Mercerization of Cellulose. 5. Crystal and Molecular Structure of Na-Cellulose I†

Hisao Nishimura,‡ Takeshi Okano,§ 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 I, a NaOH–cellulose complex formed by treating cellulose I with aqueous NaOH, has been investigated by X-ray diffraction analysis combined with stereochemical modeling. The structure crystallizes in a four-chain unit cell with dimensions a = 8.83 Å, b = 25.28 Å, c (fiber axis) = 10.29 Å, and all angles 90°. The probable space group is $P2_1$. The unit cell contains eight anhydroglucopyranose residues, eight Na^+ ions, and an undetermined amount of water. For the most likely composition, $C_6H_{10}O_5 \cdot NaOH \cdot 2H_2O$, the calculated density ρ_c = 1.38 g/cm³. An antiparallel arrangement of chains is slightly favored over a parallel one. The chain conformations of the most probable structure differ in the rotational positions of the O(6) hydroxyls: tg in the "up" chains and gt in the "down" chains. All eight Na^+ ions form secondary ionic bonds with O(2) and O(6) hydroxyls, but not with the O(3) hydroxyls.

Introduction

As reported in previous parts of this series, controlled alkali mercerization of ramie cellulose I proceeds through several solid-phase transformations.[1,2] The latter may involve as many as five crystalline alkali–cellulose complexes ("Na-celluloses") which occur as intermediate structures during the complete conversion of cellulose I to cellulose II. Preliminary characterizations of these structures, based on X-ray diffraction analysis, led to a proposal for a mechanism of mercerization.[2-4] According to this mechanism, mercerization begins in the amorphous regions of the fiber and at the interfaces between the cellulose I crystallites and, during the initial, fast phase, involves only the small crystallites. The first crystalline structure observed as a result of the conversion is Na-cellulose I. Because of the characteristics of this conversion step, it is likely that the change from the parallel-chain structure of cellulose I to the antiparallel-chain structure characteristic of cellulose II occurs either during this step or shortly after. To clarify the changes occurring in this transformation, a detailed characterization of the crystal structure of Na-cellulose I was attempted.

Experimental Section

The starting material was, as previously, commercially refined ramie (supplied by Teikoku Bouseki K.K., Japan). The procedures for the preparation of specimens for mercerization and X-ray diffraction studies have been previously described.[1] A sufficient number of fibers were packed into a glass specimen tube (1 mm in diameter) to allow complete conversion of cellulose I to Na-cellulose I to occur in approximately 15 days, using 6.3 N aqueous NaOH solution. X-ray diffraction diagrams were recorded on Kodak DEF-5 X-ray film by placing the sample tube in an evacuated flat-film camera (Warhus) equipped with 0.025-in. pinhole collimators. Typical exposure time was 8 h, with a copper target X-ray tube operated at 40 kV and 25 mA. The films were developed in Kodak Liquid X-Ray Developer for recommended times. The X-ray diffraction diagrams were calibrated with the 2.319-Å diffraction line of NaF. The films were measured for interplanar spacing data with a Supper circular film measuring device. A representative diffraction pattern of Na-cellulose I is shown on the left side of Figure 1.

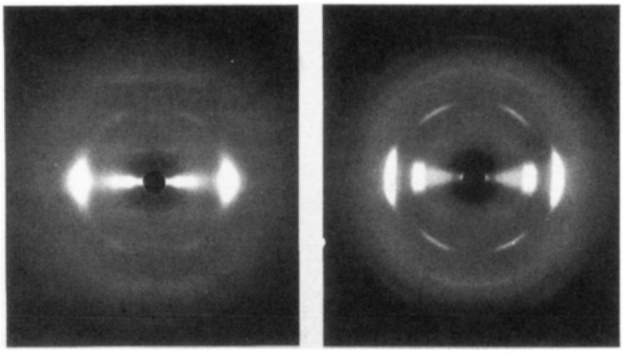

Figure 1. X-ray diffraction patterns of (left) Na-cellulose I and (right) ramie cellulose I treated with a 1.5 N aqueous NaOH solution (swollen but not converted cellulose). Fiber axes are vertical.

Because the diffraction patterns of Na-cellulose I are generally not as well resolved as those of cellulose I and other cellulose polymorphs, the measured intensities were corrected quantitatively for scattering from the NaOH solution and the swollen, amorphous cellulose. Considerable scattering from the latter two components is visible in the diffraction diagrams, as shown on the right side of Figure 1. The layer line intensity profiles of the diffraction diagrams of Na-cellulose I were recorded with a Joyce-Loebl recording microdensitometer, along with the diffraction profiles of the NaOH solution and the swollen cellulose I. After subtraction of the NaOH and amorphous cellulose contributions from the intensity profiles of Na-cellulose I, the latter were resolved (as well as possible) into individual intensity peaks using a least-squares curve resolution program.[5] Predicted, but not observed, reflections were arbitrarily assigned half of the minimum observable intensity at the corresponding diffraction angle. The integrated intensities of each resolved component were corrected for Lorentz and polarization factors,[6] reflection arcing, distance of diffracted ray from specimen to film, film scanning direction if other than radial, and an artificial temperature factor,[7] by use of a computer program specially developed for this purpose.[8] In addition, the intensity data were corrected for absorption by the 6.3 N NaOH solution (linear absorption coefficient 15.5 cm^{-1}) in view of having to use 1-mm-thick sample preparations. (The intensities on higher order layer lines, in particular, showed considerable effects due to absorption.) The experimental structure factor amplitudes, F_o, were then obtained as the square roots of the corrected intensities.

Results

A. X-ray Diffraction Data. The X-ray diffraction pattern of Na-cellulose I exhibits nine clear reflections—

* To whom correspondence should be addressed.

† Part 4 of this series: *J. Appl. Polym. Sci.* **1987**, *33*, 867.

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

§ Permanent address: Department of Forest Products, Faculty of Agriculture, The University of Tokyo, Bunkyo-ku, Tokyo 113, Japan.

0024-9297/91/2224-0759$02.50/0

Table I
Comparison of the Observed and Calculated *d* Spacings[a]

hkl	*d* spacing, Å obsd	calcd
020	12.64	12.64
200	4.40	4.42
220	4.15	4.17
011	9.57	9.53
002	5.13	5.145
102	4.43	4.45
004	2.57	2.57
104	2.49	2.47
204	2.23	2.22

[a] Unit cell parameters a = 8.83 Å, b = 25.28 Å, c = 10.29 Å, and $\alpha = \beta = \gamma = 90°$.

