Notes

The Structure of Cellulose II: A Revisit

S. Raymond,† Å. Kvick,‡ and H. Chanzy*,†

Centre de Recherches sur les Macromolécules Végétales, CNRS, B.P. 53, 38041 Grenoble Cédex 9, France, and European Synchrotron Radiation Facility, B.P. 220, 38043 Grenoble Cédex, France

Received June 5, 1995
Revised Manuscript Received July 25, 1995

The most recent molecular and crystal structure of cellulose II was established in 1976, independently by Kolpak and Blackwell[1] and Stipanovic and Sarko.[2] Using a similar methodology, both groups took X-ray diagrams of rayon fibers and analyzed them by conventional X-ray fiber crystallography together with conformational and packing analysis. In both cases, the cellulose chain was built from the Arnott and Scott[3] standard pyranose sugar geometry and the cellulose conformation as well as the crystal packing was optimized starting from such fixed geometry. In the structure of Stipanovic and Sarko, this fixed geometry was relaxed during the refinement but was not relaxed in the structure of Kolpak and Blackwell. Both the model of Kolpak and Blackwell[1] and that of Stipanovic and Sarko[2] define cellulose II as a two-chain unit cell with antiparallel chain packing and an extensive hydrogen bond network. Another important feature of these two structures is the position of the hydroxymethyl groups that present different conformations in the two molecules of the unit cell. Following, for example, the convention of Kolpak and Blackwell all the hydroxymethyl groups of the "up" chain located at the corner of the unit cell are oriented near the *gt* position[4] whereas those of the "down" center chain are near the *tg* position. This situation implies that along the corner chain there is only one type of intramolecular hydrogen bond whereas two different ones occur along the center one.

The occurrence of two types of hydroxymethyl conformations for crystalline cellulose II has been a point of controversy among authors who have analyzed the details of the infrared spectroscopy and mechanical properties of cellulose II.[5,6] A stronger argument against the two types of hydroxymethyl conformation comes from the analysis of the CP/MAS ^{13}C NMR solid spectra of cellulose II. In these spectra, the C6 resonance occurs only as a singlet near 64 ppm[7−10] as opposed to the expected doublet with a resonance band near 64 and another near 66 if both *tg* and *gt* had been present within the crystalline structure.[11]

Very recently, the crystal and molecular structures of β-D-cellotetraose[12,13] and methyl β-D-cellotrioside[14] have been determined. As both crystalline cellodextrins display diffraction patterns and ^{13}C solid state NMR spectra that are quite similar to those of cellulose II,[15] it is expected that their three-dimensional molecular structure should be a good model for cellulose II. In both cellotetraose and methyl β-D-cellotrioside crystals, the molecules are organized in pairs of antiparallel residues where all the hydroxymethyl groups are in the *gt* situation. Another important common feature of these structures is that the Cremer and Pople[16] puckering parameters of their glucosyl moieties are roughly the same within any molecule of a given pair but differ significantly between the "up" molecule and the "down" molecule. Thus in the crystals of cellotetraose as well as those of methyl β-D-cellotrioside, there are two types of molecules: the A molecules for which these parameters are considered to be "standard" with a Cremer and Pople Θ value close to 0° and the B molecules for which these parameters reflect a strained situation with a Θ value of around 10°.[12−14] The difference between the A and B molecules of these cellodextrins is also evidenced by comparing the torsion angles Φ and Ψ[17] that define the geometry of the linkage connecting successive glycosyl residues within a given structure. Whereas the values of the Φ angle for the internal glycosyl residues of a given cellodextrin molecule are within 1° for the A and B molecules, there is a difference of about 10° for the Ψ angles between the corresponding residues in molecule A and molecule B.[12−14]

