X-ray Structure of Mercerized Cellulose II at 1 Å Resolution

ARTICLE *in* BIOMACROMOLECULES · JUNE 2001

Impact Factor: 5.75 · DOI: 10.1021/bm005612q · Source: PubMed

3 AUTHORS, INCLUDING:

Yoshiharu Nishiyama
French National Centre for Scientific Research
100 PUBLICATIONS **6,231** CITATIONS

X-ray Structure of Mercerized Cellulose II at 1 Å Resolution

Paul Langan,*,† Yoshiharu Nishiyama,‡ and Henri Chanzy§

Biosciences Division, M888, Los Alamos National Laboratory, Los Alamos, New Mexico 87545; Department of Biomaterials Science, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo 113-8657, Japan; and Centre de Recherches sur les Macromolécules Végétales-CNRS, B.P. 53, 38041 Grenoble Cedex 9, France

Received September 18, 2000; Revised Manuscript Received February 8, 2001

A revised crystal structure for mercerized cellulose based on high-resolution synchrotron X-ray data collected from ramie fibers is reported (space group $P2_1$, $a = 8.10(3)$ Å, $b = 9.03(3)$ Å, $c = 10.31(5)$ Å, $\gamma = 117.10(5)°$; 751 reflections in 304 composite spots; $\theta < 21.11°$; $\lambda = 0.7208$ Å; LALS refinement with $d > 1.5$ Å, $R'' = 0.16$; SHELX97 refinement with $d > 1$ Å, $R = 0.21$). As with regenerated cellulose the crystal structure consists of antiparallel chains with different conformations but with the hydroxymethyl groups of both chains near the *gt* position. However, the conformation of the hydroxymethyl group of the center chain in the structure reported here differs significantly from the conformation in regenerated cellulose. This may be related to a large observed difference in the amount of hydroxymethyl group disorder: ∼30% for regenerated cellulose and ∼10% for mercerized cellulose.

Introduction

A linear poly(1−4)-β-D-glucan, cellulose, is normally biosynthesized into a metastable form, known as cellulose I, which consists of slender rodlike crystalline microfibrils. These microfibrils are composed of species specific compositional ratios of two crystalline allomorphs, designated cellulose Iα and cellulose Iβ.[1,2] The unit cells of these allomorphs have recently been determined by electron diffraction,[3] and a combination of biochemical techniques and electron microscopy has shown that the cellulose chains pack in a parallel arrangement,[4] although the detailed crystal structures are still under investigation.[5,6] Cellulose I can be made to undergo an irreversible transition to a stable crystalline form, cellulose II, by two distinct processes: regeneration and mercerization. Regeneration involves either preparing a solution of cellulose in an appropriate solvent or preparing of an intermediate derivative followed by coagulation and recrystallization. This process is used to produce rayon fibers. Mercerization involves intracrystalline swelling of cellulose in concentrated aqueous NaOH followed by washing and recrystallization. This process is used to improve the properties of natural yarns and fabrics.

The crystal structure of regenerated cellulose was initially determined independently by Kolpak and Blackwell[7] and Stipanovik and Sarko[8] in X-ray diffraction studies of Fortisan rayon fibers. Two antiparallel and crystallographically independent chains align along the 2-fold axes of the $P2_1$ unit cell. Both chains have equivalent backbone and sugar conformations but differ in the conformation of their primary alcohol (commonly called hydroxymethyl groups); *gt* for the chain located at the cell origin and *tg* for the center chain.[9,10] During refinement of these models, the conformations of the sugar rings were held rigid. More recent model building studies incorporating flexible sugar rings have shown that the same X-ray data can be equally interpreted as chains with different backbone and sugar conformations but with similar hydroxymethyl group conformations.[11−13] Although the X-ray diffraction data are not sufficient to allow these two models to be differentiated, neutron fiber diffraction data have allowed the original *gt*/*tg* model to be rejected in favor of the revised *gt*/*gt* model.[13]

The conformational features of the revised model are in good agreement with the X-ray diffraction single-crystal structures of two cellulose oligomers, β-cellotetraose and methyl β-cellotrioside.[14−16] Although the sugar of the origin chain is conformationally unstrained, the sugar of the center chain is conformationally strained. The relative orientation of adjacent glycosyl residues, described by the glycosidic torsion angles, ϕ and ψ is also different for the center and origin chains.[17] The hydroxymethyl groups of both chains are near the *gt* conformation, in agreement with observations resulting from a number of ^{13}C NMR studies. In the cellulose II spectra, the C6 resonance occurs as a singlet near 64 ppm and not as the expect doublet with resonances near 64 and 66 ppm if both *gt* and *tg* conformations coexist in the crystalline structure.[18−22]

The crystal structure of mercerized cellulose was initially determined by Kolpak et al.[23] in X-ray diffraction studies of cotton fibers. The molecular and crystal structure is essentially identical to that initially determined for regenerated Fortisan rayon fibers.[7,8] There was some indication that the hydroxymethyl groups were disordered and that this disorder is different for regenerated and mercerized cellulose. Given the topographical problems involved in converting from a parallel chain cellulose I to an antiparallel chain cellulose II without dissolving the chains, it has also been proposed that

† Los Alamos National Laboratory.
‡ The University of Tokyo.
§ Centre de Recherches sur les Macromolécules Végétales-CNRS.

10.1021/bm005612q CCC: $20.00
Published on Web 04/18/2001

mercerized cellulose has a parallel structure,[24,25] different from that of regenerated cellulose. Clearly, the crystal structure of mercerized cellulose has to be reexamined in light of the revisions to the structure of regenerated cellulose and in light of recently reported parallel chain models.

