Mercerization of Cellulose. 6. Crystal and Molecular Structure of Na-Cellulose IV†

Hisao Nishimura‡ and Anatole Sarko*

Department of Chemistry and the Cellulose Research Institute, College of Environmental Science and Forestry, State University of New York, Syracuse, New York 13210

Received December 18, 1989; Revised Manuscript Received June 29, 1990

ABSTRACT: The crystal structure of Na-cellulose IV, an intermediate in the mercerization of cellulose, has been investigated by using combined X-ray diffraction analysis and stereochemical model refinement. The structure is derived from Na-cellulose II by washing with water, and it represents the final step prior to complete conversion to cellulose II. Its most probable crystal structure is based on a two-chain, monoclinic unit cell with antiparallel chain packing and unit cell parameters a = 9.57 Å, b = 8.72 Å, c (fiber axis) = 10.35 Å, and γ = 122.0°. The unit cell contains two water molecules in crystallographically defined positions, located between the corner chairs. All four hydroxymethyl groups are in gt positions, and $P2_1$ symmetry is very nearly observed. All of the cellulose hydroxyl groups participate in hydrogen bonding, and each water molecule takes part in four hydrogen bonds. The process of conversion of Na-cellulose II to Na-cellulose IV suggests that cellulose chains begin aggregating into a sheet structure as a result of hydrophobic attractions. A similar sheet structure is present in all cellulose crystal structures that possess a cellobiose type fiber repeat.

Introduction

Of the five crystalline NaOH–cellulose complexes ("Na-celluloses") observed during controlled mercerization of ramie cellulose, Na-cellulose IV is the final structure in the series.[1,2] It does not actually contain NaOH, as it is derived from Na-celluloses I, II, or III by washing with water, thereby eliminating all NaOH. It is converted to cellulose II by drying. Its crystal structure became of interest in view of its water content and the possible influence of the latter on the structure of cellulose. The analysis of the crystal structure was, therefore, undertaken using X-ray fiber diffraction procedures assisted by molecular modeling.

Experimental Section

The starting material was, as in previous studies,[1-5] commercially refined ramie fibers. The specimens for mercerization and X-ray diffraction studies were prepared as follows. A bundle of fibers—sufficient to fill a 1-mm-diameter glass tube normally used for X-ray diffraction samples—was carefully straightened, smoothed, and inserted in the tube. The latter was then filled with 8 N aqueous NaOH solution. After the cellulose I structure had been completely converted to Na-cellulose II, the alkali was washed out of the tube and both ends of the tube were quickly sealed to prevent drying. All of these steps were conducted at room temperature.

Two series of samples were prepared in this fashion, differing principally in the time needed to wash out the alkali. The samples of series A contained the maximum amount of cellulose fibers that could be inserted in the sample tube. The swelling of the fibers that normally occurs during mercerization exerted sufficient lateral pressure on the fibers to slow down both the process of mercerization and the subsequent washing. Under these conditions, complete conversion of cellulose I to Na-cellulose II took about 1 month. The washing process, conducted by simply immersing the sample tube in a distilled water bath, was completed in ca. 1 week.

On the other hand, the samples of series B were prepared for more rapid washing. Only about half of the fibers used in series A were placed in the tube. To prevent the misalignment that would occur in such a loosely packed tube during mercerization, the fibers were left extending 2–3 cm out of both ends of the tube. The fiber ends were secured in a stainless steel clamp that was adjusted to keep the fibers straight, but not under tension. The fibers were kept clamped in this fashion for the entire experiment. The mercerization process for these samples was completed in ca. 2 weeks, and the subsequent washing was effected with a syringe while keeping the sample in a water bath. The washing process under these conditions took ca. 75 min.

The procedures for recording the X-ray diffraction diagrams, film measurements, recording and resolution of overlapped intensities, and the correcting of measured intensities have been previously described (see the preceding paper). In view of the 1-mm thickness of the specimens, an absorption correction based on the linear absorption coefficient of water (10.3 cm^{-1}) was additionally applied to the intensities. The structure analysis and refinement procedures are described in the following section.

Results

A. Unit Cell. A typical X-ray fiber diffraction diagram of Na-cellulose IV is shown in Figure 1. The pattern is similar to that of cellulose II,[6] suggesting similarities in the unit cells and structures of both crystalline substances. Using seven strong and well-resolved reflections—three on the equator, two on the second layer line, and one each on the third and fourth layer lines—the unit cell constants of Na-cellulose IV of both series A and B samples were refined with least-squares procedures. The results, shown in Table I, indicate that the series B (quick washing procedure) unit cell is very slightly larger in comparison with that of series A. However, the difference was too small to affect the final structural characteristics. It was, therefore, concluded that neither the time of mercerization nor the time of washing exerted any significant effect on the characteristics of the unit cell of Na-cellulose IV.

The final unit cell parameters, based on the refinement with all reflections for samples of series A, are shown in Table II, along with the correspondence of observed and calculated d spacings. All subsequent measurements were made on series A samples, because of slightly better resolution of intensities in their patterns as compared with series B. The poorer resolution in the latter was probably a result of less well-formed crystallites caused by the more rapid washing procedure.

A firm assignment of the space group could not be made, but as for cellulose II, the structure appears to be close to $P2_1$ symmetry. (A faint 003 meridional reflection is

* To whom correspondence should be addressed.

† Part 5 of this series: *Macromolecules*, preceding paper in this issue.

‡ Permanent address: Daicel USA, Inc., 2 Executive Drive, Fort Lee, NJ 07024.