Table 1. Comparison of the Cell Parameters of Cellulose II with Those of Model Cellodextrins[a]

structure	*a* (Å)	*b* (Å)	*c* (Å)	γ (deg)	ref
cellulose II	7.95	9.17	10.34	117.2	Wellard[23]
cellulose II	8.01	9.04	10.36	117.1	Kolpak−Blackwell[1]
cellulose II	7.96	9.09	10.31	117.3	Stipanovic−Sarko[2]
β-D-cellotetraose	8.026	9.03		116.96	Gessler *et al.*[12]
β-D-cellotetraose	8.045	9.003	10.386(14)[b]	115.8	Raymond *et al.*[13]
methyl β-D-cellotrioside	7.998	8.991	10.381(10)[c]	116.40	Raymond *et al.*[14]
average	8.00(4)	9.05(7)	10.36(3)	116.8(6)	
cellulose II	8.00	9.05	10.38	116.8	this work

[a] The cell parameters were transposed in order to be consistent from one structure to the next. [b] If, following the Joint Commission on Biochemical Nomenclature,[24] the glucosyl residues are labeled 1−4 starting from the reducing end, this value corresponds to the average of the center to center distance between the sugar rings 1 and 3 and 2 and 4 for both independent cellotetraose molecules. [c] This value corresponds to the average distance between the centers of the two sugar rings located at both ends of each independent molecule.

The striking similarity of unit cell parameters *a*, *b*, and γ of cellulose II with those of cellotetraose and methyl β-D-cellotrioside is shown in Table 1. Since this match is almost perfect, the lateral packing of cellulose II must be nearly the same as that in these cellodextrins. Along the chain direction, the *c* axis periodicity of cellulose II is given as 10.31 Å by Stipanovic and Sarko[2] and 10.36 Å by Kolpak and Blackwell.[1] This parameter, which corresponds to the length of a cellobiosyl residue is measured as 10.38 Å in both cellotetraose as well as methyl β-D-cellotrioside. Since this unit-cell constant appeared critical to any modification of the structure of cellulose II, it was remeasured. The

† Centre de Recherches sur les Macromolécules Végétales.
‡ European Synchrotron Radiation Facility.

0024-9297/95/2228-8422$09.00/0

Table 2. Fractional Atomic Coordinates of the Non-Hydrogen Atoms for One Glucose Residue of Each Chain of the Cellulose II Model

atom	x	y	z
	Corner Chain[a] (0, 0, 0)		
C1	−0.043 97	0.009 50	0.384 29
C2	−0.125 47	0.085 56	0.288 94
C3	−0.157 65	−0.007 48	0.160 20
C4	0.026 73	−0.001 14	0.115 67
C5	0.101 68	−0.074 74	0.218 31
C6	0.290 43	−0.065 65	0.184 49
O2	−0.299 51	0.070 78	0.335 78
O3	−0.234 32	0.063 46	0.071 81
O4	0.003 84	−0.094 48	0.000 05
O5	0.126 04	0.016 66	0.336 06
O6	0.344 24	−0.144 41	0.281 57
	Center Chain[b] (1/2, 1/2, 0)		
C1	0.540 89	0.477 59	0.349 05
C2	0.687 36	0.491 55	0.445 70
C3	0.607 79	0.443 00	0.582 30
C4	0.483 45	0.522 90	0.616 78
C5	0.344 94	0.504 05	0.510 73
C6	0.233 36	0.597 10	0.534 96
O2	0.770 90	0.388 78	0.405 34
O3	0.762 87	0.491 00	0.666 62
O4	0.373 67	0.453 77	0.731 73
O5	0.451 38	0.573 37	0.394 79
O6	0.127 19	0.596 18	0.426 02

[a] The corner chain corresponds to the A molecules in the cellodextrin crystals. [b] The center chain corresponds to the B molecules in the cellodextrin crystals.

experiment was carried out at room temperature on the 2(ID11) beam line at ESRF, Grenoble, France.[18] Monochromatic radiation with a wavelength of 0.6199 Å was used, and the data were recorded on a thin bundle of well-aligned Fortisan fibers mounted on the head of a goniometer. The sample was rotated from 0 to 9° in 1° steps around an axis perpendicular to both the sample fiber axis and the incident beam direction. The diagrams were recorded on image plate detectors positioned at a sample to plate distance of 303 mm. The image plates were scanned with a Molecular Dynamics Phosphor Imager 400E. Calibration was done with a crystalline BN needle that gave sharp diffraction lines at 2.087 and 1.808 Å. Under those conditions, the center of the 002 and 004 reflections of cellulose II could be precisely measured, giving the c parameter value of 10.38 ± 0.03 Å that was used in this paper. In the image plates, we could not detect the presence of 001 and 003 reflections. Their absence is an indication that the cellulose II structure has a 2-fold screw axis along its fiber direction, and therefore its most probable space group is $P2_1$.