In this work, we present high-resolution synchrotron X-ray data collected from mercerized ramie fibers. These data have allowed us to determine a revised model for mercerized cellulose. This model is compared to the revised model of regenerated cellulose, and the differences are discussed.

Sample Preparation

A bundle of purified ramie fibers was gently combed and fixed to a stainless steel stretching device. The fibers were slightly relaxed and then swollen in 20% NaOH solution at room temperature overnight, followed by washing in distilled water. The sample was successively treated with 12 % NaOH and washed with water at 0 °C, five times. Washing at 0 °C involved putting the alkali-swollen sample into ice water. After about 6 h, the water was changed and heated to 80 °C to remove any remaining alkali bound to the fiber. The shrunken fibers became slightly relaxed during washing so they were finally stretched in order to improve the orientation.

Data Collection

Data were collected using an online MAR image plate and a wavelength of 0.7208A on instrument ID13 at the ESRF, Grenoble, France. The fiber tilt and rotation parameters were determined using the program XFIX, part of the BBSRC funded CCP13 project, Daresbury, U.K., and the data were then transformed into reciprocal space. A 2D fit of the background and Bragg intensities in reciprocal space was made using the CCP13 program LSQINT. The background was fitted using the "roving aperture" method. A Lorentzian distribution was used to model the angular profile of the spots. The angular width of this distribution along with the particle length and width were refined together with the unit cell parameters during maximum entropy refinement. The refined unit cell parameters, $a = 8.10(3)$ Å, $b = 9.03(3)$ Å, $c = 10.31(5)$ Å, and $\gamma = 117.10(5)°$ are slightly different from those reported for regenerated cellulose ($a = 8.01(5)$ Å, $b = 9.04(5)$ Å, $c = 10.36(3)$ Å, and $\gamma = 117.1(1)°$).[7] However, to avoid singularities during matrix inversion, the positions of closely spaced reflections are made equal in LSQINT, and this may affect the accuracy of the unit cell values. The data collection parameters are summarized in Table 1. A comparison of the fit to the data is shown in Figure 1.

At the end of refinement a standard deviation, σ, was calculated for each reflection directly from the root-mean-square difference between the calculated and observed pixel intensity values over the predicted profile of the spot. The Bragg data consisted of 751 reflections corresponding to the sum of 1497 symmetry overlapped reflections. Accidental overlap due to cylindrical averaging reduced the number of independent intensities to 304. Of these intensities, I, 292 had $I > \sigma$ with $R_\sigma = 0.18$ and 126 had $I > 10\sigma$ with $R_\sigma = 0.06$, where R_σ is defined as the average value of σ/I over all reflections concerned. Figure 2 shows the mean intensity, $\langle I \rangle$, and R_σ calculated for diffraction spots grouped in resolution shells. While $\langle I \rangle$ decreases with resolution, R_σ increases, indicating that at high resolution the spot intensities are weaker and have greater associated uncertainties.

A weak equatorial reflection that could not be indexed

Table 1. Data Collection Parameters[a]

wavelength (Å)	0.7208
space group	$P2_1$
a (Å)	8.10(1)
b (Å)	9.03(1)
c (Å)	10.31(1)
γ (deg)	117.10(5)
radiation type	synchrotron X-ray
no. of reflections	1497
no. of indep reflections	751
no. of indep intens spots	304
no. of indep intens spots with $I > \sigma$	292
R_σ for $I > \sigma$	0.18
no. of indep intens spots with $I > 4\sigma$	232
R_σ for $I > 4\sigma$	0.11
no. of indep intens spots with $I > 10\sigma$	126
R_σ for $I > 10\sigma$	0.06
diffractometer	ID13
range of h, k, l	$0 \rightarrow h \rightarrow 8$, $-9 \rightarrow k \rightarrow 8$, $0 \rightarrow l \rightarrow 10$
θ_{min} (deg)	2.57
θ_{max} (deg)	21.11

[a] Medium resolution refinement: LALS. $R'' = 0.16$; 277 data including 95 reflections, 150 applied contact terms, and 19 net degrees of freedom. High-resolution refinement: SHELX97. $R = 0.211$; 318 data including 235 reflections and 83 restraints and 90 refined parameters.

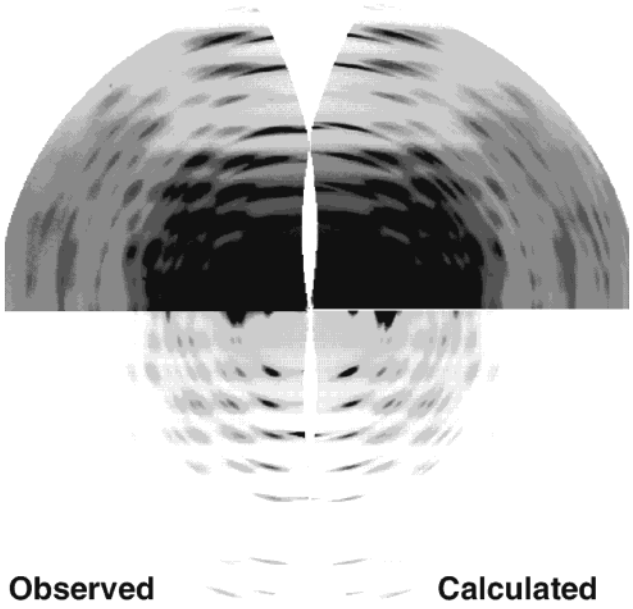

Figure 1. Left-hand side: Synchrotron X-ray diffraction data collected on an online MAR image plate from mercerized ramie fibers on station ID13 at the ESRF, Grenoble, France. Right-hand side: A 2D fit of the background and Bragg intensities using program LSQINT, part of the BBSRC-funded CCP13 project, Daresbury, U.K. The images have been remapped into cylindrical reciprocal space with the fiber axis vertical and reproduced at low display thresholds (top) and high display thresholds (bottom) so that the full dynamic range of reflections can be seen. The data collection parameters are summerized in Table 1.

