3C−H3C	O6′C	2.41	3.061(2)	124
O3C−H3C	O4C	2.72[b]	3.049(2)	100
O6B−H6B	O4B	2.62[b]	3.066(2)	108
O6B−H6B	O2′B	1.83	2.800(2)	175
O6C−H6C	O4C	2.32	2.864(2)	115
O6C−H6C	O2′C	2.06	3.014(2)	166
O3′C−H3′C	O2′C	2.57	2.995(2)	107

[a] d(D···A) < R(D) + R(A) + 0.50 Å = 3.54 Å (O−H···O), 3.72 Å (C−H···O); d(H···A) < R(H) + R(A) − 0.12 Å = 2.60 Å; D−H···A > 100.0°. Covalent radii (Å) for C: 1.70; H: 1.20; O: 1.52.[39] [b] This distance exceeds the distance criteria for hydrogen bonding in the PLATON program[39] but fits the angular criteria well. As part of a 3-center bond, the O2′C···O3′A's 144° angle, the 149° angle for the O2′C···O2A bond, and the 65° O3A-H2′C−O2A angle of 65.5° add up to 358°, close to the ideal value of 360° for a 3-center hydrogen bond.

One way to appreciate the impact of the hydrogen bonding is in terms of the sheets and stacks of molecules that they appear to organize. Another consideration is the network of cooperative hydrogen bonds. For this discussion, all O−H and C−H covalent bond lengths were set to the neutron diffraction values of 0.97 and 1.10 Å,[39] respectively, and the geometric data were recalculated. Geometric features are in Tables 2−4, along with the criteria for defining hydrogen bonds from the PLATON program.[40] We also scanned the structures with a more-inclusive H···O criterion of 2.85 Å (for the O−H···O bonds only) and found four additional interactions that are included in Tables 2 and 3, as well as Figures 8 and 9.

Two of those additional bonds are critical to the discussion of the networking in **8**, below. One is an intramolecular link involving O3C, H3C, and O4C, with an H···O distance of 2.72 Å, an O···O distance of 3.049 Å, and an O−H···O angle of 100°; the other is intermolecular with O2B, H2B, and O2A, with values of 2.73 Å, 3.248 Å and 114°. The PLATON software reports two independent networks, and we have retained that result. However, either of these two long and angular bonds would link the two networks, leaving only one

(35) Horii, F.; Hirai, A.; Kitamaru, R. *Polym. Bull. (Berlin)* **1983**, *10*, 357–61.
(36) Allen, F. H. *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58*, 380–388.
(37) Hamilton, J. A.; Sabesan, M. N. *Carbohydr. Res.* **1982**, *102*, 31–46.
(38) Pop, M. M.; Goubitz, K.; Borodi, G.; Bogdan, M.; De Ridder, D. J. A.; Peschar, R.; Schenk, H. *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58*, 1036–1043.
(39) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1991.
(40) Spek, A. L. *J. Appl. Crystallogr.* **2003**, *36*, 7–13.

Table 3. Intermolecular O−H···O hydrogen bonds

donor	acceptor	ARU[a]	H···O (Å)	O···O (Å)	O−H···O (deg)
O2A−H2A	O3B	1465	1.67	2.633(2)	170
O6A−H6A	O6B	within	1.88	2.816(2)	162
O2′A−H2′A	O6′C	1565	1.79	2.748(2)	168
O3′A−H3′A	O2C	1565	1.78	2.725(2)	163
O6′A−H6′A	O6C	1475	1.81	2.747(2)	161
O2B−H2B	O1A	1645	1.83	2.740(2)	156
O2B−H2B	O2A	1645	2.73[b]	3.248(2)	114
O2′B−H2′B	O2′A	within	1.77	2.692(2)	158
O3′B−H3′B	O3′A	within	1.75	2.713(2)	170
O6′B−H6′B	O3A	1645	2.60	3.069(2)	110
O6′B−H6′B	O6′A	1645	1.94	2.857(2)	156
O2C−H2C	O6′B	within	1.76	2.710(2)	165
O2′C−H2′C	O2A	1635	2.16	3.006(2)	144
O2′C−H2′C	O3A	1635	2.77[b]	3.633(2)	149
O3′C−H3′C	O2A	1635	1.82	2.717(2)	152
O6′C−H6′C	O2B	within	1.75	2.715(2)	170