Table II
Calculated Equatorial *d* Spacings for Orthogonal and Monoclinic Subcells (cf. Figure 2)

obsd *d* spacing, Å	F_o[c]	orthogonal subcell[a] *hkl*	orthogonal subcell[a] calcd *d*[d]	monoclinic subcell[b] *hkl*	monoclinic subcell[b] calcd *d*[d]
12.64	18.9	010	12.64	010	12.64
	(4.8)	100	(8.84)		
	(5.0)			100	(8.34)
				$\bar{1}10$	(8.34)
	(5.6)	110	(7.24)		
6.23	8.7	020	6.32	020	6.32
				110	6.10
				$\bar{1}20$	6.10
4.40	105.0	120	5.14	$\bar{2}10$	4.42
		200	4.42	120	4.39
				$\bar{1}30$	4.39
4.15	57.8	030	4.21	030	4.21
		210	4.17	200	4.17
				$\bar{2}00$	4.17
3.87	37.1	130	3.80	210	3.62
		220	3.62	$\bar{2}30$	3.62
	(10.7)			130	(3.34)
				$\bar{1}40$	(3.34)
2.99	33.7	040	3.16	040	3.16
		230	3.05	220	3.05
		140	2.98	$\bar{2}40$	3.05
		300	2.95	$\bar{3}10$	2.93
		310	2.87	$\bar{3}20$	2.93
				300	2.78
				$\bar{3}30$	2.78
	(13.1)	320	(2.67)	140	(2.68)
				$\bar{1}50$	(2.68)
	(13.6)	240	(2.57)	230	(2.57)
				$\bar{2}50$	(2.57)
	(13.7)			310	(2.55)
				$\bar{3}40$	(2.55)
	(13.8)	050	(2.53)	050	(2.53)

[a] Subcell parameters a = 8.83 Å, b = 12.64 Å, and $\gamma = 90°$. [b] Subcell parameters a = 8.83 Å, b = 13.39 Å, and $\gamma = 109.3°$. [c] Estimates for unobserved reflections shown in parentheses. [d] Unobserved reflections shown in parentheses.

three on the equator, one each on the first and second layer lines, two on the fourth layer line, and the 002 and 004 meridional reflections. On the basis of the *d* spacings of these reflections, an orthogonal unit cell with dimensions a = 8.83 Å, b = 25.28 Å, and c (fiber axis) = 10.29 Å was determined by refinement with the *d* spacing data shown in Table I. (The comparison of all observed and calculated *d* spacings is shown in Table VIII, together with the structure factor amplitudes.) It should be noted that even though $\gamma = 90°$ for this unit cell (as determined by least-squares refinement), it is monoclinic rather than orthorhombic, because as will be shown, the probable space group is $P2_1$ (c axis unique).

Although an observed density of Na-cellulose I is not available because this compound is not stable in a dry state, the most likely number of cellulose chains in the unit cell is four (eight glucopyranose residues). The corresponding calculated density, without contributions from NaOH and water, is 0.94 g/cm^3. The most probable placement of the chain axes in the a–b plane of the unit cell is at coordinates (0, 0), (0, b/2), (a/2, 0), and (a/2, b/2). On the basis of the structure analysis described below, the most probable number of Na^+ ions in the unit cell is eight. The resultant most likely molecular composition, $C_6H_{10}O_5 \cdot NaOH \cdot 2H_2O$, leads to a calculated density ρ_{calc} = 1.38 g/cm^3, which is not an unreasonable value.

No distinct odd-order 00l reflections were observed, lending agreement to a probable space group of $P2_1$. However, the diffraction pattern was not sufficiently well resolved to allow a firm determination of the space group on the basis of reflection data alone.

During the course of experimental sample preparations, a diffraction pattern of Na-cellulose I was occasionally observed that exhibited a very weak reflection at $d \simeq 2.8$ Å near the meridian. This reflection could not be indexed with the present unit cell, and to account for it, unrea-

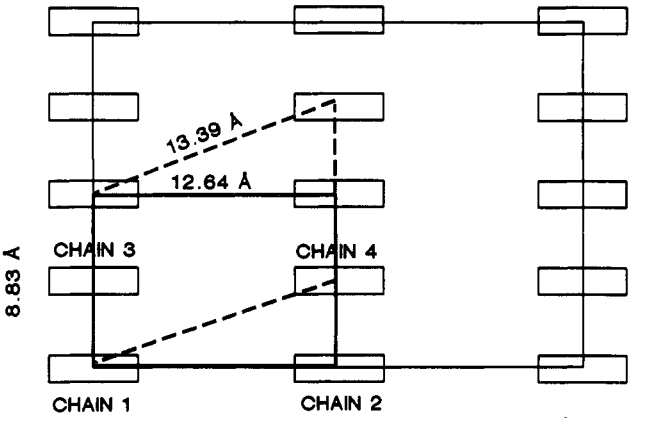

Figure 2. Outlines of the two two-chain subcells (orthogonal subcell with a = 8.83 Å, b = 12.64 Å, and $\gamma = 90°$; monoclinic subcell with a = 8.33 Å, b = 13.39 Å, and $\gamma = 109.3°$), superimposed on the complete four-chain unit cell.

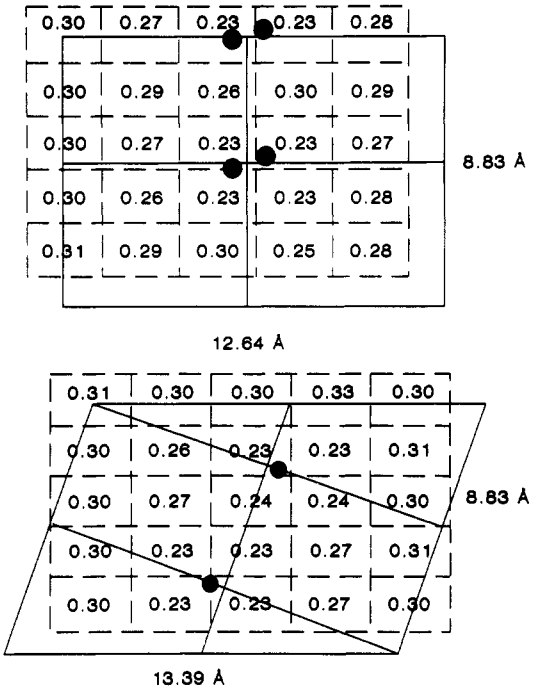

Figure 3. R'' value maps of (top) orthogonal and (bottom) monoclinic subcells for refinement of Na^+ positions within the indicated grid. The most probable location of Na^+ is shown.

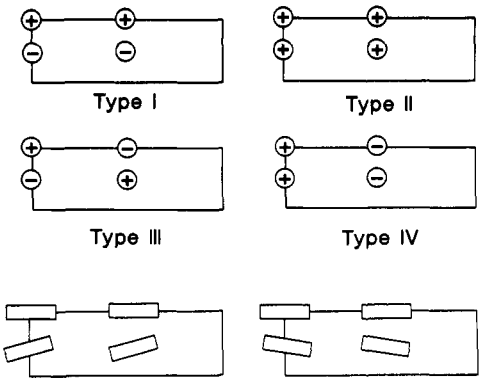

Figure 4. Possible chain-packing models for (top) chain polarity and (bottom) chain rotations.