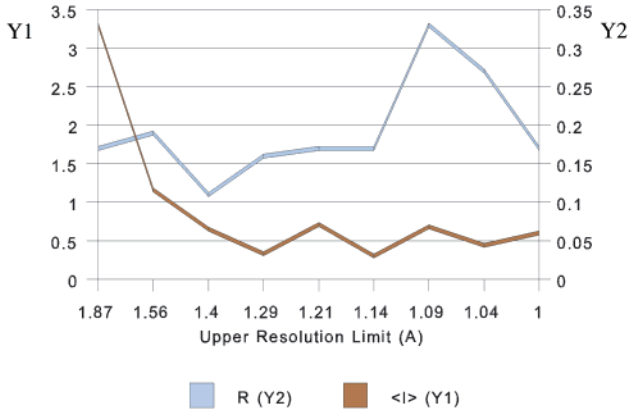

Figure 2. Mean intensity, $\langle I \rangle$, and R_σ, calculated for diffraction spots grouped in resolution shells. Each resolution shell contains approximately 30 reflections. R_σ is defined in the text.

using the unit cell of cellulose II but that could be indexed using the unit cell of cellulose I, was present in the diffraction data. This reflection was present even with samples that were treated with up to 32% NaOH and probably indicates a persistent cellulose I component. From the relative intensity of this reflection, we estimate that less than 1% of the crystalline material in the sample corresponds to cellulose I. Any accidental overlap of reflections from the cellulose I and II components would not be expected to greatly influence the structure analysis, because of the relatively small amount of cellulose I present. The crystalline cellulose II regions within the sample may also have varying degrees of disorder or inhomogeneous disorder. The diffraction data reported here correspond to the statistical average of all cellulose II crystalline regions.

Structure Solution

Structure solution proceeded in two steps. First, 2-fold helical models were built and refined against the X-ray data extending to *d* spacings >1.5 Å, using the linked atom-least-squares (LALS) procedure. [26] The LALS procedure is well suited to building starting models with different conformational features. However because bond lengths are fixed during refinement and only a global temperature factor can be refined it is inappropriate to use data with *d* spacings <1.5 Å. Hamilton's statistical test [27] was used to determine the significance of changes in agreement with the data as represented by R'', [28] after introducing or removing parameters. Over short contacts were relieved by applying a constraint term in the minimization function. Hydrogen atoms were explicitly included when covalently bound to carbon atoms but not to oxygen so that hydrogen bonding was not initially incorporated into the refinement. Starting models were refined with various net degrees of freedom as previously described.[13]

The first model, designated A, was built from the previously published coordinates of mercerized cellulose.[23] The second, designated B, was identical to A except the all hydroxymethyl groups were in the *gt* conformation. Models A and B were refined with fixed sugar conformations and 7 net degrees of freedom to give values of 0.251 and 0.192 for R'', respectively. Allowing either the sugar pucker of the origin chain or the center chain to refine involved 13 net degrees of freedom and gave R'' values of 0.213 and 0.218 for model A and R'' values of 0.169 and 0.179 for model B, respectively. Allowing the sugar puckers of both chains to refine involved 19 net degrees of freedom and resulted in R'' values of 0.206 and 0.160 for models A and B, respectively. All models could be rejected with respect to model B with the sugar pucker of both chains allowed to refine, at a confidence level of 95% or better.

Refinements with 19 net degrees of freedom and both chains either parallel up or parallel down [10] resulted in R'' values of 0.271 and 0.275, respectively, when the hydroxymethyl groups of the origin and center chains were in the *gt* and *tg* conformations, respectively. Similarly, with both chains either parallel up or parallel down the values of R'' were 0.329 and 0.315 when the hydroxymethyl groups of both chains were in the *gt* conformation. All parallel chain models could be rejected with respect to model B at a confidence level of greater than 99.5%.

Model B was then further refined against all of the data using SHELX97.[29] At first only the scale factor and an overall temperature factor were refined against data extending to a resolution of 1.5A. Two residuals were monitored during refinement, each calculated from $\sum(|F_o| - |F_c|)/\Sigma F_o|$ where F_c and F_o are the calculated and observed amplitudes, respectively. The first, designated R_{free}, was calculated from 10% of the reflections which were excluded from the course of refinement. The second, designated R, was calculated from the remaining 90% of the reflections with $F_o > 4\sigma$. The resultant values for R and R_{free} were 0.224 and 0.172. The position parameters of all atoms were then allowed to refine with an overall temperature factor. The coordinates of hydrogen atoms were allowed to ride on the coordinates of the covalently bonded atoms. Hydrogen atoms had thermal parameters tied to 1.5 and 1.2 times the thermal parameters of the covalently bound oxygen or carbon atoms, respectively. The initial orientations of hydroxyl groups were estimated using the AFIX algorithm with $m = 8$ in SHELX97. All bond lengths and bond angles were restrained using the DFIX and DANG options in SHELX97 with default standard deviations. The resultant values for R and R_{free} were 0.176 and 0.117.