A cellulose II model was built using the two different types of chains: A and B. The A molecule was generated by averaging the geometry of the two internal glucosyl residues of the A molecule of cellotetraose and the middle glucosyl residue of the two independent A molecules in methyl β-D-cellotrioside. These four glucosyl residues together with their torsion angles Φ and Ψ were averaged by superposition by using a molecular similarity program.[19] The A chain of cellulose was then propagated from the averaged glucosyl residue. A similar approach was used for building the B chain.

A cellulose II model was then made by positioning the two antiparallel cellulose molecules A and B on the 2-fold screw axes of the average unit cell presented in Table 1 with cell parameters a = 8.00 Å, b = 9.05 Å, c = 10.38 Å, and γ = 116.8° and the $P2_1$ space group. This model was refined against the X-ray data listed by Kolpak and Blackwell,[1] using the linked approach least square (LALS) program.[20] Only three variables were used: the translation of the two chains with respect to one another and the independent rotation of the two chains on their 2-fold screw axes. All attempts to refine further the model by stepping the torsion angles Φ and Ψ as well as the rotation angles χ and χ' of the hydroxymethyl group of the individual chains were unsuccessful since it led systematically to increased crystallographic residual R together with some unacceptable short contacts. Given these conditions, the optimized cellulose II structure was refined to an R factor of 0.196[21] by using the observed data from the set of intensities given by Kolpak and Blackwell.[1] This value became R'' = 0.164 by using all the data set with w = 1 for the observed reflections and 0.5 for the unobserved. The list of the coordinates of our structure is given in Table 2 whereas its conformational characteristics are presented in Table 3. This table compares also our conformational parameters with those of Kolpak/Blackwell and Stipanovic/Sarko. The list of the shortest oxygen to oxygen distances is given in Table 4. Figure 1 presents a stereopair of the packing of our cellulose II structure in its unit cell. In Figure 2, a superposition of the A and B molecules illustrates the differences in their conformation.

A comparison of our structure and that of Kolpak and Blackwell is shown in Figure 3. In a projection perpendicular to the chain axis, the two structures are indistinguishable (Figure 3A). In a projection including the chain direction, the two structures are nearly identical for the corner chains (Figure 3B) but differ markedly for the center chains (Figure 3C). In this case, the difference is not only in the position of the hydroxymethyl moieties that are in the *gt* position in our structure as opposed to *tg* in that of Kolpak and Blackwell but also in the translation of the center chain with respect to the corner chain. In our structure the difference in z coordinates between the center of a glucosyl residue of a corner chain and that of a center

Table 3. Selected Geometrical Parameters of the Two Glucose Residues Constituting the Two Chains of Cellulose II: A Comparison between Our Model (Chains A and B) and Those Calculated from the Coordinates of the Cellulose II Structures of Kolpak/Blackwell[1](K/B) and Stipanovic/Sarko[2] (S/S)

	chain A corner (deg) (this work)	chain B center (deg) (this work)	K/B corner (deg)	K/B center (deg)	S/S[a] center (deg)	S/S[a] corner (deg)
τ(C1′−O4−C4)	117.7	116.5	114.8	114.8	116.8	116.8
Φ(O5−C1′−O4−C4)	−93.2	−92.0	−95.9	−96.4	−97.6	−97.6
Ψ(C1′−O4−C4−C3)	96.4	86.4	93.6	94	95.8	95.8
χ(O5−C5−C6−O6)	63.2 (*g*)	53.3 (*g*)	52.3 (*g*)	168.8 (*t*)	64.0 (*g*)	165.9 (*t*)
χ′(C4−C5−C6−O6)	−176.9 (*t*)	171.2 (*t*)	173.8 (*t*)	−69.7 (*g*)	176.2 (*t*)	−76.6 (*g*)
Θ param[b]	1.2	11.6	1.9	2.3	3.3	3.3