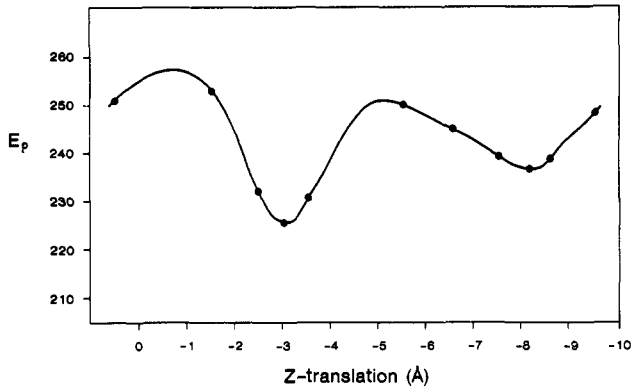

Figure 5. Typical plot of the packing energy E_p as a function of the z translation of the second chain relative to the first chain.

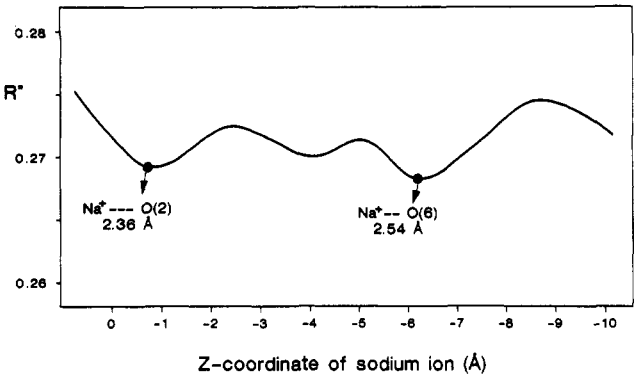

Figure 6. Typical plot of the R'' value as a function of the z coordinate of the Na^+ ion.

sonable unit cells would be generated. It was also observed that such a pattern was more frequently produced when fewer fibers were placed in the specimen tube, causing the conversion to Na-cellulose I to be faster. Under these conditions, the appearance of mixed Na-cellulose I + Na-cellulose IIA patterns had been previously noted.[1,2] The diffraction diagram of Na-cellulose IIA (fiber repeat ca. 15.4 Å) exhibits its strongest reflection at d = 2.87 Å on the fifth layer line near the meridian. Consequently, the appearance of an ~2.8-Å reflection in some patterns of Na-cellulose I was taken as an artifact due to the presence of a small amount of Na-cellulose IIA. (It should be noted that patterns containing this reflection and referenced as Na-cellulose I have been previously described by Hayashi et al.[9])

B. Structure Analysis. Because of the presence of both NaOH and water in the unit cell, the complete structure analysis and refinement were approached in several stages. In the first stage, all probable molecular

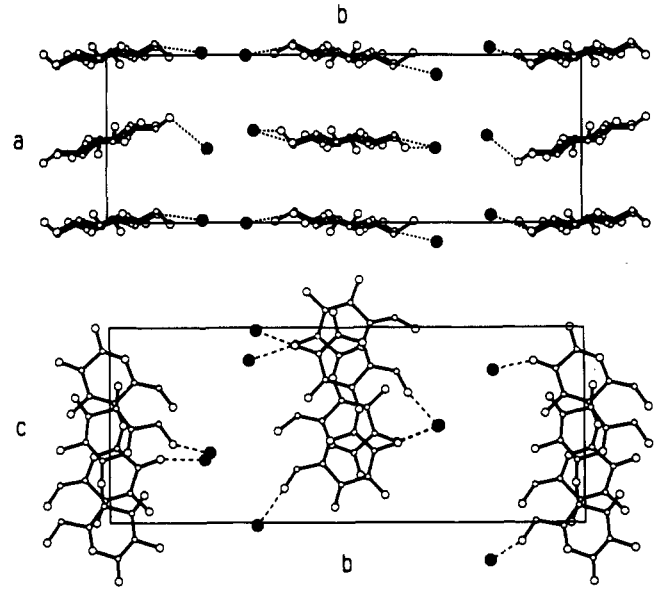

Figure 7. a–b (top) and b–c (bottom) projections of the Na-cellulose I structure. Filled circles denote Na^+ ions; dashed lines denote ionic secondary bonds.

Table III
Summary of Two-Dimensional X-ray Intensity Refinements of Orthogonal and Monoclinic Subcells

chain pol[a]	chain rot.[b]
orthogonal subcell[c]	
chain 1: +	36–38°
chain 3: +	(a) 23–24° or (b) 50–51°
chain 1: +	34–38°
chain 3: –	(a) 129–130° or (b) 155–157°
monoclinic subcell[c]	
chain 1: +	55–59°
chain 3: +	(a) 41–42° or (b) 70–72°
chain 1: +	55–58°
chain 3: –	(a) 147–148° or (b) 176–178°

[a] "+" designates "up" and "–" designates "down" chains (cf. Figure 3). [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] See footnotes a and b of Table II.

models for the isolated chain were identified. This was followed, in the second stage, by refinement against X-ray diffraction intensities in two dimensions, using two-chain subcells and without including NaOH or water. In the third stage, the two-dimensional refinement included all four chains, followed by the inclusion of Na^+ ions. Finally, a full three-dimensional refinement was carried out, using combined stereochemical modeling and X-ray intensity refinement. The following results were obtained, in sequence.

Molecular Models. All probable molecular models for a chain with a dimer residue repeat of 10.29 Å and incorporating a twofold screw axis coincident with the chain axis were determined by using stereochemical refinement procedures. (For the detailed description of procedure, cf. ref. 10.) All models were characterized by a *virtual bond* length of 5.425 Å (the length of the vector connecting successive glycosidic oxygens) and allowed all probable O(6) rotational positions, i.e., tg, gt, and gg. (For the description of the O(6) rotation terminology, cf. ref 11.) The most probable rotational angles marking the positions of O(6) were tg = 166.3°, gt = 64.9°, and gg = –63.3°. (These angles were kept invariant in some of the following refinement steps, as noted.) All models allowed the formation of the usual, intrachain O(3)- - -O(5)′ hydrogen bond, 2.62 Å in length. The tg model allowed