Incorporating the remaining reflections between 1.5 and 1.0 Å introduced systematic deviations from a flat analysis of variance. Both the scale factor, K, defined as the mean F_o^2 over the mean F_c^2, and the crystallographic residual, R, start to rise at above a resolution of 1.5Å when calculated for reflections grouped in resolution shells; see Figures 3 and 4. This indicates that, on average, the measured intensities are too large at high resolution. When K is calculated for reflections grouped in shells of decreasing $F_c/F_c(\max)$, where $F_c(\max)$ is the maximum value of F_c, it starts to rise below a value of about 0.15; see Figure 5. This indicates that, on average, the measured weak intensities are too large. These observations, when considered alongside the correlation between intensity, uncertainty and resolution shown in Figure 2, indicate that the intensities of weak reflections have been systematically measured as too large at high resolution. One explanation for this is that at high

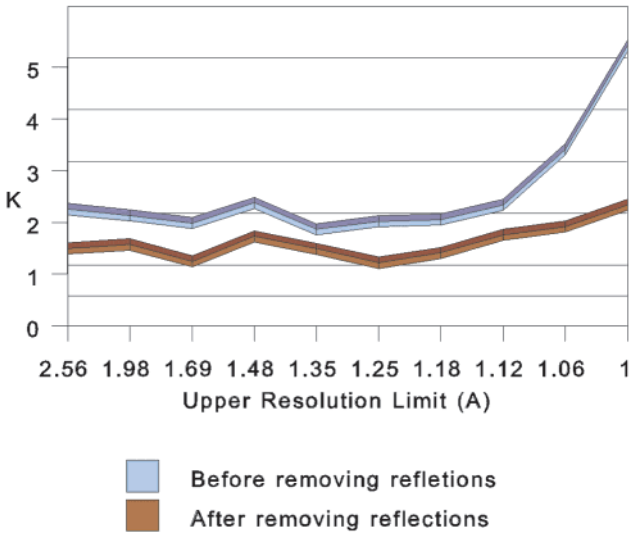

Figure 3. *K*, the mean of F_o^2 over the mean of F_c^2, calculated for reflections grouped in resolution shells, before and after reflections were removed to reestablish a flat analysis of variance.

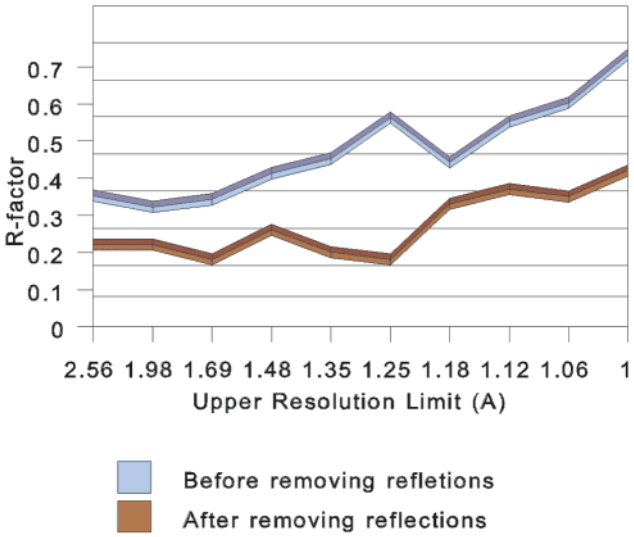

Figure 4. *R*, the crystallographic residual, calculated for reflections grouped in resolution shells, before and after reflections were removed to reestablish a flat analysis of variance.

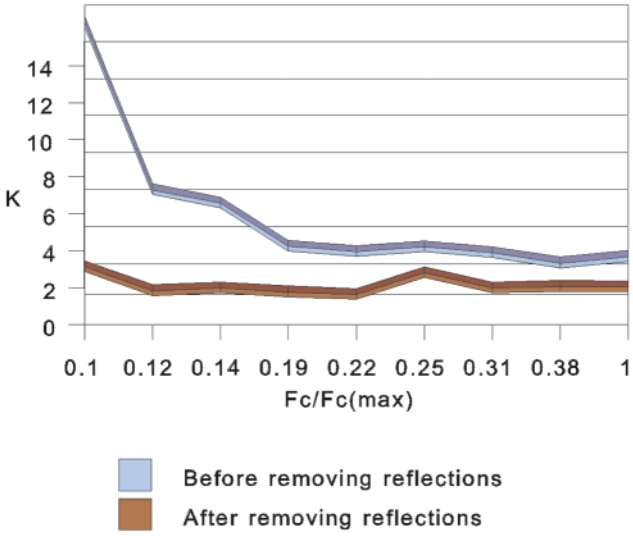

Figure 5. *K*, the mean of F_o^2 over the mean of F_c^2, calculated for reflections grouped in shells of F_o/F_c(max), where F_c(max) is the maximum value of F_c, before and after reflections were removed to reestablish a flat analysis of variance.

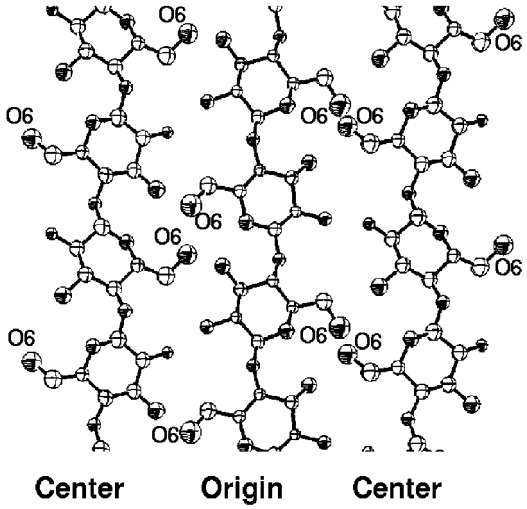

Figure 6. Structure of mercerized ramie cellulose, plotted using ORTEP,[33] showing thermal ellipsoids with 50% probability surfaces enclosing the atom centers. The view shows antiparallel chains packed in the crystallographic (110) plane. Only the hydroxymethyl group oxygen atoms have been labeled for clarity. The other atoms can be identified from Figure 7.

resolution, where spots have large extents because of arcing effects, LSQINT has fitted reflections to background features that have been underestimated. The background fitting method used in LSQINT will always tend to slightly underestimate the background.