[a] There is an inversion in the definition of the center and corner chains between the structure of Stipanovic and Sarko[2] with that of Kolpak and Blackwell.[1] In this work, we have followed the convention of Kolpak and Blackwell. [b] Puckering parameters following the definition of Cremer and Pople.[16]

Table 4. List of Short Distances O···O Indicating Hydrogen Bondings: A Comparison between This Work and That of Kolpak and Blackwell[1] (K/B) and That of Stipanovic and Sarko[2] (S/S)

Intramolecular Short Distances O···O (Å)

	chain A corner (this work)	chain B center (this work)	K/B corner	K/B center	S/S[a] center	S/S[a] corner
O3′···O5	2.8	2.83	2.69	2.69	2.70	2.70
O6′···O2				2.73	2.76	

Intermolecular Short Distances O···O (Å) between Chains of the Same Type

	chain A/ chain A (this work)	chain B/ chain B (this work)	K/B corner/ corner	K/B center/ center	S/S[a] corner/ corner	S/S[a] center/ center
O2···O6* [b]	2.69	2.62	2.76			2.97
O3···O6* [b]				2.67	2.65	

Intermolecular Short Distances O···O (Å) between Chains of Different Types

	chain A/chain B (this work)	K/B	S/S[a]
O2···O2* [b]	2.77	2.77	2.62
O6···O6* [b]	2.66		
$O6_{tg}$···O3* [b]		3.25	2.79
$O6_{gt}$···O3* [b]		3.08	2.80

[a] There is an inversion in the definition of the center and corner chains between the structure of Stipanovic and Sarko[2] and that of Kolpak and Blackwell.[1] In this work, we have followed the convention of Kolpak and Blackwell. [b] * indicates a different chain.

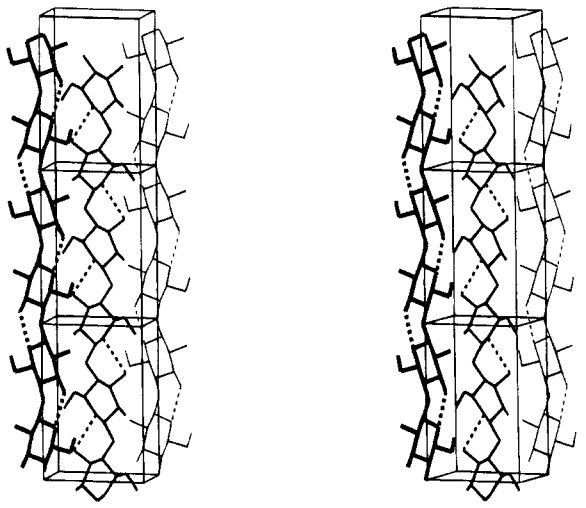

Figure 1. Stereoviews of the structure of cellulose II with its intramolecular short O···O distances (dashed lines) corresponding to hydrogen bonds. In our convention that is the same as that of Kolpak and Blackwell[1] but diffrent from that of Stipanovic and Sarko;[2] molecules A are located at the corner of the unit cell, and molecules B, at the center.

chain is 2.42 Å, very similar to the value of 2.40 Å found for the cellodextrin crystals but somewhat different from 2.21 Å in the Kolpak/Blackwell structure (2.13 Å in the structure of Stipanovic/Sarko). Thus, in our structure the corner and the center molecules are staggered by nearly half the length of a glucosyl residue.