0024-9297/91/2224-0771$02.50/0

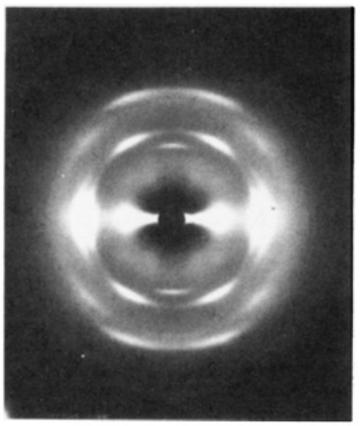

Figure 1. X-ray diffraction pattern of Na-cellulose IV. Fiber axis is vertical.

Table I
Unit Cell Refinement Based on the Strongest Reflections of Series A and Series B Diffraction Diagrams

Series A: Unit Cell Parameters a = 9.43 Å, b = 8.68 Å, c = 10.37 Å, and γ = 121.1°

	d spacing, Å	
hkl	obsd	calcd
$\bar{1}10$	7.89	7.85
110	4.45	4.44
200	4.03	4.04
002	5.17	5.19
102	4.37	4.36
013	3.13	3.13
104	2.47	2.47

Series B: Unit Cell Parameters a = 9.46 Å, b = 8.72 Å, c = 10.38 Å, and γ = 121.1°

$\bar{1}10$	7.87	7.88
110	4.46	4.46
200	4.05	4.05
002	5.17	5.19
102	4.39	4.37
013	3.14	3.14
104	2.47	2.47

observed, indicating that the latter space group is not allowed.) With four anhydroglucopyranose residues per unit cell, the calculated crystal density ρ_c = 1.453 g/cm^3, to be compared with the observed density of ~1.5 g/cm^3.[2] The latter value, although somewhat approximate, suggests the presence of water in the unit cell. The calculated densities for selected water contents of the unit cell are shown in Table III.

B. Structure Determination. The intensity tracings obtained from the diffraction diagrams of series A samples could be resolved into 25 individual intensity groupings. A total of 60 hkl planes were predicted to contribute to these intensities as calculated from the unit cell. An additional 19 reflections were predicted but not observed, and their intensities were estimated at one-half of the minimum observable intensity at the corresponding diffraction angle. Therefore, a total of 79 individual reflection intensities could be used in the structure refinement.

Initial Refinement. Because of the similarity of the cellulose II and Na-cellulose IV unit cells, the structure determination of the latter was considerably simplified. A cellulose II type molecular chain, with a twofold screw axis coincident with the chain axis and relating the positions of all but the O(6) atoms of the glucose residues, was placed in both the corner and the center of the unit cell. The two chains were in antiparallel orientation. The positions of all four O(6) atoms, as determined by the rotation about the C(5)–C(6) bond, remained independently variable. On the basis of three most probable rotational positions for each O(6)—gt, gg, and tg—a total of 12 starting models were generated. Of these, three were "pure" models in which all four O(6)'s were in the same position, six were "mixed" models in which corner and center chains differed in O(6) rotations, and another three were "mixed" models in which an alternating sequence of two different O(6) positions occurred along the chain. All 12 models were subjected to an initial refinement against the complete intensity data, using as variables only chain rotations and center chain translation along the c axis. No water molecules were included in the unit cell, and an isotropic temperature factor was arbitrarily set at 5. The primary refinement criteria were the usual weighted and unweighted residuals:

$$R'' = \left\{ \frac{\sum w_i (|F_i|_o - |F_i|_c)^2}{\sum w_i |F_i|_o^2} \right\}^{1/2} \quad (1)$$

$$R = \frac{\sum ||F_i|_o - |F_i|_c|}{\sum |F_i|_o} \quad (2)$$

(The weights w for the observed reflections were set at 1. For the unobserved reflections w = 0.5 when the calculated structure factor exceeded the observed value; otherwise w = 0.)

Additional criteria in the refinement were the number and nature of short nonbonded contacts and the number

Table II
Comparison of Observed and Calculated d Spacings of Na-Cellulose IV, Series A, All Reflections (Refined Unit Cell Parameters: a = 9.57 Å, b = 8.72 Å, c = 10.35 Å, γ = 122.0°[a])

	d spacing, Å	
hkl	obsd	calcd
$\bar{1}10$	7.89	7.95
110	4.45	4.42
200	4.03	4.06
$\bar{1}11$	6.22	6.31
201	3.75	3.78
$\bar{3}31$	2.56	2.57
002	5.17	5.18
102	4.37	4.36
112	3.39	3.36
022	2.96	3.01
212	2.58	2.57
122	2.50	2.49
$\bar{3}32$	2.35	2.36
013	3.13	3.13
113	2.74	2.72
203	2.61	2.63
023	2.49	2.52
$\bar{2}33$	2.22	2.22
$\bar{3}33$	2.12	2.10
004	2.58	2.59
104	2.47	2.47
114	2.24	2.23
204	2.19	2.18

[a] For comparison, the unit cell parameters of cellulose II are a = 9.09 Å, b = 7.96 Å, c = 10.31 Å, and γ = 117.3°.