[a] ARU = asymmetric residue unit. [1645] = 1 + *x*, −1 + *y*, *z*; [1545] = *x*, −1 + *y*, *z*; [1635] = 1 + *x*, −2 + *y*, *z*; [1565] = *x*, 1 + *y*, *z*; [1465] = −1 + *x*, 1 + *y*, *z*; [1475] = −1 + *x*, 2 + *y*, *z*. [b] This distance exceeds the PLATON criterion.[40]

Table 4. C−H···O Hydrogen Bonds

donor	acceptor	ARU[a]	H···O (Å)	C···O (Å)	C−H···O (deg)
C2S−H22S	O3′B	within	2.60	3.435(4)	132
C13A−H13A	O3′A	intramolecular	2.46	3.044(2)	112
C13C−H13C	O3′C	intramolecular	2.40	2.968(4)	110
C2A−H2AA	O3C	1465	2.44	3.543(3)	178
C6A−H61A	O6′C	1565	2.47	3.532(3)	162
C5′A−H5′A	O1C	1465	2.43	3.476(3)	159
C6′A−H62A	O4B	1565	2.58	3.583(3)	151
C2B−H2BB	O3′C	1565	2.53	3.594(3)	162
C5′B−H5′B	O4C	1565	2.43	3.506(2)	165
C2C−H2CC	O3A	1645	2.48	3.571(3)	170
C5C−H5CC	O4′B	1545	2.42	3.059(3)	170
C1'C−H1'C	O6A	1645	2.56	3.438(3)	136
C2′C−H2"C	O3B	1545	2.50	3.473(3)	147
C8C−H81C	O4′A	1645	2.59	3.674(4)	170

[a] See Table 3 for ARU definitions.

Table 5. O6 Orientations

residue	nominal	X(O5−C5−C6−O6) (deg)	X(C4−C5−C6−O6) (deg)
A (reducing)	*gg*	−63.3	59.1
A′ (nonreducing)	*gt*	66.7	−174.1
B (reducing)	*gg*	−82.3	37.6
B′ (nonreducing)	*gt*	55.8	174.6
C (reducing)	*tg*	165.9	−75.6
C′ (nonreducing)	*gt*	50.9	170.0

network. Thus, two of the extra-long hydrogen bonds are included in the depiction of the networks in Figure 11 (below), but the two network-combining bonds are not shown there. We did not combine the networks, partly because that would depend on the somewhat dubious long bonds, but also because it greatly complicated the graphical depiction of already complicated networks.

Intramolecular. As shown in Figures 8 and 9, all O5′ atoms are acceptors for the O3 hydroxyl hydrogens, an arrangement found in all cellulose forms so far and frequently in small analogs. The O3-H is also donated to O6′ in a three-centered or bifurcated arrangement. The liaison with O6′ is less frequently discussed because of its longer length, 2.37 Å to 2.53 Å in **8**, but it appears to stabilize the O6′ atoms in the *gt* orientation. The reducing-end O6 atoms, on the other hand, take either *gg* or *tg* orientations. In our C molecule, O6 has a *tg* orientation and donates its hydrogen to O2′, as in the minor disordered schemes in cellulose I.[1,2] The B molecule also has an O6···O2′