Table IV
Two-Dimensional X-ray Intensity Refinements of the Complete Unit Cell Including Na^+ Ions

chain	chain pol[a]	chain rot.,[b] deg	O(6) position	Na^+ ion positions, Å		
				ion	x	y
		Model 32(b) (R'' = 0.023, CT = 6.75)				
1	+	40.2	tg	$Na^+(1)$	0.01	5.52
2	-	141.4	tg	$Na^+(2)$	4.18	6.07
3	+	53.8	tg	$Na^+(3)$	4.21	6.38
4	-	140.6	tg	$Na^+(4)$	0.22	7.81
				$Na^+(5)$	0.30	18.28
				$Na^+(6)$	4.16	17.87
				$Na^+(7)$	4.40	20.89
				$Na^+(8)$	0.16	21.14
		Model 26(b) (R'' = 0.026, CT = 6.96)				
1	+	40.8	tg	$Na^+(1)$	-0.09	5.92
2	-	144.3	tg	$Na^+(2)$	4.12	5.87
3	-	161.4	gt	$Na^+(3)$	4.44	6.42
4	+	35.9	gt	$Na^+(4)$	0.20	8.04
				$Na^+(5)$	0.03	18.53
				$Na^+(6)$	4.21	17.63
				$Na^+(7)$	4.38	20.95
				$Na^+(8)$	0.43	21.17
		Model 18(b) (R'' = 0.026, CT = 5.56)				
1	+	33.0	tg	$Na^+(1)$	-0.67	4.29
2	+	43.8	gt	$Na^+(2)$	4.22	4.33
3	+	36.8	gt	$Na^+(3)$	3.96	7.38
4	+	57.2	tg	$Na^+(4)$	-0.06	6.90
				$Na^+(5)$	-0.15	17.30
				$Na^+(6)$	4.67	18.82
				$Na^+(7)$	4.46	19.77
				$Na^+(8)$	-0.29	20.13
		Model 35(b) (R'' = 0.028, CT = 6.73)				
1	+	41.9	gt	$Na^+(1)$	-0.39	5.68
2	-	144.9	tg	$Na^+(2)$	4.25	5.80
3	+	54.6	tg	$Na^+(3)$	4.20	6.53
4	-	140.4	gt	$Na^+(4)$	0.36	8.00
				$Na^+(5)$	0.39	18.62
				$Na^+(6)$	4.11	17.85
				$Na^+(7)$	4.53	21.08
				$Na^+(8)$	0.39	21.15
		Model 15(a) (R'' = 0.028, CT = 2.93)				
1	+	33.2	gt	$Na^+(1)$	0.56	4.26
2	+	37.2	tg	$Na^+(2)$	4.58	4.91
3	+	15.1	tg	$Na^+(3)$	4.12	7.57
4	+	38.2	gt	$Na^+(4)$	-0.13	7.10
				$Na^+(5)$	0.02	17.53
				$Na^+(6)$	4.61	17.32
				$Na^+(7)$	4.07	20.00
				$Na^+(8)$	-0.16	20.30
		Model 28(a) (R'' = 0.029, CT = 1.01)				
1	+	32.0	tg	$Na^+(1)$	0.09	4.96
2	-	145.1	gt	$Na^+(2)$	4.25	4.45
3	-	119.1	gt	$Na^+(3)$	4.29	7.59
4	+	38.6	tg	$Na^+(4)$	-0.11	7.83
				$Na^+(5)$	0.13	17.47
				$Na^+(6)$	4.52	17.73
				$Na^+(7)$	4.05	20.70
				$Na^+(8)$	-0.02	20.16
		Model 25(a) (R'' = 0.029, CT = 3.81)				
1	+	32.3	gt	$Na^+(1)$	0.31	4.26
2	-	148.7	tg	$Na^+(2)$	4.12	4.78
3	-	122.6	tg	$Na^+(3)$	4.36	7.27
4	+	37.5	gt	$Na^+(4)$	-0.10	7.21
				$Na^+(5)$	0.07	17.58
				$Na^+(6)$	4.55	17.01
				$Na^+(7)$	4.05	19.98
				$Na^+(8)$	-0.09	20.21
		Model 33(b) (R'' = 0.029, CT = 4.43)				
1	+	41.6	gt	$Na^+(1)$	-0.25	6.04
2	-	141.4	gt	$Na^+(2)$	4.80	4.97
3	+	58.8	tg	$Na^+(3)$	4.66	7.70
4	-	141.3	tg	$Na^+(4)$	0.13	7.91
				$Na^+(5)$	-0.20	17.70
				$Na^+(6)$	4.56	18.03

Table IV (Continued)

chain	chain pol[a]	chain rot.,[b] deg	O(6) position	Na⁺ ion positions, Å: ion	x	y
				$Na^+(7)$	5.06	21.07
				$Na^+(8)$	0.13	20.56
		Model 15(b) (R'' = 0.030, CT = 3.39)				
1	+	40.6	gt	$Na^+(1)$	−0.28	6.27
2	+	36.7	tg	$Na^+(2)$	4.39	5.13
3	+	57.7	tg	$Na^+(3)$	4.79	7.90
4	+	36.9	gt	$Na^+(4)$	0.29	7.57
				$Na^+(5)$	−0.02	17.73
				$Na^+(6)$	4.62	17.22
				$Na^+(7)$	4.95	20.43
				$Na^+(8)$	0.30	20.58
		Model 5(a) (R'' = 0.030, CT = 1.15)				
1	+	33.4	gt	$Na^+(1)$	0.31	4.46
2	+	36.3	tg	$Na^+(2)$	4.64	4.79
3	−	121.7	tg	$Na^+(3)$	4.38	7.89
4	−	143.0	gt	$Na^+(4)$	−0.07	7.38
				$Na^+(5)$	−0.31	20.60
				$Na^+(6)$	4.22	20.16
				$Na^+(7)$	4.62	17.49
				$Na^+(8)$	0.02	17.64
		Model 23(a) (R'' = 0.032, CT = 1.35)				
1	+	32.0	gt	$Na^+(1)$	0.21	4.63
2	−	146.8	gt	$Na^+(2)$	4.56	5.32
3	−	122.5	tg	$Na^+(3)$	4.33	8.03
4	+	35.4	tg	$Na^+(4)$	−0.56	7.39
				$Na^+(5)$	0.04	17.96
				$Na^+(6)$	4.54	17.23
				$Na^+(7)$	4.24	20.46
				$Na^+(8)$	−0.19	20.80
		Model 14(a) (R'' = 0.038, CT = 1.99)				
1	+	33.1	gt	$Na^+(1)$	0.62	4.27
2	+	37.9	tg	$Na^+(2)$	4.74	4.88
3	+	16.9	gt	$Na^+(3)$	4.34	7.84
4	+	37.4	tg	$Na^+(4)$	−0.06	7.11
				$N1^+(5)$	0.17	17.81
				$Na^+(6)$	4.39	17.48
				$Na^+(7)$	4.28	20.24
				$Na^+(8)$	−0.10	20.50

[a] See footnote *a* of Table III. [b] See footnote *b* of Table III.

Table V
Parameter Set for the Calculation of Nonbonded Energy[a]

interaction	d°_{ij}, Å	w_{ij}
O–O	3.60	3.0
O–C	3.60	3.0
O–H	3.25	1.4
C–C	3.70	3.0
C–H	3.30	1.35
H–H	3.20	0.50
O–O (H bond)	2.80	20.0
Na^+–Na^+	3.70	3.0
Na^+–O	3.60	3.0
Na^+–C	3.70	3.0
Na^+–H	3.30	1.35
Na^+–O (ionic bond)	2.40	10.0

[a] See eq 4.

the formation of an additional intramolecular hydrogen bond, O(2)- - -O(6)′, with a length of 2.86 Å.