To restore a flat analysis of variance and remove any systematic bias in the refinement we excluded all reflection with *d* spacings <1.5 Å and also F_c/F_c(max) < 0.15; see Figures 3−5. This reduced the number of data from 304 to 209 and resulted in values of 0.226 and 0.220 for *R* and R_{free}. It should be pointed out that this is similar to the conventional practice in fiber studies of leaving out very weak or absent reflections with $F_o > F_c$ from the refinement in order to avoid biasing the refinement by overestimating F_o. Allowing individual atomic thermal parameters to refine resulted in values of 0.218 and 0.213 for *R* and R_{free}. A final refinement, including the reflections which were set aside for the R_{free} calculation, resulted in a value of 0.211 for *R* and involved 90 parameters, 235 reflections and 82 restraints. Residual density in a Fourier difference map had a rms deviation from a mean of 0.125 e/Å^3 and maximum and minimum peak values of +0.454 and −0.433 e/Å^3.

All torsion angles, bond angles, and bond lengths of the high resolution SHELX97 refined structure were equal, within given uncertainties, to those of the medium resolution LALS structure. The root-mean-square displacement of nonhydrogen atoms in the high resolution structure from their positions in the low resolution structure is 0.13 Å. The high resolution refinement has not significantly changed the atomic position parameters but it has allowed us to model their thermal displacement parameters. The final structure is shown in Figure 6. The atomic parameters are shown in Table 2. During the first stage of refinement with SHELX97 when only data extending to a resolution of 1.5 Å were used, R_{free} was smaller than *R*. This is unusual and indicates that because the R_{free} data set was relatively small at that stage it

Table 2. Fractional Atomic Coordinates and Thermal Parameters (Å^2) for Mercerized Ramie Cellulose, with Esd's in Brackets[a]

atom	*x*	*y*	*z*	U_{iso}
		Origin Chain		
C1o	−0.043(6)	0.007(7)	0.381(2)	0.06(3)
H1o	−0.1257	−0.1115	0.3924	0.072
C2o	−0.125(7)	0.086(7)	0.286(3)	0.05(2)
H2o	−0.0411	0.2049	0.2764	0.056
C3o	−0.151(7)	−0.003(9)	0.156(3)	0.05(2)
H3o	−0.2396	−0.1205	0.1677	0.066
O2o	−0.299(7)	0.062(8)	0.334(4)	0.029(18)
C4o	0.034(7)	0.008(8)	0.112(3)	0.05(2)
H4o	0.1198	0.1243	0.0929	0.055
O3o	−0.224(10)	0.069(12)	0.066(4)	0.06(3)
C5o	0.118(9)	−0.057(9)	0.216(3)	0.04(2)
H5o	0.0297	−0.1731	0.2320	0.046
O1o	0.011(6)	−0.091(6)	−0.001(2)	0.038(17)
O5o	0.133(6)	0.034(8)	0.333(3)	0.047(18)
C6o	0.298(13)	−0.053(13)	0.183(5)	0.07(3)
H6oA	0.2935	−0.0928	0.0947	0.083
H6oB	0.3960	0.0603	0.1874	0.083
O6o	0.337(12)	−0.155(12)	0.270(8)	0.08(3)
		Center Chain		
C1c	0.468(6)	0.522(6)	−0.150(2)	0.04(2)
H1c	0.5652	0.6374	−0.1577	0.049
C2c	0.318(7)	0.506(8)	−0.054(3)	0.028(19)
H2c	0.2243	0.3886	−0.0506	0.034
C3c	0.398(7)	0.559(8)	0.082(3)	0.04(2)
H3c	0.4800	0.6792	0.0821	0.052
O2c	0.230(8)	0.603(9)	−0.092(4)	0.038(19)
C4c	0.509(6)	0.468(7)	0.119(3)	0.030(19)
H4c	0.4232	0.3516	0.1370	0.036.
O3c	0.250(8)	0.523(9)	0.170(4)	0.05(2)
C5c	0.640(7)	0.474(10)	0.011(3)	0.034(18)
H5c	0.7246	0.5920	−0.0036	0.040
O1c	0.621(5)	0.539(6)	0.232(2)	0.034(16)
O5c	0.541(8)	0.413(8)	−0.108(3)	0.050(17)
C6c	0.757(12)	0.392(17)	0.037(5)	0.06(3)
H6cA	0.8256	0.4356	0.1166	0.075
H6cB	0.6794	0.2739	0.0479	0.075
O6c	0.886(13)	0.418(15)	−0.067(8)	0.10(3)

[a] All esds are estimated using the full covariance matrix. The coordinates of hydrogen atoms were allowed to ride on the coordinates of the covalently bound atoms, with their thermal parameters tied to 1.5 or 1.2 times the thermal parameters of covalently bound oxygen or carbon, respectively

may not have been strictly statistically meaningful. When the data between 1.5 and 1 Å were included in the refinement, R_{free} behaved as expected.

Discussion

In the crystal structure of mercerized cellulose presented here the D-glucopyranoses of each chain are in the $^4C_1$ chair conformation with endocylic bond angles that do not deviate significantly from standard values. The calculated Cremer and Pople[30] puckering parameters indicate that, relative to the ideal value for an unstrained α-D-glycopusanose ($\theta = 2.7°$),[31] although the sugar of the origin chain is relatively unstrained ($\theta = 5°$), the sugar of the center chain is conformationally strained ($\theta = 10.2°$). The relative orientation of adjacent glycosyl residues, described by the glycosidic torsion angles ϕ and ψ, is also different for the center and origin chains. The hydroxymethyl groups of both chains are near the *gt* conformation. The center and origin chains have a relative displacement of ∼2.4 Å.