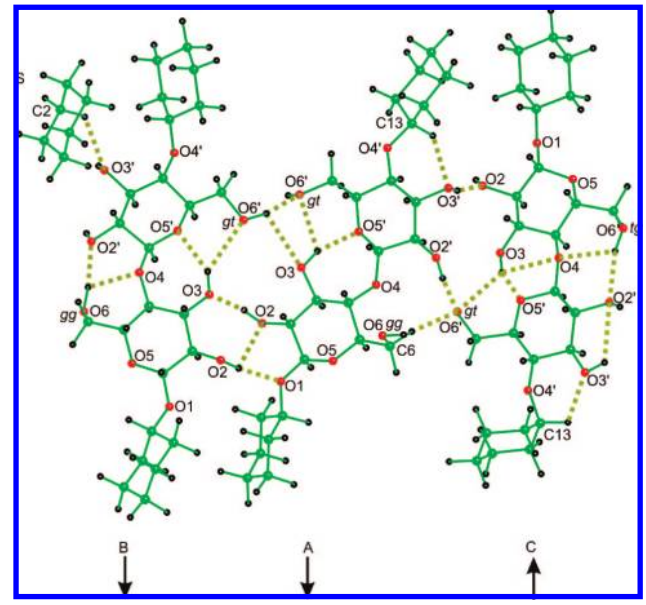


Figure 8. Hydrogen bonding and O6 orientation (*gt*, *gg*, or *tg*) in the three molecules connected by hydrogen bonds in the (120) sheets, as well as S (cyclohexane). Four C−H···O hydrogen bonds are shown.

link, but its O6 is only 22° from the ideal *gg* orientation and 98° from the ideal *tg* orientation. The O6···O2′ link based on O6 *gg* has not been observed previously. O6B and O6C also have secondary, three-centered hydrogen bonds to the inter-residue glycosidic oxygen atom, O4. Only one hydroxyl orientation remotely resembles the "clockwise" or "counter-clockwise" hydrogen bonds favored for isolated-molecule molecular modeling studies. On C, an approximately counter-clockwise O3′-H donates to O2′, but O2′ is a double acceptor and its hydrogen is involved in a three-centered, intermolecular arrangement in the (210) sheet.

(120) Sheet O−H···O Bonds. Although the glucose rings are evenly spaced and their planes are well aligned with the (120) plane (Figure 5), the intermolecular hydrogen bonding does not link the molecules continuously. As shown in Figure 8, the sheet segments in the (120) plane contain only three molecules, with the A molecule in the center. Two donor hydroxyls in bonds between B and A are on B, and A reciprocates with one.

Although O6′A and O6′B are linked together on the nonreducing ends, the molecules are offset and O2A donates to O3B and O2B donates to O1A as well as to O2A (a very weak, bifurcated link). Both O−H···O hydrogen bonds between A and C have A atoms as donors.

(210) Sheet O−H···O Bonds. Unlike the (120) sheet, all molecules in the (210) sheet in Figure 9 are hydrogen-bonded to their neighbors, continuously throughout the crystal. A novel arrangement exists in this sheet, namely, that like atoms in A and B are linked. Thus, O2′B donates to O2′A, O3′B donates to O3′A, and O6′A donates to O6C. As in the (120) sheet, the A and B molecules are turned toward each other. This is in contrast to other known structures of cellulose and its analogs. Therein, when there are adjacent parallel molecules within a given sheet, the molecules are related by translational symmetry and each has the same rotational orientation. Therefore, O6 atoms on one molecule can interact with O3 or O2 atoms on adjacent molecules.

Because B and C are antiparallel, they can have reciprocal heterogeneous pairings: O2B receives from O6′C and O2C