Table III
Calculated Density of the Unit Cell of Na-Cellulose I as a Function of Water Content

no. of H_2O in unit cell	density, g/cm^3	no. of H_2O in unit cell	density, g/cm^3
0	1.453	4	1.614
1	1.493	6	1.695
2	1.534	8	1.776
3	1.574		

Table IV
Structural Characteristics of Acceptable Initial Models

model	R''	interchain hydrogen bonds and lengths,[a] Å
model 1		
chain rotations[b]	0.202	$O(2)_3$---$O(6)_3$ = 2.81
corner: 20.0°		$O(2)_4$---$O(6)_4$ = 2.81
center: 130.3°		$O(2)_2$---$O(6)_3$ = 2.45
center chain translation: −8.97 Å		$O(2)_1$---$O(6)_4$ = 2.45
O(6) rotations:[c] all gt		$O(6)_1$---$O(2)_4$ = 2.69
		$O(6)_2$---$O(2)_3$ = 2.69
model 2		
chain rotations	0.215	$O(2)_3$---$O(6)_3$ = 3.00
corner: 21.0°		$O(2)_4$---$O(6)_4$ = 3.00
center: 127.7°		$O(6)_1$---$O(3)_4$ = 2.76
center chain translation: −1.33 Å		$O(6)_2$---$O(3)_3$ = 2.76
O(6) rotations: all gt		$O(3)_1$---$O(6)_4$ = 3.10
		$O(3)_2$---$O(6)_3$ = 3.10
model 3		
chain rotations	0.215	$O(2)_3$---$O(6)_3$ = 2.95
corner: 20.8°		$O(2)_4$---$O(6)_4$ = 2.95
center: 128.4°		$O(2)_2$---$O(6)_3$ = 2.37
center chain translation: −9.03 Å		$O(2)_1$---$O(6)_4$ = 2.37
O(6) rotations: tg_1, tg_2, gt_3, gt_4		
model 4		
chain rotations	0.220	O(2)---$O(6)_2$ = 3.01
corner: 24.7°		$O(2)_2$---$O(6)_3$ = 2.73
center: 125.3°		$O(6)_2$---$O(3)_3$ = 2.97
center chain translation: −2.32 Å		$O(6)_2$---$O(6)_4$ = 2.96
O(6) rotations: tg_1, gt_2, tg_3, gt_4		
model 5		
chain rotations	0.221	$O(2)_2$---$O(6)_2$ = 3.02
corner: 24.5°		$O(6)_1$---$O(6)_3$ = 2.51
center: 126.3°		$O(6)_2$---$O(2)_3$ = 2.76
center chain translation: −9.04 Å		$O(2)_1$---$O(6)_4$ = 2.41
O(6) rotations: tg_1, gt_2, tg_3, gt_4		

[a] All subscripts indicate residue numbers: 1,2, corner chain; 3,4, center chain. [b] Chain at 0° rotation has O(4) at coordinates (0, −y, z) (before translation to its position in the unit cell). Positive rotation is clockwise looking down the c axis. [c] O(6) is at 0° when the bond sequence O(5)–C(5)–C(6)–O(6) is cis. Rotation of C(6)–O(6) is positive-clockwise looking from C(5)–C(6), and pure gt = 60°, tg = 180°, and gg = −60°.

and characteristics of inter- and intrachain hydrogen bonds. On the basis of these combined criteria, the initial refinement quickly eliminated all models containing O(6) in the gg position, as had been expected from similarity with cellulose II. Of the remainder, the five models described in Table IV possessed both the lowest residuals and no short nonbonded contacts and were therefore selected for further refinement. (The rejected models were all marked by values of R'' higher than 0.23.)

Water Molecules. A comparison of the experimental and calculated densities suggested that the most probable number of water molecules in this type of chain arrangement was either 2 or 4 per unit cell (cf. Table III). Both models were analyzed. The approximate placement of water molecules (as oxygen atoms) was determined in two stages. First, their (x, y) coordinates were determined by using equatorial intensity data only. The strategy was identical with that used for locating Na^+ ions in the Na-cellulose I structure,[5] in that the base plane of the unit cell was divided into 30 rectangles 1.64 × 1.46 Å^2 in size, and the effect on the R'' factor of placing and refining the position of an oxygen atom in each such cell was determined. As illustrated in Figure 2, the results of the (x, y) refinement clearly suggested only two locations for the water molecules. On the basis of this analysis, the three water models shown in a–b projection in Figure 3 were

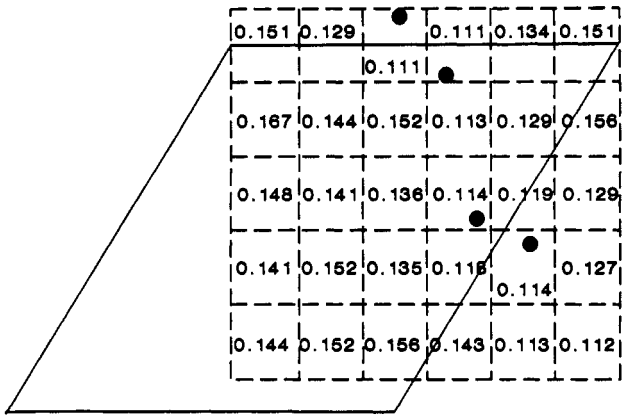

Figure 2. Two-dimensional R'' value map for locating water molecules.

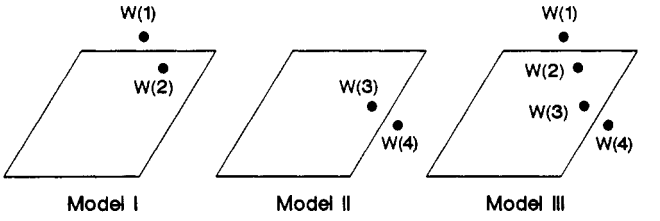

Figure 3. The three most probable models for water locations.

Table V
Preliminary X-ray Intensity Refinement for Water-Containing Models

water location model[a]	cellulose chain model[b]	R''	B[c]	E_p[d]
model I	model 1	0.150	10.5	71
	model 2	0.128	10.8	73
	model 3	0.158	8.4	133
	model 4	0.176	8.2	146
	model 5	0.160	8.1	168
model II	model 1	0.155	8.2	118
	model 2	0.131	9.9	96
	model 3	0.173	8.4	175
	model 4	0.169	8.8	86
	model 5	0.167	8.2	164
model III	model 1	0.126	10.2	183
	model 2	0.094	12.9	164
	model 3	0.130	9.3	259
	model 4	0.129	10.5	214
	model 5	0.134	8.7	285

[a] Refer to Figure 3. [b] Refer to Table IV. [c] Isotropic temperature factor. [d] E_p = packing energy, calculated from the nonbonded term of eq 4 in the preceding paper (arbitrary units).

used in subsequent refinements.