Two-Dimensional Refinement: Two-Chain Subcells. As shown in Figure 2, two different two-chain subcells that were in accord with the observed equatorial reflections could be described within the complete four-chain cell: *subcell A*, termed "orthogonal", and *subcell B*, termed "monoclinic". Both subcells were used for the initial refinement of the chain positions in the *a*–*b* plane, using all three chain models and both parallel and antiparallel chain packing, for a total of six packing models. In the course of the refinement, the ring components of the monomer residues remained invariant in positions constrained by a twofold molecular screw axis, while only the chain rotations about the twofold axis and all four O(6) rotations were allowed to vary independently. The equatorial reflections given in Table II were used in these refinements. The weighted X-ray residual

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

was used as the refinement criterion, with the weights w_i set equal to 1 for observed reflections, 0.5 for unobserved reflections when $F_c > F_o$, or 0 when $F_c < F_o$. For overlapped reflections, F_c was obtained by using

$$|F_i|_c = [\sum |F_n|_c^2]^{1/2} \quad (2)$$

where the summation was taken over the n reflections contributing to the given intensity peak. The usual, unweighted residual

$$R = \frac{\sum (|F_i|_o - |F_i|_c|}{\sum |F_i|_o} \quad (3)$$

was also calculated in the process.

Because the structure contained an undefined amount of water molecules whose positions in the lattice were thought not to be fixed, their contribution to the scattering factors was taken into account by using water-weighted

Table VI
X-ray Refinement for the c-Axis Translation of Chains

model	c-axis transl, Å		chain pol and conf model[a]
model I	chain 1	0.0	model 28(a) type
R'' = 0.284	chain 2	−3.21	antiparallel
E = 25.7[b]	chain 3	−2.53	CT = 1.01
B = 20.1 [c]	chain 4	2.42	
model II	chain 1	0.0	model 28(a) type
R'' = 0.287	chain 2	−7.43	antiparallel
E = 20.6	chain 3	−3.02	CT = 1.01
B = 21.8	chain 4	−8.97	
model III	chain 1	0.0	model 28(a) type
R'' = 0.287	chain 2	−3.06	antiparallel
E = 20.9	chain 3	−3.08	CT = 1.01
B = 20.4	chain 4	−4.42	
model IV	chain 1	0.0	model 28(a) type
R'' = 0.287	chain 2	−0.39	antiparallel
E = 28.4	chain 3	−1.42	CT = 1.01
B = 20.0	chain 4	2.54	
model V	chain 1	0.0	model 28(a) type
R'' = 0.287	chain 2	−4.37	antiparallel
E = 30.3	chain 3	−1.41	CT = 1.01
B = 19.8	chain 4	−8.81	
model VI	chain 1	0.0	model 28(a) type
R'' = 0.288	chain 2	−1.57	antiparallel
E = 22.0	chain 3	−2.72	CT = 1.01
B = 20.3	chain 4	−2.70	
model VII	chain 1	0.0	model 15(a) type
R'' = 0.289	chain 2	−4.82	parallel
E = 27.6	chain 3	3.46	CT = 2.93
B = 19.8	chain 4	−6.36	
model VIII	chain 1	0.0	model 18(b) type
R'' = 0.289	chain 2	0.94	parallel
E = 28.9	chain 3	−2.61	CT = 5.56
B = 18.4	chain 4	1.64	
model IX	chain 1	0.0	model 15(b) type
R'' = 0.289	chain 2	−0.76	parallel
E = 33.3	chain 3	0.99	CT = 3.39
B = 19.3	chain 4	−3.51	
model X	chain 1	0.0	model 28(a) type
R'' = 0.289	chain 2	−8.49	antiparallel
E = 23.9	chain 3	−2.78	CT = 1.01
B = 20.0	chain 4	−2.58	
model XI	chain 1	0.0	model 25(a) type
R'' = 0.289	chain 2	−2.50	antiparallel
E = 23.5	chain 3	−3.32	CT = 3.81
B = 18.4	chain 4	2.87	
model XII	chain 1	0.0	model 33(b) type
R'' = 0.289	chain 2	0.83	antiparallel
E = 24.9	chain 3	−2.21	CT = 4.43
B = 19.7	chain 4	−2.09	
model XIII	chain 1	0.0	model 33(b) type
R'' = 0.289	chain 2	−2.45	antiparallel
E = 34.0	chain 3	1.13	CT = 4.43
B = 19.8	chain 4	−5.35	

[a] See Table IV for the model characteristics. [b] The nonbonded packing energy term of eq 4, restricted to interchain contacts. [c] Isotropic temperature factor.

atomic scattering factors.[12] For the calculation of these scattering factors, the amount of water displaced by hydrogen, carbon, and oxygen were taken as 5.15, 16.44, and 9.13 Å^3, respectively. (Similarly, water-weighted atomic scattering factors for the Na^+ ion were calculated with a displacement volume of 3.59 Å^3. The latter value was calculated from the crystal ionic radius for Na^+.[13])

Initial results obtained with the six models quickly eliminated the O(6) gg models, as the O(6) atoms tended to rotate away from the gg positions in all cases. This left eight models to be evaluated in both two-chain subcells: all O(6) tg, all O(6) gt, a mixed model with O(6) tg in the chain at (0, 0) and O(6) gt in the chain at (a/2, 0), and a similar gt + tg mixed model, all in parallel- and antiparallel-packing modes. The results obtained while keeping the O(6) positions constant and letting only the rotational positions of the chains vary are summarized in Table III. As shown, two probable chain rotation positions were found for each model, indicated by (a) and (b), respectively, in Table III. Although comparisons based on the R factors were at this stage meaningless, antiparallel models with mixed O(6) rotations (gt in one chain, tg in the other) were slightly favored in both subcells. On the basis of these refinements it was clear, however, that there were only two likely chain-packing schemes for the complete, four-chain unit cell in both packing polarities.

The probable positions of the Na^+ ions in the a–b plane were determined next by dividing the base plane of both subcells into 25 equal rectangular areas, 0.9 × 1.3 Å^2 for the orthogonal subcell and 0.85 × 1.35 Å^2 for the monoclinic subcell. A Na^+ ion was placed in each of the rectangular areas in turn, and its position was refined in the x and y directions, keeping the chain positions invariant and using only equatorial intensity data. The results, shown in Figure 3 for both subcells, clearly indicated the most probable positions for the Na^+ ions. Four minima were predicted for the orthogonal subcell, while for the monoclinic subcell there were two minima. No differences between parallel and antiparallel models were noted; however, the orthogonal subcell was favored over the monoclinic subcell because the former suggested better symmetry for the four-chain unit cell. The four predicted locations of Na^+ ions for the orthogonal subcell were consistent with $P2_1$ symmetry for the complete unit cell, but the same was not true for the monoclinic subcell. The case for the $P2_1$ symmetry was also strengthened by the observation that the 001 and 003 meridional reflections were both absent. As a consequence, the orthogonal subcell was picked at this point as the more probable basis for the complete unit cell model. The number of Na^+ ions in the four-chain cell was, consequently, expected to be eight.