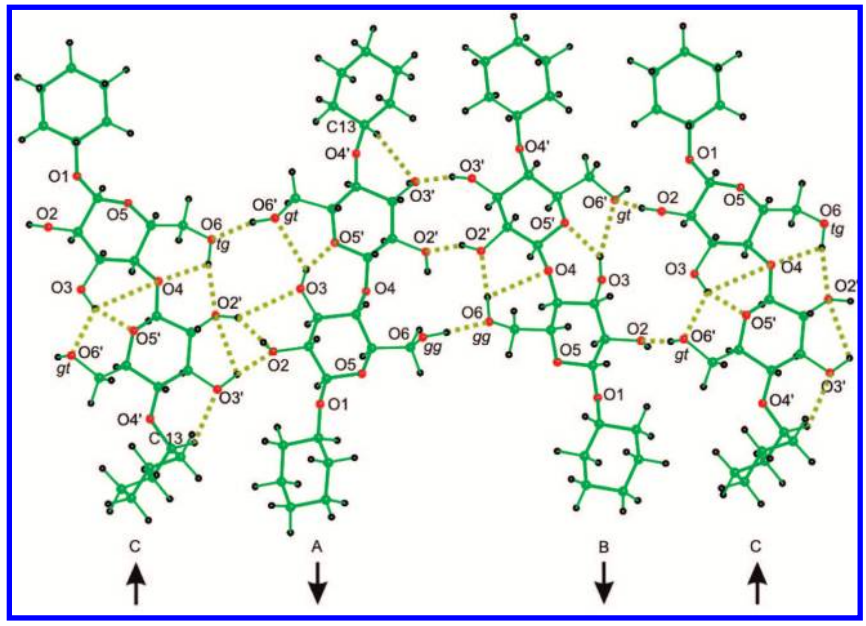


Figure 9. Four molecules of **8** in the (210) plane, in which all neighboring molecules are hydrogen bonded to each other. The second C molecule is shown so that the B−C hydrogen bonds are included. The S and T molecules are not shown, nor are E and F. Intramolecular hydrogen bonds are repetitions from Figure 7.

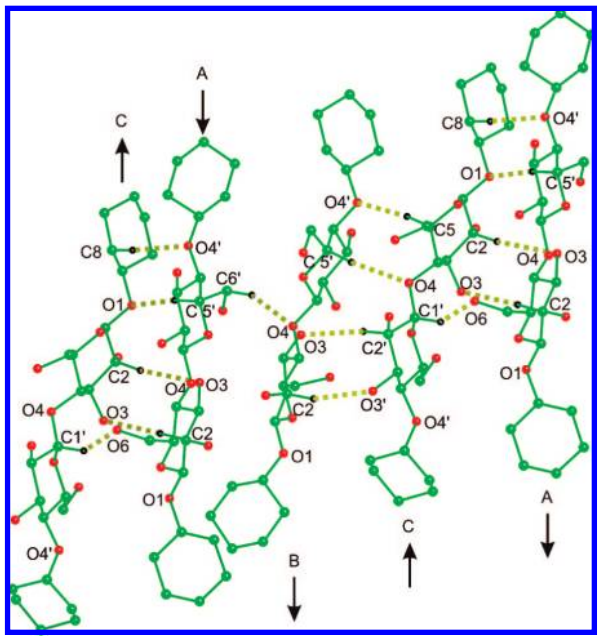


Figure 10. Ten intermolecular C−H···O hydrogen bonds in the [110] stack; intramolecular hydrogen bonds are not shown, nor are hydrogen atoms that are not involved in hydrogen bonds.

donates to O6′B. Between antiparallel C and A, O6′A donates to O6C, and O2′C donates to both O2A and O3A (minor component). The latter bond does not meet the PLATON H···O distance criterion but would complete a homodromic cycle. O3′C also donates to O2A.

C−H···O Hydrogen Bonds. Ten intermolecular C−H···O bonds are among molecules in the [110] stack, as shown in Figure 10, and the other four are indicated in Figure 8. One intermolecular bond in Figure 8 is from the S cyclohexyl solvent molecule to the B molecule, and another links C6A to O6′C. Two intramolecular bonds in Figure 8 involve the A and C cyclohexyl groups on nonreducing ends. Considering all the C−H···O bonds, seven involve O3 or O3′ as acceptors, with other acceptors being O4 and O4′ (four instances), O6 and O6′ (one each), and O1. In cellotetraose, only the glycosidic oxygen atoms are acceptors for C−H donors. Besides C2 on S, C2 and C2′ atoms donate four times, with others being C5′ (three instances), C13 (twice), C1′, C6, C6′ and C8.

Networks. The finite networks are illustrated in Figure 11. Network 1 consists of nine O−H···O and four C−H···O hydrogen bonds that involve five different cellobiosyl residues. Its intermolecular hydrogen bonds connect molecules in adjacent [110] stacks. Network 2 connects six different cellobiose residues and involves 17 different O−H···O hydrogen bonds and seven C−H···O links. While Network 2 has a linear component, it also has a large homodromic cycle consisting of O6′B, O6′A, O6C, O2′C, O2A, and O3B which completes the link to O6′B. Another homodromic cycle adds O3A to the O6′A, O6C and O2′C atoms of the larger cycle (see also Figure 9).

Molecular Conformation. Ring Shapes. All 16 of the ring structures have normal chair conformations, listed in Table 6. The glucose rings have somewhat larger puckering amplitudes (Q) than the cyclohexyl rings on average, and the values of θ, the extent of deviation from a perfect chair form, are also larger.