Table 3. Conformational Parameters of Models for Mercerized and Regenerated Cellulose II[a]

	origin chain				center chain				
	ϕ	ψ	χ'	θ	ψ	ψ	χ'	θ	*z*
B	−97(4)	95(4)	−165(7)	5.0	−94(8)	87(4)	−175(8)	10.2	0.23
LNC	−95.4	92.3	−172.4	3.9	−91.3	89.4	148.4	10.8	0.23
KWB	−96.0	93.6	−175.1	1.9	−96.4	94.0	−69.9	2.3	0.23
KK	−97.0	102.0	−169.0	5.9	−97.0	90.0	−179.0	17.6	0.23

[a] A comparison between the model reported in this study for mercerized cellulose, B, with the previously reported model of Kolpak, Weih, and Blackwell (KWB)[23] and the revised model for regenerated cellulose reported by Langan, Nishiyama, and Chanzy (LNC).[13] A model of cellulose II resulting from MD simulation reported by Kroon-Batenburg and Kroon is also shown, (KK).[32] The parameters are defined as follows: ϕ(O5−C1−O1−C4), ψ(C1−O1−C4−C3), and χ'(C4−C5−C6−O6). θ is the puckering parameter defined by Cremer and Pope[30] (averaged for KK), and *z* is the relative fractional displacement of the origin and center chains in the *c* axis direction.

A comparison of selected parameters of the structure presented here, designated B, with the earlier model for mercerized cellulose,[23] designated KWB, is given in Table 3. Table 3 also includes parameters from a model, designated LNC, that was obtained in a revision of the structure of regenerated cellulose.[13] LNC was obtained by LALS refinement of the original fiber diffraction model for regenerated cellulose and models based on the crystal structures of cellodextrins[14−16] against the original X-ray fiber diffraction data collected by Kolpak and Blackwell[7] from regenerated Fortisan fibers but with more degrees of freedom. This revised model was subsequently used to provide phases for neutron fiber diffraction studies of the hydrogen bonding system in cellulose II. However only the X-ray data from regenerated cellulose were used to refine the parameters for LNC given in Table 3. Models LNC and B display the same alternation between center and origin chains of glycosidic linkage and sugar pucker. Although the revised models for both mercerized and regenerated cellulose have all hydroxymethyl groups near the *gt* position, the exact values for the center chain are significantly different, with χ' having values of −175(8) and +148(6)°, respectively.[9] To test the significance of this difference we refined model B again using LALS but with χ' forced to the value observed in model LNC, 148°. The value of R'' increased from 0.160 to 0.172 and the resulting model could be rejected with respect to model B at a confidence level of 99.5%.

From Table 4, which shows the displacement of the atoms in mercerized cellulose from their positions in regenerated cellulose, it can be seen that these differences in conformational parameters reflect a relatively large displacement of the hydroxymethyl group of the central chain. This can be seen in Figure 7 where a representation of mercerised cellulose is shown with atom spheres that are proportional in size to the displacement values given in Table 4. The model published by Kroon-Batenburg and Kroon from MD simulations of cellulose II, is in better agreement with the revised model for mercerized cellulose than the revised model for regenerated cellulose.[32]

The revised model of mercerized cellulose differs significantly from the original model, KWB. In KWB the origin

Table 4. Displacement (Å) of Atoms in Mercerized Cellulose from Their Positions in Regenerated Cellulose[23]

atom	displacment	atom	displacment
	Origin Chain		
C1o	0.07	C5o	0.15
C2o	0.06	O1o	0.15
C3o	0.06	O5o	0.17
O2o	0.09	C6o	0.13
C4o	0.08	O6o	0.14
O3o	0.22		
	Center Chain		
C1c	0.14	C5c	0.35
C2c	0.14	O1c	0.17
C3c	0.13	O5c	0.29
O2c	0.19	C6c	0.44
C4c	0.17	O6c	0.16
O3c	0.18		

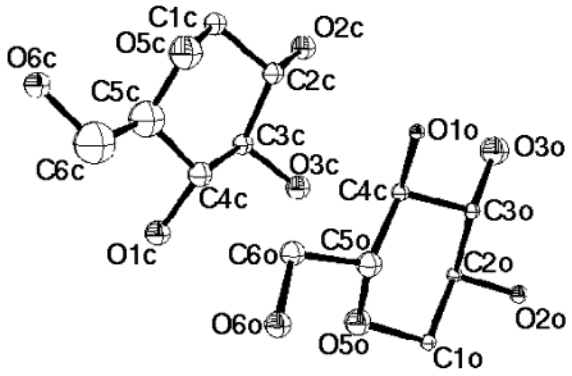

Figure 7. The asymmetric unit of the structure of mercerized Ramie cellulose, plotted using ORTEP,[33] showing surfaces enclosing the atom centers which are proportional in radius to the displacement of atoms in mercerized Ramie cellulose from their positions in regenerated cellulose, as given in Table 4.

and center chains have similar sugar puckers and glycosidic linkages. The hydroxymethyl group is near the *gt* conformation in the origin chain and near the *tg* conformation in the center chain. However the relative displacement of the chains in the *c*-axis direction is almost the same, ~2.4 Å.

Even with unreliable estimates of the uncertainties in LALS, the relative displacements of the chains in the *c*-axis direction are the same, being 2.358(31) and 2.387(10) Å for regenerated and mercerized cellulose, respectively. The *c* axis parameters are also the same to within experimental error, being 10.36(5) and 10.31(5) Å for regenerated and mercerized cellulose, respectively. Differences in conformation between mercerized and regenerated cannot be explained by a difference in the displacement of chains in the *c* axis direction. Although the unit cell *b* parameter, which reflects the packing of antiparallel sheets, is the same in mercerized and regenerated cellulose (9.03(3) and 9.04(5) Å, respectively), the unit cell a parameter, which reflects the packing of parallel chains in sheets, is not (8.10(3) and 8.01(5) Å, respectively).