In the second step of determining the water positions the z coordinates were estimated by using complete three-dimensional data. The chain rotations and translation, the (x, y, z) coordinates of the water oxygens, and the isotropic temperature factor were the refinable parameters at this stage. The procedure was repeated for all five of the surviving chain models shown in Table IV. The results, given in Table V, strongly suggest chain model 2 as the most probable structure, but a clear choice among the three water models could not be made. One of the contributing reasons to the inability to choose the best water model was the high intermolecular packing energy E_p (column 5 in Table V) for all three models, but especially for water model III. (The intermolecular packing energy E_p is calculated from the nonbonded term of the total E expression, eq 4 in the preceding paper).

As a result of these refinements, all but the chain model 2 were eliminated, but all three water models were retained. These three models were now subjected to a final round

Table VI
Pre- and Postrefinement Characteristics of the Final Three Models (Chain Model 2[a])

water model[b]	characteristic	prerefinement value	postrefinement value
model I	R''	0.128	0.150
	E_p	73	40
	short contacts[c] (Å)		
	$C(2)_2$---W(2)[d]	2.59	(3.10)
model II	R''	0.131	0.175
	E_p	96	46
	short contacts (Å)		
	$C(6)_3$---W(4)	2.38	2.35
	$C(6)_4$---W(3)	2.16	2.46
	$C(6)_2$---W(4)	2.53	2.65
	$H(6A)_3$---W(4)	1.81	1.75
	$H(6A)_4$---W(3)	1.86	(2.26)
	$H(6A)_2$---W(4)	1.70	1.74
	$H(6A)_1$---W(3)	2.12	(2.71)
model III	R''	0.094	0.153
	E_p	164	69
	short contacts (Å)		
	$C(3)_4$---W(3)	2.46	(3.27)
	$C(2)_4$---W(3)	2.64	(>3.60)
	$C(6)_3$---W(4)	2.20	2.29
	$C(6)_4$---W(3)	2.49	2.51
	$C(6)_2$---W(4)	2.18	(3.03)
	$C(2)_2$---W(2)	2.60	(3.23)
	$C(2)_1$---W(1)	2.52	(3.46)
	$H(6A)_3$---W(4)	1.61	1.98
	$H(6A)_2$---W(4)	1.36	(2.38)
	$C(6)_1$---W(1)	(3.22)	2.69
	$C(6)_2$---W(2)	(3.22)	2.56
	$H(6B)_3$---W(4)	(2.28)	2.04
	$H(6B)_4$---W(3)	(2.48)	2.01
	$H(6B)_1$---W(1)	(2.84)	2.08
	$H(6B)_2$---W(2)	(2.87)	1.85

[a] Refer to Table IV. [b] Refer to Figure 3. [c] Distance less than 2.80 Å for O---C, and 2.20 Å for O---H and C---H contacts. Values in parentheses are acceptable. [d] W = oxygen of water. Refer to Figure 2 for numbering.

of refinement.

Final Refinement. To minimize the effects of unacceptable stereochemistry and short nonbonded contacts, the function

$$\Phi = fR'' + (1-f)E \qquad (3)$$

was minimized in all following refinements. (In this function, E is the energy calculated from eq 4 in the preceding paper, R'' is expressed as a percentage, and the fraction f could be varied to weight the contributions of the two terms. Typically, f = 0.90–0.95.) The refinable parameters were chain rotations and translation, four independent O(6) rotations, the (x, y, z) coordinates of water oxygens, and the isotropic temperature factor.

The final results are shown in Table VI, in terms of both best pre- and postrefinement characteristics of the three water models. Clearly, model I is superior to the other two. Its one short C(2)- - -W(2) contact, at 2.59 Å before refinement, was easily eliminated at a modest increase in the R'' factor from 0.128 to 0.15. On the other hand, the multiple short contacts in the other two models could not be significantly improved even when the R'' factor was allowed to rise as high as 0.175. Consequently, the most probable structural model for Na-cellulose IV was taken as the combination of chain model 2 and water model I, in which all O(6) were gt and two water molecules were located in the 100 plane between the corner chains.

In this model, the water molecules are located in crystallographically defined positions. An alternative model was still considered possible—one in which the water was located randomly. To test this hypothesis, the chain model 2 structure was refined without water, but with water-weighted atomic scattering factors to simulate the random presence of water in the structure.[7] The resultant R'' = 0.221 clearly did not support such a structure.