Two-Dimensional Refinement: Complete Unit Cell. The refinements were carried out with 64 initial models within four chain-packing polarities shown in Figure 4. For each packing model, two chain rotation models as previously determined ((a) and (b); see preceding section and Table III) and eight different O(6) rotation mixes were the possible choices. The different O(6) mixes were all tg, all gt, and all six possible combinations of two gt + two tg in the four chains. Additionally, each model contained eight Na^+ ions in their most probable positions (see preceding section). The chain rotations and the (x, y) coordinates of the Na^+ ions were refined (against equatorial intensities), with the O(6) rotations being held constant at the respective most probable gt and tg positions.

In addition to the R'' factor, another index—termed CT—was defined for use as a refinement criterion. This index measured the closeness of the Na^+ ions to the positions related by twofold rotation axes, which are the two-dimensional equivalents of twofold screw axes in $P2_1$ symmetry. The value of CT is zero when $P2_1$ symmetry is strictly observed, and it becomes larger with increasing deviations from this symmetry. This criterion was adopted for two reasons: (1) because of a small number of equatorial reflections and a relatively large number of variables and (2) because all of the most probable models obtained in previous refinement steps favored the presence of $P2_1$ symmetry in the four-chain unit cell. (The definition of the CT index is given in the Appendix.)

As an arbitrary boundary value for R'' separating the more probable models from the less probable, R'' = 0.03 (equatorial reflections only) was chosen. Additionally,

Table VII
Characteristics of the Final Antiparallel and Parallel Models

characteristic	antiparallel model	parallel model
glycosidic bridge angle, deg	116.6	116.6
virtual bond,[a] Å	5.425	5.425
chain 1[b]		
chain pol	up	up
rot. angle,[c] deg	32.2	33.2
c-axis transl, Å	0.0	0.0
O(6) rot. angles[d] (deg): residues 1 and 2	166.7, 165.2	25.5, 92.4
chain 2		
chain pol	down	up
rot. angle, deg	148.1	39.8
c-axis transl, Å	−7.80	−4.64
O(6) rot. angles (deg): residues 3 and 4	88.6, 76.2	169.8, 160.6
chain 3		
chain pol	down	up
rot. angle, deg	121.4	23.7
c-axis transl, Å	−3.06	3.23
O(6) rot. angles (deg): residues 5 and 6	56.5, 56.8	160.1, 168.7
chain 4		
chain pol	up	up
rot. angle, deg	39.4	46.3
c-axis transl, Å	−9.47	−6.38
O(6) rot. angles (deg): residues 7 and 8	165.0, 171.3	85.3, 46.6
x,*y*,*z* coordinates of Na^+ ions, Å		
$Na^+(1)$	−0.14, 5.08, −6.96	−0.04, 6.59, −1.46
$Na^+(2)$	4.87, 5.35, −6.59	0.32, 18.74, −6.87
$Na^+(3)$	3.91, 7.83, −0.16	5.81, 4.63, −0.01
$Na^+(4)$	0.03, 7.44, −1.73	4.27, 19.96, −4.52
$Na^+(5)$	1.03, 17.57, −5.18	0.18, 4.98, −6.23
$Na^+(6)$	4.85, 17.50, −5.22	−0.33, 20.28, −2.39
$Na^+(7)$	4.15, 20.25, −1.98	3.72, 8.57, −1.26
$Na^+(8)$	−0.41, 20.43, −2.25	4.15, 16.87, −6.25
CT index (including *z* coordinate)	5.51	5.92
R″	0.131	0.129
R	0.147	0.157
anisotropic temp factor components: B_x, B_y, B_z	1.7, 34.9, 34.0	2.5, 27.8, 35.0
short contacts and lengths,[e,f] Å	none	$O(6)_1$–$O(5)_1$ = 2.51
packing energy[g] (arb units)	22.2	27.1
conf energy[h] (arb units)	63.5	65.1
interchain hydrogen bonds	none	none
intrachain hydrogen bonds[f] and lengths, Å	$O(3)$–$O(5)$ = 2.62	$O(3)$–$O(5)$ = 2.62
	$O(6)_2$–$O(2)_1$ = 2.88	$O(6)_4$–$O(2)_3$ = 2.98
	$O(6)_8$–$O(2)_7$ = 2.77	$O(6)_6$–$O(2)_5$ = 2.81
	$O(6)_1$–$O(2)_2$ = 2.85	$O(6)_3$–$O(2)_4$ = 2.79
	$O(6)_7$–$O(2)_8$ = 2.89	$O(6)_5$–$O(2)_6$ = 2.99
ionic secondary bonds[f] and lengths, Å	$O(2)_1$–$Na^+(1)$ = 2.42	$O(6)_6$–$Na^+(3)$ = 2.49
	$O(2)_2$–$Na^+(8)$ = 2.38	$O(6)_5$–$Na^+(4)$ = 2.49
	$O(6)_5$–$Na^+(2)$ = 2.62	$O(2)_1$–$Na^+(5)$ = 2.46
	$O(6)_6$–$Na^+(7)$ = 2.44	$O(2)_2$–$Na^+(6)$ = 2.50
	$O(2)_8$–$Na^+(3)$ = 2.31	$O(6)_7$–$Na^+(7)$ = 2.54
	$O(6)_7$–$Na^+(3)$ = 2.47	$O(6)_8$–$Na^+(8)$ = 2.31
	$O(2)_7$–$Na^+(6)$ = 2.37	$O(3)_8$–$Na^+(7)$ = 2.17
	$O(6)_8$–$Na^+(6)$ = 2.39	$O(3)_7$–$Na^+(8)$ = 2.41
	$O(2)_3$–$Na^+(4)$ = 2.70	
	$O(2)_4$–$Na^+(5)$ = 2.46	

[a] Length of the vector connecting successive glycosidic oxygens. [b] See Figure 3 for chain numbering. [c] See footnote *b* of Table III. [d] 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°. [e] Distances less than 2.60 Å for O---O, 2.80 Å for O---C, and 2.20 Å for O---H and C---H contacts. [f] Subscripts indicate residue number. [g] The nonbonded term of eq 4, restricted to interchain contacts. [h] The nonbonded term of eq 4, restricted to intrachain contacts.

when *R*″ was >0.03 and <0.04 and CT was lower than 2.0, the model was also included among the more probable ones.

A total of 12 models surviving these selections are shown in Table IV. Included are four parallel and eight antiparallel models. All of these were selected for further refinement in three dimensions, performed in the next step.