The revised cellulose II structure that is presented in this study has the advantage of being supported by precise cellodextrin crystallographic data. Without exhibiting any short contacts our structure can be refined against the observed reflections in the data set of Kolpak and Blackwell[1] to give an R factor that is slightly higher than theirs: 0.196 as opposed to their 0.171.[22] By including in the refinement their unobserved reflections and using the same weight factors as in their work, we obtained the same R'' factor of 0.164 as they did. Another advantage of our structure is that all the hydroxymethyl moieties of both independent chains are in the *gt* conformation. This situation is consistent with the NMR spectroscopy of cellulose II. The implication of two different puckering parameters and Ψ angles for the glucosyl moieties that constitute the corner and the center chains could explain also why a splitting is observed in the C1 as well as the C4 signals of the ^{13}C solid state spectra of cellodextrins as well as those of cellulose II.[7-10]

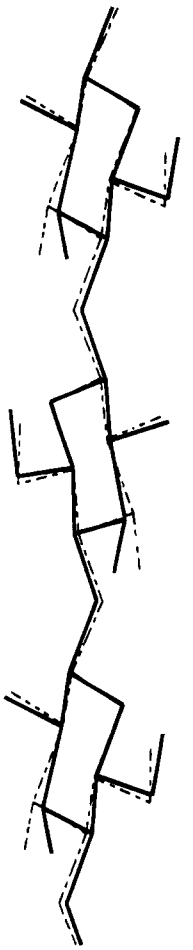

Figure 2. Superposition of molecules A and B of the structure of cellulose II. The A molecule is denoted by the continuous line whereas the B molecule is denoted by the dashed line.

A comparison of our structure with that of Stipanovic and Sarko was not made. The structure proposed by these authors was refined with a c axis parameter of 10.31 Å that is smaller than the 10.38 Å measured in this work and deduced from our cellodextrin data. In view of such discrepancy, the use of their X-ray data with our model would probably imply a re-indexation of their X-ray patterns, and this was beyond our goal. In their refinement, these authors have also relaxed the $P2_1$ symmetry to $P1$. This lower symmetry implies that there are more parameters that can be adjusted in the structure refinement. As a consequence, their residual R factor is substantially lower than the one of Kolpak and Blackwell as well as lower than ours. In the present work, relaxing the $P2_1$ symmetry has also the effect of lowering the residual R but the comparison of such a low-symmetry structure with that of Kolpak and Blackwell becomes meaningless. The distinction between space groups $P2_1$ and $P1$ may be resolved by the

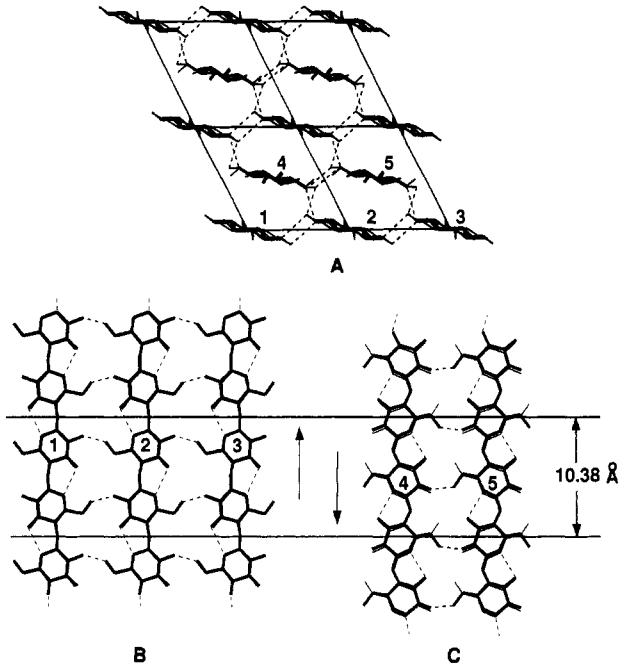

Figure 3. Superposition of our structure of cellulose II (bold line) over that of Kolpak and Blackwell[1] (thin line). The short O···O distances that correspond to hydrogen bonds are given by dashed lines. (A) Projection in the *ab* plane. (B) Projection of the 1–3 molecules parallel to the c axis. (C) Projection of molecules 4 and 5 parallel to the *c* axis.

collection of a better data set. The new possibilities given by the synchrotron machines should allow such collection, and we will attempt it in the future.