Table VII
Characteristics of the Most Probable Antiparallel Model after Final Refinement (Chain Model 2, Water Model I)

characteristic[a]	value
glycosidic bridge angles, deg	
1,3 residues	117.1
2,4 residues	117.7
chain rotation, deg	
corner	18.8
center	128.7
center chain translation, Å	−1.32
O(6) rotations, deg	
residue 1	68.9
residue 2	57.1
residue 3	53.6
residue 4	49.9
virtual bonds,[b] Å	
residues 1,3	5.45
residues 2, 4	5.47
conformation angles ϕ,ψ,[c] deg	
residues 1, 3	28.1, −30.8
residues 2, 4	28.0, −29.8
fractional atomic coordinates of water oxygens (x, y, z)	
W(1)	−0.089, 0.445, −0.895
W(2)	0.107, 0.578, −0.329
intrachain hydrogen bonds, Å	
$O(3)_2$- - -$O(5)_1$, $O(3)_4$- - -$O(5)_3$	2.60
$O(3)_1$- - -$O(5)_2$, $O(3)_3$- - -$O(5)_4$	2.59
interchain hydrogen bonds, Å	
$O(2)_3$- - -$O(6)_3$	2.84
$O(2)_4$- - -$O(6)_4$	2.80
$O(6)_1$- - -$O(3)_4$	2.82
$O(6)_2$- - -$O(3)_3$	2.69
hydrogen bonds to water, Å	
$O(2)_1$- - -W(1)	2.94
$O(3)_1$- - -W(1)	2.41
$O(6)_1$- - -W(1)	2.65
$O(6)_4$- - -W(1)	2.83
$O(2)_2$- - -W(2)	2.35
$O(3)_2$- - -W(2)	2.56
$O(6)_2$- - -W(2)	2.65
$O(6)_3$- - -W(2)	2.76
residuals (isotropic temp factor)	
R	0.167
R''	0.143
B	11.9
residuals (anisotropic temp factor)	
R	0.152
R''	0.130
B_x, B_y, B_z	2.8, 10.8, 8.5

[a] Refer to Figure 3 for residue numbering and to Figure 2 for water numbering. [b] Length of the vector connecting successive bridge oxygens. [c] $\phi = 0°$ when the bond sequence H(1)–C(1)–O(4)′–C(4)′ is cis; $\psi = 0°$ when the bond sequence C(1)–O(4)′–C(4)′–H(4)′ is cis; the angle is positive when the far bond is clockwise relative to the near bond.

In the final cycles of refinement, the restriction of 2_1 symmetry along the chains (relative to ring atoms) was removed, and all bond angles and torsion angles along the chain were allowed to vary. (It should be noted that owing to a relatively small data set, the variables of only one chain were refined in this manner and the second chain was simply a copy of the first chain, with the exception that all four O(6) rotations remained independent. All bond lengths remained invariant.) In addition, three components of an anisotropic temperature factor, B_x, B_y, and B_z, were refined.[8]

The final results are shown in Table VII in terms of model characteristics and in Table VIII in terms of a

Table VIII
Observed and Calculated Structure Factor Amplitudes for the Final Two-Water, Antiparallel Model ($R'' = 0.130$, Anisotropic Temperature Factor)

d spacing, Å					*d* spacing, Å				
obsd	calcd[a]	*hkl*	F_o[b]	F_c	obsd	calcd[a]	*hkl*	F_o[b]	F_c
				Equator					
7.89	8.11	100	57.3	62.1	3.00	3.13	$\bar{3}10$	31.2	24.1
	7.95	$\bar{1}10$				3.10	$\bar{3}20$		
	7.40	010				2.96	210		
4.45	4.78	$\bar{2}10$	167	165		2.89	$\bar{2}30$		
	4.42	110				(2.85)	120	(13.5)	6.10
	4.34	$\bar{1}20$				(2.81)	$\bar{1}30$	(13.7)	2.19
4.03	4.06	200	154	155		(2.71)	300	(14.2)	25.3
	3.98	$\bar{2}20$			2.60	2.65	$\bar{3}30$	20.9	8.92
	3.70	020				(2.47)	030	(15.4)	12.8
				First Layer Line					
	(10.35)	001				(3.48)	021	(10.5)	16.3
6.22	6.39	101	11.0	10.3		(3.00)	$\bar{3}11$	(12.4)	11.9
	6.31	$\bar{1}11$				(2.97)	$\bar{3}21$	(12.5)	2.91
	6.02	011				(2.84)	211	(13.0)	21.1
	(4.34)	$\bar{2}11$	(8.28)	11.4		(2.79)	$\bar{2}31$	(13.4)	5.35
	(4.07)	111	(8.92)	15.2		(2.74)	121	(13.5)	17.4
	(4.01)	$\bar{1}21$	(9.09)	13.1		(2.71)	$\bar{1}31$	(13.7)	19.2
3.75	3.78	201	8.84	18.7	2.56	2.62	301	28.1	35.5
	3.71	$\bar{2}21$				2.57	$\bar{3}31$		
				Second Layer Line					
5.17	5.18	002				(2.68)	$\bar{3}12$	(12.6)	19.3
4.37	4.36	102	45.5	41.7		(2.66)	$\bar{3}22$	(12.7)	13.0
	4.34	$\bar{1}12$			2.58	2.57	212	27.7	22.1
	4.24	012			2.50	2.53	$\bar{2}32$	19.6	21.4
3.39	3.51	$\bar{2}12$	38.8	35.3		2.49	122		
	3.36	112				2.47	$\bar{1}32$		
	3.33	$\bar{1}22$			2.35	2.40	302	22.8	23.2
3.19	3.19	202	14.3	24.5		2.36	$\bar{3}32$		
	3.15	$\bar{2}22$				(2.23)	032	(15.7)	13.1
2.96	3.01	022	20.2	15.8					
				Third Layer Line					
3.38	3.45	003			2.49	2.52	023	30.7	12.2
3.13	3.18	103	38.8	43.4		(2.32)	$\bar{3}13$	(12.8)	7.15
	3.17	$\bar{1}13$				(2.31)	$\bar{3}23$	(13.0)	14.7
	3.13	013			2.22	2.25	213	43.8	47.8
2.74	2.80	$\bar{2}13$	12.9	17.5		2.22	$\bar{2}33$		
	2.72	113				2.20	123		
	2.70	$\bar{1}23$				2.18	$\bar{1}33$		
2.61	2.63	203	36.6	39.5	2.12	2.13	303	42.2	33.4
	2.61	$\bar{2}23$				2.10	$\bar{3}33$		
				Fourth Layer Line					
2.58	2.59	004				2.23	114		
2.47	2.47	104	18.0	18.2		2.22	$\bar{1}24$		
	2.46	$\bar{1}14$			2.19	2.18	204	24.4	25.7
	2.44	014				2.17	$\bar{2}24$		
2.24	2.28	$\bar{2}14$	28.1	27.5	2.14	2.12	024	18.3	23.0

[a] Unobserved reflections are shown in parentheses. [b] Values enclosed in parentheses are estimates for unobserved reflections.

comparison between calculated and observed structure factor amplitudes. The atomic coordinates of the final structure are listed in Table IX.