Glycosidic Linkages and Geometries of Extrapolated Helices. Geometries for the linkages between the various rings are given in Table 7. Both heavy-atom and hydrogen-based definitions are given for the same torsion angles. They are generally in line with an assumption of tetrahedral geometry requiring that $\phi_H = \phi_{O5} + 120°$ and $\psi_H = \psi_{C5} + 120°$. The conformations for the glucose−glucose linkages are quite similar to linkages in cellulose itself and in some of the other small molecules including dimethyl cellobioside. Figure 12 shows the linkage geometries from **8** plotted on a conformational energy surface for cellobiose.[41] Energies were calculated for 181 different combinations of exocyclic group orientations or some other, smaller variations, using energy minimization with HF/6-31G(d) electronic structure theory. Besides values of ϕ and ψ from **8**,

(41) French, A. D.; Johnson, G. P. *Can. J. Chem.* **2006**, *84*, 603–612.

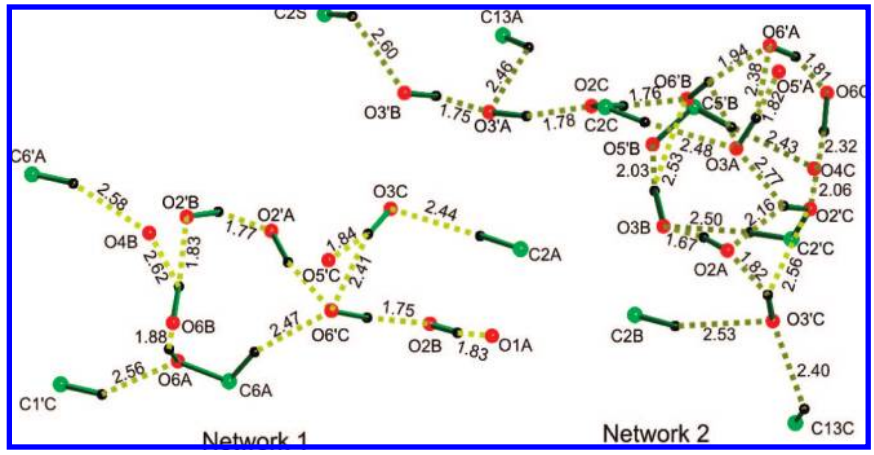


Figure 11. Hydrogen bond networks in **8**. An intramolecular hydrogen bond from O3C to O4C (see Figures 8 and 9) that exceeds the distance criterion would, given appropriate symmetry operations, connect these two networks. The same is true for a very weak bond between O2B and O2A. C−H···O hydrogen bonds are indicated by their green carbon atoms as well as the C atom labels.

Table 6. Puckering Parameters for the Rings

ring label	molecule	note	Φ (deg)	θ (deg)	Q (Å)
A	glucosyl	reducing	49	2.6	0.583
A′	glucosyl	nonreducing	50	6.8	0.602
B	glucosyl	reducing	312	10.1	0.592
B′	glucosyl	nonreducing	338	3.3	0.598
C	glucosyl	reducing	2	6.9	0.573
C′	glucosyl	nonreducing	51	11.6	0.606
		average for glucose rings		6.9	0.592
A	cyclohexyl	C7A−C12A	157	178.7[a]	0.581
A′	cyclohexyl	C13A−C18A	260	179.0[a]	0.561
B	cyclohexyl	C7B−C12B	199	178.0[a]	0.584
B′	cyclohexyl	C13B−C18B, disordered	320	178.9[a]	0.572
C	cyclohexyl	C7C−C12C	234	179.0[a]	0.579
C′	cyclohexyl	C13C−C18C, disordered, axial	140	2.3	0.569
E	cyclohexyl	C13E-C18E, minor	219	179.0[a]	0.563
F	cyclohexyl	C13F−C18F, minor, axial	151	2.6	0.574
S	cyclohexyl	C1S−C6S separate molecule	252	1.7	0.561
T	cyclohexyl	C1T−C6T separate molecule	259	178.7[a]	0.562
		average for cyclohexanes		1.5[b]	0.571

[a] Values for θ of 0° and 180° for cyclohexane are equivalent. The reported values are from the PLATON software for the structure of **8**. [b] If θ was >90°, the value used for the average was 180° − θ.