The X-ray data presented in this study do not allow the direct determination of the positions of hydroxyl hydrogen atoms with any confidence. However, a recent neutron diffraction study of the hydrogen-bonding arrangement in cellulose II [13] revealed that the O6o atom and its bound hydrogen atom are in positions to donate to three possible acceptor atoms on the center chain (a major component, O6o...O6c, and two minor components, O6o...O5c and O6o...O3c). In the neutron study, it was unclear to what extent disorder of the O6c atom influenced this four-centered arrangement. The O6c atom of the central chain in regenerated cellulose was reported to be ~30% in a *tg* conformation and 70% in *gt*.

One explanation for the observed four-centered hydrogen-bonding arrangement is that when O6c is in a *gt* conformation, it accepts from O6o but when it is in the *tg* position, O3c accepts from O6o. A crystallographic averaged position for the hydrogen atom attached to O6o, statistically disordered in this way, would correspond to the hydrogen atom's observed position in the neutron studies. The four centered arrangement is then a statistical effect rather than a true four-centered hydrogen bond. The degree of disorder of O6c will affect the crystallographically determined positions of the hydroxmethyl group atoms of the center chain. However, disorder may also influence the most energetically optimum cooperative hydrogen network, and therefore the detailed crystal and molecular structure, providing an explanation for the differences in structure observed for mercerized and regenerated cellulose.

During refinement of the initial model for mercerized cellulose the authors also investigated the possibility of alternative orientations for the hydroxymethyl groups.[23] With all other parameters fixed, the orientation of the hydroxymethyl groups for the *gt*/*tg* models of both regenerated[7] and mercerized[23] cellulose were stepped and R'' mapped. A broad minimum in R'' was observed for mercerized cellulose whereas several minima were observed for regenerated cellulose. Although models incorporating these alternate hydroxymethyl group conformations were rejected for steric reasons, it was noted that disorder may be possible, but most likely exterior to the cellulose crystallites.

Figure 8 shows maps of R'' for the revised model of mercerized cellulose presented in this study, and the revised model of regenerated cellulose,[13] obtained by keeping all parameters fixed but allowing the occupancy of either the O6c or O6o hydroxymethyl group to be shared with an alternative orientation. The total occupancy of the O6 group and its alternate orientation add to unity. In all cases, as the occupancy of the alternate orientation increases toward unity a strong minimum develops at the same value of χ' as the fixed O6 group i.e., the O6 group and its alternate orientation coincide. However, in all cases these are not the deepest minima.

For regenerated cellulose, R'' displays well-defined minima in χ', at +75° (*gg*) and −60° (*tg*) for O6o and at −100° (*tg*) for O6c, when the alternate orientation has an occupancy of ~30%. These minima indicate an improvement in the agreement with the data at a confidence level of ~99.5%. For mercerized cellulose, R'' displays less distinct minima in χ', for O6c and O6o, when the alternate orientation has an occupancy of about 10%. This minimum indicates an improvement in the agreement with the data at a confidence level of ~90%.

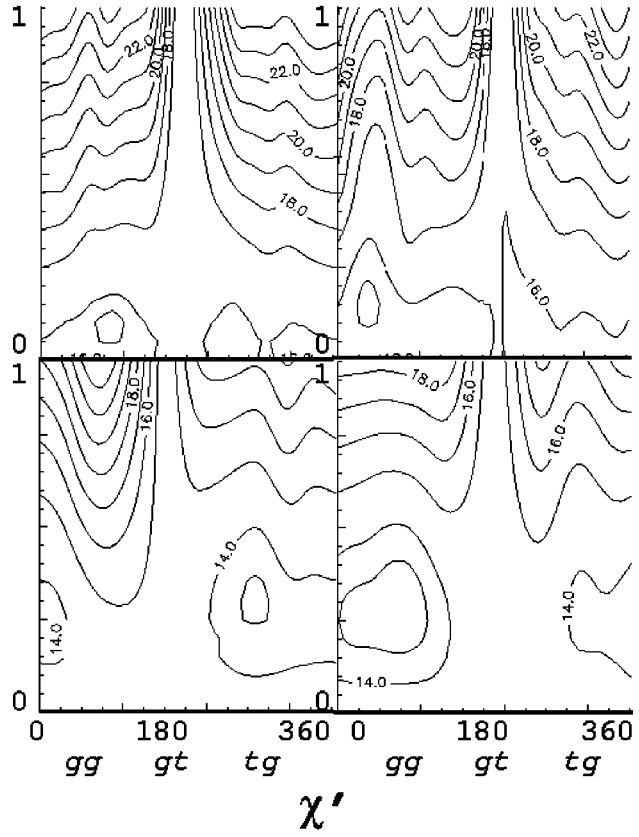

Figure 8. Maps of R'' generated for the revised model of mercerized cellulose presented in this study (top half) and the revised model of regenerated cellulose [13] (bottom half) as the occupancy of either the O6o group (right half) or the O6c group (left half) is shared between an alternative orientation. The occupancy of the alternate orientation is plotted vertically, and its value, χ', horizontally. Every second contour level is labeled with the corresponding R'' value for clarity.