Three-Dimensional Refinement: Chains. Because chains 1 and 2 (cf. Figure 2) are sufficiently isolated from one another (as are chains 3 and 4), nonbonded repulsion within these pairs of chains was expected to be absent. On the other hand, repulsion could occur between chains 1 and 3 (as well as 2 and 4). Therefore, prior to further X-ray refinement, repulsion-free positions for these pairs of chains were determined by stereochemical refinement. In this refinement, the energy *E* of the model, as calculated from the following expression, is minimized:

$$E = \sum\left(\frac{a_i - a_i^\circ}{\mathrm{SD}_i}\right)^2 + \frac{1}{W^2}\sum w_{ij}(d_{ij} - d_{ij}^\circ)^2 \qquad (4)$$

(In this equation, a_i are the values of the bonded parameter bond lengths, bond angles, and torsion angles, with a_i° their standard values. The d_{ij} and d_{ij}° are, respectively, the nonbonded contact distances and their equilibrium

Table VIII
Observed and Calculated Structure Factor Amplitudes for the Final Antiparallel Model, $R'' = 0.131$ (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					
	(25.28)	010	*c*			(3.34)	$1\bar{7}0$	(35.8)	8.1
12.64	12.64	020	63.2	56.0		(3.34)	$\bar{1}70$		
	(8.84)	100	(16.0)	7.1		(3.33)	$2\bar{5}0$	(35.8)	31.4
	(8.43)	030	(16.7)	6.5		(3.33)	$\bar{2}50$		
	(8.34)	$1\bar{1}0$	(16.7)	19.4	2.99	3.16	080	113	107
	(8.34)	$\bar{1}10$				3.05	$2\bar{6}0$		
	(7.24)	$1\bar{2}0$	(18.7)	23.2		3.05	$\bar{2}60$		
	(7.24)	$\bar{1}20$				2.98	$1\bar{8}0$		
6.23	6.32	040	29.1	31.2		2.98	$\bar{1}80$		
	6.10	$1\bar{3}0$				2.95	300		
	6.10	$\bar{1}30$				2.93	$3\bar{1}0$		
4.40	5.14	$1\bar{4}0$	351	350		2.93	$\bar{3}10$		
	5.14	$\bar{1}40$				2.87	$3\bar{2}0$		
	5.06	050				2.87	$\bar{3}20$		
	4.42	200				2.81	090		
	4.39	$1\bar{5}0$				2.80	$2\bar{7}0$		
	4.39	$\bar{1}50$				2.80	$\bar{2}70$		
	4.35	$2\bar{1}0$				2.78	$3\bar{3}0$		
	4.35	$\bar{2}10$				2.78	$\bar{3}30$		
4.15	4.21	060	193	198		(2.68)	$1\bar{9}0$	(43.8)	4.5
	4.17	$2\bar{2}0$				(2.68)	$\bar{1}90$		
	4.17	$\bar{2}20$				(2.67)	$3\bar{4}0$	(43.8)	54.9
3.87	3.91	$2\bar{3}0$	124	123		(2.67)	$\bar{3}40$		
	3.91	$\bar{2}30$				(2.57)	$2\bar{8}0$	(45.5)	13.4
	3.80	$1\bar{6}0$				(2.57)	$\bar{2}80$		
	3.80	$\bar{1}60$				(2.55)	$3\bar{5}0$	(45.8)	35.2
	3.62	$2\bar{4}0$				(2.55)	$\bar{3}50$		
	3.62	$\bar{2}40$				(2.53)	0,10,0	(46.1)	11.6
	3.61	070							
				First Layer Line					
	(10.29)	001				(3.42)	$2\bar{4}1$	(20.4)	33.0
9.57	9.53	011	14.4	46.5		(3.42)	$\bar{2}41$		
	7.98	021				(3.41)	071	(20.4)	7.7
6.20	6.71	101	17.7	28.1		(3.18)	$1\bar{7}1$	(22.1)	21.9
	6.52	031				(3.18)	$\bar{1}71$		
	6.48	$1\bar{1}1$				(3.17)	$2\bar{5}1$	(22.1)	40.4
	6.48	$\bar{1}11$				(3.17)	$\bar{2}51$		
	5.92	$1\bar{2}1$				(3.02)	081	(23.4)	21.0
	5.92	$\bar{1}21$				(2.92)	$2\bar{6}1$	(24.1)	34.0
	(5.39)	041	(12.4)	37.0		(2.92)	$\bar{2}61$		
	(5.25)	$1\bar{3}1$	(12.7)	24.3		(2.86)	$1\bar{8}1$	(24.7)	14.3
	(5.25)	$\bar{1}31$				(2.86)	$\bar{1}81$		
	4.60	$1\bar{4}1$	*d*			(2.83)	301	(25.1)	32.0
	4.60	$\bar{1}41$				(2.82)	$3\bar{1}1$	(25.1)	37.0
	4.54	051	*d*			(2.82)	$\bar{3}11$		
	4.06	201	*d*			(2.76)	$3\bar{2}1$	(25.7)	17.3
	4.04	$1\bar{5}1$	*d*			(2.76)	$\bar{3}21$		
	4.04	$\bar{1}51$				(2.71)	091	(26.1)	19.6
	4.01	$2\bar{1}1$	*d*			(2.70)	$2\bar{7}1$	(26.1)	11.8
	4.01	$\bar{2}11$				(2.70)	$\bar{2}71$		
	3.90	061	*d*			(2.69)	$3\bar{3}1$	(26.4)	32.9
	3.87	$2\bar{2}1$	*d*			(2.69)	$\bar{3}31$		
	3.87	$\bar{2}21$				(2.59)	$1\bar{9}1$	(26.4)	30.7
3.62	3.66	$2\bar{3}1$	27.7	38.2		(2.59)	$\bar{1}91$		
	3.66	$\bar{2}31$				(2.59)	$3\bar{4}1$	(26.4)	31.3
	3.57	$1\bar{6}1$				(2.59)	$\bar{3}41$		
	3.57	$\bar{1}61$							
				Second Layer Line					
5.13	5.145	002	*e*			2.96	$2\bar{4}2$		
	5.04	012	*f*			2.96	$\bar{2}42$		
	(4.77)	022	(5.7)	24.4		2.96	072		
4.43	4.45	102	37.4	43.1		(2.80)	$1\bar{7}2$	(21.7)	18.0
	4.39	032				(2.80)	$\bar{1}72$		
	4.38	$1\bar{1}2$				(2.79)	$2\bar{5}2$	(22.1)	27.6
	4.38	$\bar{1}12$				(2.79)	$\bar{2}52$		
4.20	4.19	$1\bar{2}2$	27.4	30.4		(2.69)	082	(23.1)	2.0
	4.19	$\bar{1}22$			2.53	2.62	$2\bar{6}2$	43.1	66.2
	(3.99)	042	(11.4)	20.0		2.62	$\bar{2}62$		
	(3.93)	$1\bar{3}2$	(12.0)	31.5		2.58	$1\bar{8}2$		
	(3.93)	$\bar{1}32$				2.58	$\bar{1}82$		
3.21	3.64	142	65.5	84.4		2.56	302		