Figure 12 also shows points for other cellulose analogs. The A, B, and C conformations of the central cellobiose residues of **8** are intermingled with the conformations from cellotetraose and methyl cellotrioside, close to the dashed diagonal line representing approximate 2-fold screw-axis pseudo symmetry.

One way to utilize conformational information from small molecules is to extrapolate from their linkage and residue geometries to obtain the parameters for helical polymers.[13] Helix parameters for cellulose molecules from that procedure are given in Table 8. The linkage in the B molecule corresponds to a slightly right-handed (2.11 residues per turn) helix, while linkages in the A and C molecules correspond to almost perfect models of 2-fold screw axis symmetry in cellulose chains.

While the central cellobiose linkages have nearly 2-fold screw symmetry, similar to those in cellotetraose, the conformations of the linkages to the cyclohexyl groups deviate substantially from pseudo 2-fold screw symmetry. The closeness to 2-fold symmetry can be roughly estimated by first adding the ϕ and ψ values for a given linkage and subtracting −240° (if defined by O5′ and C5 atoms) or 0° (if H1′ and H4 are in the definitions). The distance of any point to that line is the remainder divided by 2 and multiplied times $\sqrt{2}$. Thus, the distance in ϕ,ψ space for the most distorted linkage in cellotetraose = 9° × 1.414 /2 or about 6.4°, with the other deviations being 0.6° to 3.6°. Distances for **8** are 1.4°, 9.4°, and 3.3°, for A, B, and C. The linkages to the cyclohexyl rings of **8** have a minimum deviation of 13.1° (CR in Figure 12) and a maximum of 46.5° (BR). The two maximum deviations are for the reducing-end linkages of the A and B molecules. They can have less-negative values of ψ_{C12} because of the lack of a primary alcohol group on the cyclohexyl substituent. The xylobiosyl moiety in crystalline *O*-(4-*O*-methyl-α-D-glucopyranosyluronic acid)-(1→2)-*O*-β-D-xylopyranosyl-(1→4)-D-xylopyranose trihydrate, which also lacks hydroxymethyl groups, has a comparable conformation (−79.6°, −81.9°).[42] Quantum mechanical energy surfaces for the dimer of tetrahydropyran show these two conformations within the 1 kcal/mol contour, whereas the surface for an analogous dimer of 5-methyl tetrahydropyran shows such conformations as having energies of 2−4 kcal/mol, depending on the quantum method and whether molecule A or B of the xylobiose is specified.[43]

The linkages at the nonreducing ends of the A and B molecules (AN and BN in Figure 12) differ substantially from inter-residue linkages in crystals of small analogs of cellulose. They also correspond to slightly elevated energies on energy maps for either cellobiose or the tetrahydropyran analogs presumably because there is no exoanomeric effect to guide the ϕ(C18−C13−O4′−C4′) torsion angle.

In the case of the nonreducing end of molecule C, the axial linkage to the cyclohexyl group makes comparisons with maltose geometry more appropriate. Because of the disorder of the cyclohexyl group at this position, the coordinates may not be well-determined, as indicated by the suspiciously high and low values of the glycosidic bond angle, major component value of 127°, minor component 98°. Their torsion angles are most similar to the nonflipping linkages in permethylated γ-cyclodextrin.[44,45] No inter-residue hydrogen bonding is present in that molecule.

Bond Lengths. Table 9 shows the data for the backbone C−O bonds from this reasonably high-quality determination. As expected based on the anomeric effect, the C1−O1 bond lengths

(42) Moran, R. A.; Richards, G. F. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1973**, *29*, 2770–2783.
(43) French, A. D.; Johnson, G. P. *Cellulose* **2004**, *11*, 449–462.
(44) Aree, T.; Uson, I.; Schulz, B.; Reck, G.; Hoier, H.; Sheldrick, G. M.; Saenger, W. *J. Am. Chem. Soc.* **1999**, *121*, 3321–3327.
(45) French, A. D.; Johnson, G. P. *Carbohydr. Res.* **2007**, *342*, 1223–1237.