Conclusion

This study provides a revised model for mercerized cellulose. The results presented here confirm that, like regenerated cellulose, mercerized cellulose consists of antiparallel chains packed in a $P2_1$ monoclinic unit cell. The conformations of the chains in mercerized and regenerated [13] cellulose are also similar and indicate that these different ways of preparing cellulose II result in similar crystal and molecular structures. There are however significant differences in the conformation of the hydroxymethyl groups of the center chains. This may be related to the large observed difference in the amount of hydroxymethyl group disorder. Whether this disorder is confined to the surface of the crystallites or is pervasive cannot be determined from this study, but it must be related to material processing. The conformation of the chains differs significantly from the original model for mercerized cellulose.[23]

Although this study involved relatively high resolution synchrotron X-ray data, we were unable to obtain reliable positions for the hydroxyl hydrogen atoms involved in hydrogen bonding because of the relative weak scattering power of hydrogen for X-rays. Hydrogen atoms and their isotopes have relatively large scattering powers for neutrons. Neutron diffraction data have already been collected from mercerized cellulose fibers. [13] We are in the process of collecting neutron diffraction data from regenerated cellulose fibers so that any difference in the hydrogen bonding arrangements in these two types of cellulose II, can be identified and related to differences in conformation and disorder.

Acknowledgment. We thank the ESRF, Grenoble for the provision of facilities. Y.N thanks the French Government and the Japanese Society for the Promotion of Science for financial support. We also thank Loes Kroon-Batenburg for useful discussions.

References and Notes

(1) Atalla, R. H.; VanderHart, D. L. *Science* **1984**, *223*, 283.
(2) VanderHart, D. L.; Atalla, R. H. *Macromolecules* **1984**, *17*, 1465.
(3) Sugiyama, J.; Vuong, R; Chanzy, H. *Macromolecules* **1991**, *24*, 4168.
(4) Koyama, M.; Helbert, W.; Imai, T.; Sugiyama, J; Henrissat, B. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9091.
(5) Nishiyama, Y.; Okano, T.; Langan, P.; Chanzy, H. *Int. J. Biol. Macromol*. **1999**, *26*, 279.
(6) Finkenstadt, V. L.; Millane, R. P. *Macromolecules* **1998**, *31*, 7776.
(7) Kolpak, K. J.; Blackwell, J. *Macromolecules* **1976**, *9*, 273.
(8) Stipanovic, A.; Sarko, A. *Macromolecules* **1976**, *9*, 851.
(9) 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 and 180°.
(10) When the *z* coordinate of O5 is greater than that of C6, the chain is defined as "up", otherwise as "down". We refer to the "up" and "down" chain as the origin and center chain respectively and atoms labels are corresponding post-fixed with an "o" or "c".
(11) Gessler, K.; Krauss, N.; Steiner, T.; Betzl, C.; Sarko, A.; Saenger, W. *J. Am. Chem. Soc*. **1995**, *117*, 11397.
(12) Raymond, S.; Kvick, Å.; Chanzy, H. *Macromolecules* **1995**, *28*, 8422.
(13) Langan, P.; Nishiyama, Y.; Chanzy, H. *J. Am. Chem. Soc.* **1999**, *121*, 9940.
(14) Gessler, K.; Krauss, N.; Steiner, T. Betzel, C. Sandman, C.; Saenger, W. *Science* **1944**, *266*, 1027.
(15) Raymond, S.; Heyraud, A.; Tran Qui, D.; Kvick, Å; Chanzy, H. *Macromolecules* **1995**, *28*, 2096.
(16) Raymond, S.; Henrissat, B.; Tran Qui, D.; Kvick, Å.; Chanzy, H. *Carbohydr. Res*. **1995**, *277*, 209.
(17) The glycosidic torsion angles, ϕ and ψ, which describe the relative orientation of adjacent glycosyl residues in the same chain are defined as (O5−C1−O1−C4) and (C1−O1−C4−C3), respectively.
(18) Dudley, R. L.; Fyfe, C. A.; Stephenson, P. J.; Deslandes, Y.; Hamer, G. K.; Marchessault, R. H. *J. Am. Chem. Soc*. **1983**, *105*, 2469.
(19) Fyfe, C. A.; Stephenson, P. J.; Veregin, R. P.; Hamer, G.; Marchessault, R. H. *Carbohydr. Chem*. **1984**, *3*, 663.
(20) Isogai, A.; Usuda, M. Kato, T. Uryu, T.; Atalla, R. H. *Macromolecules* **1989**, *22*, 3168.
(21) Horii, F.; Hirai, A.; Kitamaru, R.; Sakurada, I. *Cellulose Chem. Technol*. **1985**, *19*, 513.
(22) Horii, F.; Hirai, A.; Kitamaru, R.; Sakurada, I. *Polym. Bull*. **1983**, *10*, 3168.
(23) Kolpak, F. J.; Weih, M.; Blackwell, J. *Polymer* **1978**, *19*, 123.
(24) Fengel, D. *Papier* **1993**, *12*, 695.
(25) Kroon-Batenburg, L. M. J.; Bouma, B.; Kroon, J. *Macromolecules* **1996**, *29*, 5695.
(26) Campbell Smith, P. J.; Arnott, S. *Acta Crystallogr*. **1978**, *A34*, 3.
(27) Hamilton, W. C. *Acta Crystallogr*. **1965**, *18*, 502.
(28) $R'' = [\sum_i w_i(|F_i|_o - |F_i|_c)^2/\sum_i w_i|F_i|_o^2]^{1/2}$, where F_o, F_c, and w are the observed and calculated structure factor amplitudes and their weights, respectively.
(29) Sheldrick, G. M. *SHELX-97, a program for the Refinement of Single-Crystal Diffraction Data*; University of Gottingen: Gottingen, Germany, 1997.
(30) Cremer, D; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354.
(31) Dowd, M.K;French, A. D.;Reilly, P. J. *Carbohydr. Res.* **1994**, *264*, 1.
(32) Kroon-Batenburg, L. M. J.; Kroon, J. *Glycoconj. J*. **1997**, *14*, 677.
(33) Barnes, C. L. *J. Appl. Crystallogr*. **1997**, *30*, 568−568.

BM005612Q