Discussion

A. Structural Characteristics. As is shown by the data presented in Table VII, the structure deviates only little from $P2_1$ symmetry. Both residues along the chain possess nearly identical conformations, almost the same bridge angles, *virtual bond* lengths, and ϕ,ψ angles. The similarity extends to the intrachain O(3)- - -O(5)′ hydrogen bonds as well as the interchain hydrogen bonds of a given chain. The largest departures from 2_1 symmetry are found in the O(6) rotations: although all four are nominally in the gt position, the rotation angles vary from 49.9° to 68.9°. The positions of the water oxygens are, similarly, close to 2_1 symmetry dictated positions, as indicated by their fractional coordinates in Table VII. It is therefore concluded that even though the crystal structure of Na-cellulose IV does not possess strict $P2_1$ symmetry, it is very close to it. In this respect, its structure is similar to all other cellulose structures which are based on a ~10.3-Å fiber repeat.

As can be further concluded from the chain rotation parameters shown in Table VII (and particularly clearly from the projections of the structure shown in Figure 4), the two 200 planes passing through the unit cell and containing corner and center chains, respectively, are different and exhibit different interchain hydrogen bonds. This difference is undoubtedly caused by the presence of water molecules in the plane of the corner chains. While for the center chains the interchain hydrogen bonds in the 200 plane are O(2)- - -O(6), for the corner chains the corresponding hydrogen bonds are O(2)- - -O(W)- - -O(6) and O(3)- - -O(W)- - -O(6). Additional corner-to-center, O(6)- - -O(3) hydrogen bonds in the 110 plane are also

Table IX
Cartesian Atomic Coordinates of the Final Structure (Å)

corner chain				center chain			
atom	x	y	z	atom	x	y	z
C(1)	−0.218	0.348	3.947	C(1)	4.060	1.413	−5.265
C(2)	−0.178	1.483	2.936	C(2)	4.704	0.478	−4.255
C(3)	−0.602	0.967	1.572	C(3)	4.070	0.685	−2.891
C(4)	0.252	−0.233	1.187	C(4)	4.145	2.156	−2.506
C(5)	0.233	−1.289	2.290	C(5)	3.562	3.037	−3.609
C(6)	1.164	−2.444	2.020	C(6)	3.728	4.511	−3.338
O(2)	−1.046	2.519	3.375	O(2)	4.527	−0.863	−4.694
O(3)	−0.429	1.995	0.599	O(3)	4.767	−0.089	−1.917
O(4)	−0.275	−0.809	0.000	O(4)	3.391	2.359	−1.319
O(5)	0.650	−0.703	3.530	O(5)	4.228	2.766	−4.849
O(6)	0.962	−3.511	2.943	O(6)	3.293	5.314	−4.433
H(1)	−1.200	−0.009	4.058	H(1)	3.041	1.188	−5.376
H(2)	0.802	1.856	2.875	H(2)	5.731	0.689	−4.194
H(3)	−1.612	0.682	1.605	H(3)	3.064	0.383	−2.923
H(4)	1.240	0.078	1.017	H(4)	5.146	2.423	−2.336
H(5)	−0.745	−1.656	2.396	H(5)	2.540	2.821	−3.715
H(6A)	2.159	−2.113	2.079	H(6A)	4.737	4.718	−3.131
H(6B)	0.985	−2.801	1.049	H(6B)	3.157	4.761	−2.494
C(1)′	0.210	−0.337	9.127	C(1)′	4.054	2.221	−10.446
C(2)′	0.168	−1.464	8.106	C(2)′	3.414	3.150	−9.425
C(3)′	0.581	−0.947	6.739	C(3)′	4.040	2.935	−8.058
C(4)′	−0.267	0.255	6.348	C(4)′	3.969	1.465	−7.667
C(5)′	−0.225	1.298	7.465	C(5)′	4.565	0.608	−8.784
C(6)′	−1.155	2.451	7.183	C(6)′	4.399	−0.864	−8.502
O(2)′	1.035	−2.503	8.541	O(2)′	3.588	4.492	−9.859
O(3)′	0.404	−1.988	5.781	O(3)′	3.332	3.718	−7.100
O(4)′	0.265	0.833	5.165	O(4)′	4.729	1.264	−6.483
O(5)′	−0.650	0.718	8.705	O(5)′	3.895	0.869	−10.023
O(6)′	−1.172	3.411	8.238	O(6)′	4.745	−1.686	−9.614
H(1)′	1.192	0.017	9.239	H(1)′	5.072	2.450	−10.557
H(2)′	−0.811	−1.838	8.048	H(2)′	2.387	2.939	−9.367
H(3)′	1.593	−0.665	6.765	H(3)′	5.044	3.240	−8.083
H(4)′	−1.259	−0.048	6.182	H(4)′	2.970	1.189	−7.501
H(5)′	0.755	1.659	7.568	H(5)′	5.586	0.830	−8.887
H(6A)′	−2.128	2.087	7.026	H(6A)′	3.407	−1.057	−8.219
H(6B)′	−0.834	2.929	6.305	H(6B)′	5.025	−1.118	−7.698
			Water Molecules				
W(1)	−0.725	4.335	−9.259	W(2)	0.871	4.496	−3.404

present. All hydroxyl groups participate in hydrogen bonding, and both water molecules of the unit cell take part in four hydrogen bonds each. The structure is thus nearly maximally hydrogen bonded, in keeping with the general structural characteristics seen in crystalline celluloses.