Table 7. Linkage Geometries (Major Components Only) (deg)

linkage					
central	ϕ (O5′−C1′−O4−C4)	ψ (C5−C4−O4−C1′)	ϕ (H1′−C1′−O4−C4)	ψ (H4−C4−O4−C1′)	τ (C1′−O4−C4)
A	−89.1	−152.9	31.3	−34.4	115.2
B	−101.7	−151.6	18.4	−32.7	116.1
C	−92.0	−152.7	29.3	−35.1	115.5
reducing end	ϕ (O5−C1−O1−C7)	ψ (C12−C7−O1−C1)	ϕ (H1−C1−O1−C7)	ψ (H7−C7−O1−C1)	τ (C1−O1−C7)
A	−92.5	−85.5	27.2	34.4	114.8
B	−82.4	−91.8	36.9	28	113.2
C	−101.6	−156.9	19.3	−39.5	116.3
nonreducing	ϕ (C18−C13−O4′−4′)	ψ (C5′−C4′−O4′−C13)	ϕ (H13−C13−O4′−C4′)	ψ (H4′−C4′−O4′−C13)	τ (C13−O4′−C4′)
A	−132.7	−154.6	−13.8	−35.8	116.7
B	−146.6	−131.9	−30.1	−12.1	113.6
C	(C14−C13−O4′-C4′) 61.1[a]	−176.3	−57.7[a]	−59.3	127.4

[a] Axial linkage.

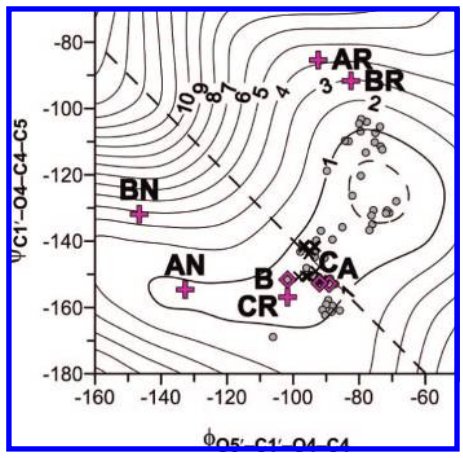


Figure 12. Partial HF/6-31G(d) energy surface for cellobiose with experimental structures indicated by markers.[40] Gray dots correspond to other small analogs of cellulose such as lactose or cellobiose. The 14 × marks, many overlapped and hidden, correspond to linkages from methyl cellotrioside[28] and cellotetraose.[19] Magenta ◆ with A, B, and C labels are for the interglucose linkages in **8**. The magenta + with AR, BR, and CR represent linkages in **8** to the cyclohexyl groups on the reducing ends of A, B, and C, respectively, and AN and BN are for linkages to the cyclohexyl groups on the nonreducing ends. The axial−equatorial CN linkage is not shown. The diagonal, dashed line shows the approximate location of structures having 2-fold screw-axis pseudo symmetry, and the circular dashed line is for 0.25 kcal/mol. Solid contour lines are at intervals of 1 kcal/mol. The contours are not necessarily relevant for the linkages to the cyclohexyl groups.

Table 8. O1−O4 Distances and Extrapolated Helix Geometries

residue	O1−O4 distance (Å)	*n* (residues/turn)	*h* (rise/residue) (Å)	helix radius (based on O4 atoms) (Å)
A (reducing)	5.474	2.02	5.15	0.93
A′ (nonreducing)	5.454	−2.01	5.11	0.95
B (reducing)	5.454	2.11	5.18	0.86
B′ (nonreducing)	5.468	2.11	5.18	0.89
C (reducing)	5.474	2.02	5.16	0.91
C′ (nonreducing)	5.440	2.02	5.11	0.94

are considerably shorter than the other C−O bonds in this sequence. The O5−C1 bonds are just a little shorter on average than the C5−O5 and aglycon distances. We also have included the C5−C6 bond lengths that are somewhat shorter, on average, than other C−C distances in **8** or the cyclohexanes.