Acknowledgment. The authors are indebted to W. T. Winter for valuable suggestions during this work.

References and Notes

(1) Kolpak, F. J.; Blackwell, J. *Macromolecules* **1976**, *9*, 273.
(2) Stipanovic, A. J.; Sarko, A. *Macromolecules* **1976**, *9*, 851.
(3) Arnott, S.; Scott, W. E. *J. Chem. Soc., Perkin Trans. 2* **1972**, *2*, 324.
(4) The conformation of the hydroxymethyl group is defined by two letters, the first referring to the torsion angle χ (O5–C5–C6–O6) and the second to the torsion angle χ' (C4–C5–C6–O6). Thus an ideal *gt* conformation would be defined as the set of two angles: 60°, 180°.
(5) Marchessault, R. H.; Liang, C. Y. *J. Polym. Sci.* **1960**, *43*, 71.
(6) Kroon-Batenburg, L. M. J.; Kroon, J.; Northolt, M. G. *Polym. Commun.* **1986**, *27*, 290.
(7) Dudley, R. L.; Fyfe, C. A.; Stephenson, P. J.; Deslandes, Y.; Hamer, G. K.; Marchessault, R. H. *J. Am. Chem. Soc.* **1983**, *105*, 2469.
(8) Fyfe, C. A.; Stephenson, P. J.; Veregin, R. P.; Hamer, G.; Marchessault, R. H. *J. Carbohydr. Chem.* **1984**, *3*, 663.
(9) Isogai, A.; Usuda, M.; Kato, T.; Uryu, T.; Atalla, R. H. *Macromolecules* **1989**, *22*, 3168.
(10) Horii, F.; Hirai, A.; Kitamaru, R.; Sakurada, I. *Cellulose Chem. Technol.* **1985**, *19*, 513.
(11) Horii, F.; Hirai, A.; Kitamaru, R.; Sakurada, I. *Polym. Bull.* **1983**, *10*, 357.
(12) Gessler, K.; Krauss, N.; Steiner, T.; Betzel, C.; Sandmann, C.; Saenger, W. *Science* **1994**, *266*, 1027.
(13) Raymond, S.; Heyraud, A.; Tran Qui, D.; Kvick, Å.; Chanzy, H. *Macromolecules* **1995**, *28*, 2096.
(14) Raymond, S.; Henrissat, B.; Tran Qui, D.; Kvick, Å.; Chanzy, H. *Carbohydr. Res.*, in press.
(15) Henrissat, B.; Pérez, S.; Tvaroska, I.; Winter, W. T. In *The Structure of Cellulose. Characterization of the Solid States*; Atalla, R. H., Ed.; ACS Symposium Series 340; American Chemical Society: Washington, DC, 1987; pp 38–67.
(16) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354.
(17) The torsion angles Φ and Ψ are defined as Φ ($O5_i$–$C1_i$–$O4_{i-1}$–$C4_{i-1}$) and Ψ (C_i–$O4_{i-1}$–$C4_{i-1}$–$C3_{i-1}$).
(18) Kvick, Å.; Wulff, M. *Rev. Sci. Instrum.* **1992**, *63*, 1073.
(19) From the Quanta 4.0 package, Molecular Simulation Inc., The University of York, York, England.
(20) Smith, P. J. C.; Arnott, S. *Acta Crystallogr.* **1978**, *A34*, 3.
(21) The crystallographic residuals R and R'' are calculated as follows: $R = \sum||kF_o| - |F_c||/\sum|kF_o|$; $R'' = \{\sum w(|kF_o| - |F_c|)^2]/\sum w(kF_o)^2\}^{1/2}$, where k is the scale factor.
(22) The R and R'' factors of the structure of Kolpak and Blackwell[1] were recalculated as in ref 21.
(23) Wellard, H. J. *J. Polym. Sci.* **1954**, *13*, 471.
(24) IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). *Eur. J. Biochem.* **1983**, *131*, 5.

MA950793N