Considerable similarities with the cellulose II structure are apparent, as would be expected on the basis of similar unit cells. For example, in the 200 plane containing the center chains the O(6) positions are gt and the interchain hydrogen bonds are O(2)- - -O(6) in both cellulose II and Na-cellulose IV. Similarly, in the 110 plane the same O(6)- - -O(3) hydrogen bond is present in both structures.

B. Conversion Characteristics and Sheet Formation. The Na-cellulose IV structure is an intermediate one between Na-cellulose II and cellulose II—formed from the former and converted into the latter. Its transformation into cellulose II can take place very easily: the removal of water molecules from between the corner chains causes the corner O(6) atoms to rotate from the gt to the tg positions, thus replacing the hydrogen bonds to water by the O(3)- - -O(6) and O(2)- - -O(6)′ inter- and intrachain bonds of cellulose II.[6] The center chains, which are free of any substantial influence by water, remain essentially the same in this transformation. This, coupled with a decrease in the unit cell parameters, allows the formation of a maximal degree of hydrogen bonding in cellulose II with relative ease.

The transformation of Na-cellulose II to Na-cellulose IV during the formation of the latter is more complex, owing to the threefold helical conformation of the chains in Na-cellulose II.[2] By examining X-ray diffraction patterns recorded at intermediate stages during this conversion—initially showing the presence of both structures, followed by the sharpening of the diffraction diagram of Na-cellulose IV—some interesting features could be discerned. Of particular interest was the observation that the 110 diffraction line of Na-cellulose IV appeared early and with strong intensity and reached a final sharpness quite early during the transformation. By contrast, the $\bar{1}10$ and 200 lines were much less resolved, as can be seen in the equatorial diffraction tracings shown in Figure 5. This observation points out that the crystallite size in the direction normal to the 110 planes is larger than in the other directions, which, in turn, suggests that the formation of the crystallites of Na-cellulose IV proceeds in the following two steps. In the first step, the threefold chain conformation reverts to the twofold one as a result of a decrease in the concentration of NaOH. Sheets begin to form in the $\bar{1}10$ plane. In the second step, stacking of sheets takes place in a roughly perpendicular direction to the $\bar{1}10$ plane, by linking the sheets via hydrogen bonds between the chains and water molecules. This gradually increases the crystallite size in the other directions (perpendicular to $\bar{1}10$ and 200). However, as can be seen in the a–b projection of the Na-cellulose IV structure (cf. Figure 4), there are no direct interchain hydrogen bonds in the $\bar{1}10$ plane. This suggests that the driving force in the initial aggregation of cellulose chains and the formation of a sheet structure is van der Waals force, or hydrophobic

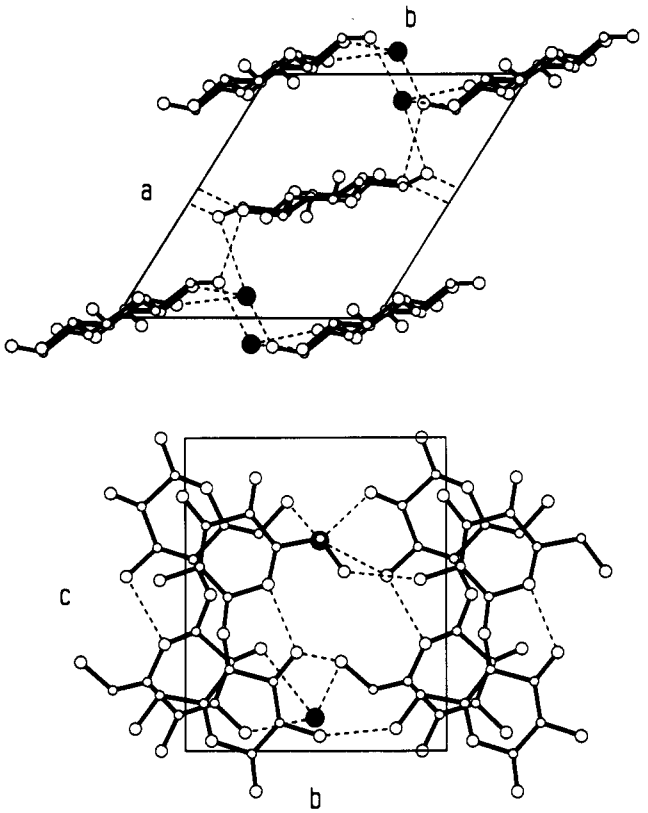

Figure 4. Projections of the Na-cellulose IV unit cell on the (top) *a–b* plane and (bottom) *b–c* plane.

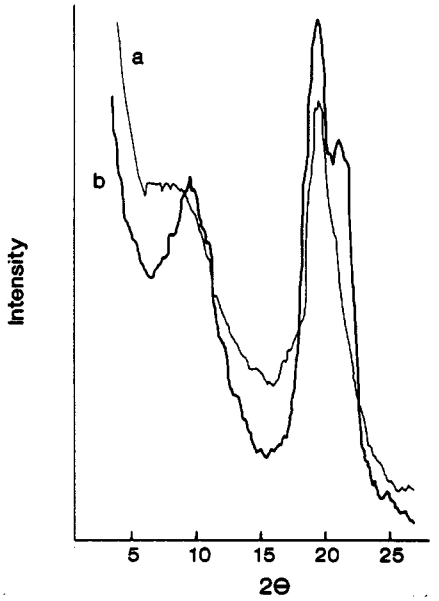

Figure 5. Equatorial intensity profiles of the diffraction patterns of (a) completely formed Na-cellulose IV and (b) Na-cellulose IV in an early stage of formation.

bonds. It is supported by the fact that, in the direction normal to the 110 plane, the hydrocarbon moieties of the chains (C–H) are facing one another, while the hydrophilic groups (O–H) point toward the outside of the sheet, as illustrated in Figure 6.