Conclusions

The synthetic addition of nonpolar cyclohexyl groups to cellobiose resulted in a crystallizable compound that has numerous unusual features. Whether there is a practical use for a material that has a broad and fairly high melting point and planar, nonpolar surfaces is as yet unknown. Our first response was that it has a reverse amphipathic structure, with the polar groups buried in the middle of the individual layers and with nonpolar head groups. Unlike lipid bilayers, however, the layer is spanned by complete molecules, depending on the nonbonded forces only for cohesion in directions perpendicular to the molecular long axes.

Table 9. C−O Lengths in Ring and Interring Bonds, C5−C6 Lengths (Å)

molecule	C5−O5	O5−C1	C1−O1(4)	O1(4)−C7	C5−C6
A (reducing)	1.4309	1.4193	1.3929	1.4494	1.5125
A′ (nonreducing)	1.4374	1.4275	1.3958	1.4411	1.5119
B (reducing)	1.4338	1.4275	1.3890	1.4396	1.5220
B′ (nonreducing)	1.4406	1.4217	1.3975	1.4322	1.5100
C (reducing)	1.4362	1.4303	1.3883	1.4432	1.5286
C′ (nonreducing)	1.4325	1.4348	1.3918	1.4441	1.5206
average	1.4352	1.4269	1.3926	1.4416	1.5176

One way to judge the relative degree of influence of the hydrogen-bonding and nonpolar interactions is to examine which aspects of cellulosic materials are retained and which are fundamentally altered by the presence of the nonpolar cyclohexyl moieties. The arrangement of the molecules in N-shaped groups is a major indicator of the influence of the nonpolar groups on the structure. The four corners of the N correspond to stacks of cyclohexyl rings from alternating sources. Relative to the mutually antiparallel, flanking C and B molecules, the A molecule is inclined so that its reducing end cyclohexanes can stack with the cyclohexanes of B, and A's nonreducing end cyclohexane can stack with the reducing end cyclohexane of C. These molecules have found it advantageous to pack with all three of the reducing end cyclohexanes interacting with each other, while two of the three nonreducing end groups are instead stacked with solvent. This leads to the unusual "down-down-up" molecular orientation.

Despite the presence of the nonpolar groups, there is no major loss of hydrogen bonding compared to other carbohydrates. Besides the 28 conventional O−H···O hydrogen bonds, 14 C−H···O attractions were identified. However, some of the intermolecular hydrogen bonds are different from those seen before, with a good example being the O3′B···O3′A, O2′B···O2′A, and O6A···O6B interactions. Also, a novel intramolecular link was identified between a *gg* O6 and O2′, as well as the first sighting of a *tg* O6 in a cellulose analog.

Some of the normal features of cellulose and its analogs were preserved. We found the retention of the slightly altered quarter stagger to be most remarkable, stabilized by 10 of the C−H···O

hydrogen bonds. The often-found, 2-fold screw-axis pseudo symmetry was also present, along with the O3···O5′ and ···O6′ hydrogen bonds. The ability to accommodate both the cyclohexane packing and a general organization of the carbohydrate in an array that is similar in broad detail to cellulose II was also remarkable.

The replacement of the two terminal glucose groups of cellotetraose with cyclohexane resulted in a molecule that had a core very much like the disaccharide core of cellotetraose, but the outer linkages were very different in their geometries from those of cellotetraose.

The crystallographic data have been deposited with the Cambridge Crystallographic Data Centre, CCDC Nos. 707928 and 707929.

Acknowledgment. The authors thank Professor Henri Chanzy, and Professor Rengaswami Chandrasekaran and Dr. Srinivas Janaswamy for helpful input on up−up-down structures. Dr. Paul Langan and Professor Ed Stevens offered comments at an early point in this study. Comments on a near-final draft were received from Professors Don Kiely and Charles Lake. This work was supported by the Austrian Science Fund (FWF, P-17426) and by normal research funding from the Agricultural Research Service, U.S. Department of Agriculture.

Supporting Information Available: Crystallographic data of **8** in CIF format; Figures S1 and S2 showing the disorder and cyclohexane packing, respectively; ^{1}H and ^{13}C NMR spectra of 8 in solution included as Figure S3. This material is available free of charge via the Internet at http://pubs.acs.org.

JA805147T