The similarities in this type of sheet structure extend to all cellulose structures based on the approximate twofold chain conformation. As shown in Figure 7, where the structures of cellulose I, Na-cellulose I, Na-cellulose IV, cellulose II, and Na-cellulose IIB are compared, four of the five structures (with Na-cellulose IIB as the sole exception) contain a sheet in which hydrogen bonding is

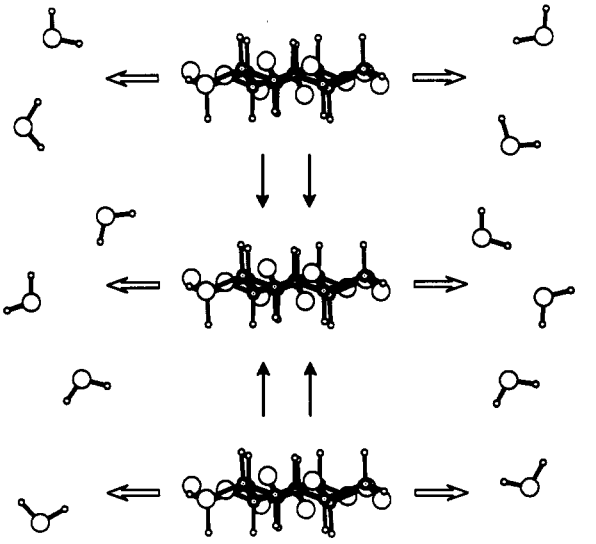

Figure 6. Possible mechanism of sheet formation in the crystallization of celluloses (in the presence of water). Hydrophobic attractions are indicated by solid arrows, and hydrophilic attractions by open arrows.

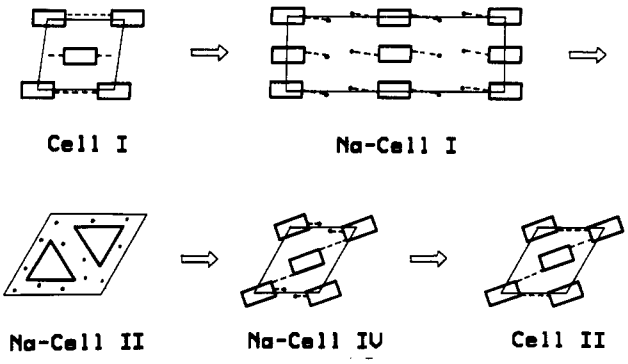

Figure 7. Similar structural characteristics of celluloses I and II and Na-celluloses I, II, and IV. Directions of hydrogen bonding are indicated by dashed lines.

absent. In celluloses I and II and in Na-cellulose IV it is the $1\bar{1}0$ plane; in Na-cellulose I it is the 020 plane. As the formation of all of these structures takes place in the presence of water, it appears that a hydrophobic attraction governs the aggregation of cellulose chains. The twofold chain conformation is obviously favored under such conditions. On the other hand, in Na-cellulose II—which forms only under conditions of high NaOH orientation—the chain conformation is a threefold one, the chains are separated from one another, and there is no sheet structure present. Here the strong interactions between cellulose and NaOH are apparently overriding any other forces with the consequent adoption of a different conformation and chain packing.

The role of the O(6) hydroxyls in the sheet-type structures appears to be one of hydrogen bond facilitator. Both gt and tg rotational positions are probable and can occur simultaneously in the same structure, depending on hydrogen bond requirements. This suggests that both positions would also be present in the amorphous regions of cellulose as well as on the surfaces of crystallites.

Finally, the stacking of chains in all twofold cellulose structures and a nonsheet type structure for a threefold chain conformation both imply that there is a definite connection between the conformation of cellulose chains and their surroundings. Therefore, it may not be adequate to perform conformational energy calculations for isolated cellulose chains and to expect the results of such com-

putations to lead to the definitive identification of global energy minima or the prediction of most probable aggregation characteristics.

Acknowledgment. This study was supported by the National Science Foundation under Grants CHE8107534, DMB8320548, and DMB8703725.

References and Notes

(1) Okano, T.; Sarko, A. *J. Appl. Polym. Sci.* **1984**, *29*, 4175.
(2) Okano, T.; Sarko, A. *J. Appl. Polym. Sci.* **1985**, *30*, 325.
(3) Nishimura, H.; Sarko, A. *J. Appl. Polym. Sci.* **1987**, *33*, 855.
(4) Nishimura, H.; Sarko, A. *J. Appl. Polym. Sci.* **1987**, *33*, 867.
(5) Nishimura, H.; Okano, T.; Sarko, A. *Macromolecules*, preceding paper in this issue.
(6) Stipanovic, A. J.; Sarko, A. *Macromolecules* **1976**, *9*, 851.
(7) Fraser, R. D. B.; MacRae, T. P.; Suzuki, E. *J. Appl. Crystallogr.* **1978**, *11*, 693.
(8) Deslandes, Y.; Marchessault, R. H.; Sarko, A. *Macromolecules* **1980**, *13*, 1466.

Registry No. Sodium cellulose, 9069-34-5.