Table VIII (Continued)

d spacing, Å					*d* spacing, Å				
obsd	calcd[a]	*hkl*	F_o[b]	F_c	obsd	calcd[a]	*hkl*	F_o[b]	F_c
	3.64	$\bar{1}42$				2.54	$\bar{3}12$		
	3.61	052				2.54	$\bar{3}12$		
	3.35	202				2.51	$\bar{3}22$		
	3.34	$\bar{1}52$				2.51	$\bar{3}22$		
	3.34	$\bar{1}52$				2.47	092		
	3.32	$\bar{2}12$				2.46	$\bar{2}72$		
	3.32	$\bar{2}12$				2.46	$\bar{2}72$		
	3.26	062				2.45	$\bar{3}32$		
	3.24	$\bar{2}22$				2.45	$\bar{3}32$		
	3.24	$\bar{2}22$				(2.38)	$\bar{1}92$	(27.1)	12.9
	3.12	$\bar{2}32$				(2.38)	$\bar{1}92$		
	3.12	$\bar{2}32$				(2.37)	$\bar{3}42$	(27.1)	37.7
	3.06	$\bar{1}62$				(2.37)	$\bar{3}42$		
	3.06	$\bar{1}62$							
				Third Layer Line					
	(3.43)	003				2.55	$\bar{1}63$		
	3.40	013	*c*			2.55	$\bar{1}63$		
3.25	3.31	023	53.5	57.7		2.49	$\bar{2}43$		
	3.20	103				2.49	$\bar{2}43$		
	3.18	033				2.49	073		
	3.17	$\bar{1}13$				(2.39)	$\bar{1}73$	(22.4)	23.1
	3.17	$\bar{1}13$				(2.39)	$\bar{1}73$		
	3.10	$\bar{1}23$				(2.39)	$\bar{2}53$	(22.4)	27.1
	3.10	$\bar{1}23$				(2.39)	$\bar{2}53$		
	3.02	043			2.26	2.32	083	51.5	51.0
	2.99	$\bar{1}33$				2.28	$\bar{2}63$		
	2.99	$\bar{1}33$				2.28	$\bar{2}63$		
	(2.85)	$\bar{1}43$	(14.0)	16.8		2.25	$\bar{1}83$		
	(2.85)	$\bar{1}43$				2.25	$\bar{1}83$		
	(2.84)	053	(14.4)	11.1		2.24	303		
2.68	2.71	203	31.1	39.7		2.23	$\bar{3}13$		
	2.70	$\bar{1}53$				2.23	$\bar{3}13$		
	2.70	$\bar{1}53$				2.20	$\bar{3}23$		
	2.69	$\bar{2}13$				2.20	$\bar{3}23$		
	2.69	$\bar{2}13$				(2.17)	093	(26.4)	11.9
	2.66	063				(2.17)	$\bar{2}73$	(26.4)	27.6
	2.65	$\bar{2}23$				(2.17)	$\bar{2}73$		
	2.65	$\bar{2}23$				(2.16)	$\bar{3}33$	(26.7)	20.5
2.55	2.58	$\bar{2}33$	39.1	41.7		(2.16)	$\bar{3}33$		
	2.58	$\bar{2}33$							
				Fourth Layer Line					
2.57	2.57	004	*e*			2.29	054		
	2.56	014	*f*			2.22	204		
2.49	2.52	024	*f*			2.22	$\bar{1}54$		
	2.47	104				2.22	$\bar{1}54$		
	2.46	034				2.22	$\bar{2}14$		
	2.46	$\bar{1}14$				2.22	$\bar{2}14$		
	2.46	$\bar{1}14$				2.20	064		
	(2.42)	$\bar{1}24$	(10.0)	7.0		2.19	$\bar{2}24$		
	(2.42)	$\bar{1}24$				2.19	$\bar{2}24$		
	(2.38)	044	(11.7)	6.0		2.15	$\bar{2}34$		
	(2.37)	$\bar{1}34$	(12.0)	6.9		2.15	$\bar{2}34$		
	(2.37)	$\bar{1}34$				2.13	$\bar{1}64$		
2.23	2.30	$\bar{1}44$	59.5	45.6		2.13	$\bar{1}64$		
	2.30	$\bar{1}44$							

[a] Unobserved reflections shown in parentheses. [b] Estimates for unobserved reflections shown in parentheses. [c] Out of measurable range. [d] Unmeasurable because of overlap with a strong equatorial reflection. [e] Meridional reflection, unmeasurable on fiber diagrams. [f] Unmeasurable because of difficulty in separating from the meridional reflection.

values; the SD_i, W, and w_{ij} are appropriate standard deviations and weights.[10] For values of d_{ij}° and w_{ij}, see Table V.)

A two-chain orthogonal subcell was employed as before, and the only parameters that were refined were the translational positions of the chains along the *c* axis. Sodium ions were not included in the refinement. As shown in Figure 5 for one of the 12 models, in a plot of the packing energy E_p as a function of *z* translation (E_p was calculated from the nonbonded term of eq 4), a single minimum was usually obtained (with some models showing two energy minima). No hydrogen bonding between the chains are predicted in any of the models.

The *c*-axis translations of chains 2–4 relative to chain 1 were then determined with X-ray refinement using three-dimensional intensity data. Only the *c*-axis positions of chains 2–4 were allowed to vary while the relative translational shifts between chains 1 and 3 (and 2 and 4) were restricted to the vicinity determined by stereochemical refinement. Sodium ions were again not included in the refinement.

The criterion of $R'' < 0.290$ was arbitrarily chosen as the indicator of probable models. From the overall results for all models, 13 were thus chosen as the more probable models (cf. Table VI). Of these, 10 were antiparallel and three were parallel chain models. However, a further

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

atom	x	y	z	atom	x	y	z
	Chain 1				Chain 2		
C(1)	−0.129	0.387	3.940	C(1)	−0.130	12.254	−11.742
C(2)	0.154	1.497	2.940	C(2)	0.148	11.142	−10.742
C(3)	−0.363	1.089	1.573	C(3)	−0.367	11.553	−9.375
C(4)	0.197	−0.271	1.181	C(4)	0.198	12.911	−8.983
C(5)	−0.061	−1.300	2.274	C(5)	−0.056	13.941	−10.076
C(6)	0.608	−2.626	1.992	C(6)	0.618	15.264	−9.794
O(2)	−0.476	2.691	3.372	O(2)	−0.487	9.951	−11.174
O(3)	0.021	2.073	0.617	O(3)	0.013	10.567	−8.419
O(4)	−0.458	−0.728	0.000	O(4)	−0.455	13.370	−7.802
O(5)	0.481	−0.831	3.510	O(5)	0.485	13.469	−11.312
O(6)	−0.057	−3.343	0.954	O(6)	−0.089	16.353	−10.383
H(1)	−1.164	0.256	4.060	H(1)	−1.165	12.388	−11.862
H(2)	1.190	1.655	2.884	H(2)	1.184	10.980	−10.686
H(3)	−1.411	1.032	1.605	H(3)	−1.415	11.614	−9.408
H(4)	1.229	−0.192	1.005	H(4)	1.230	12.827	−8.808
H(5)	−1.095	−1.446	2.379	H(5)	−1.089	14.091	−10.182
H(6A)	0.619	−3.206	2.867	H(6A)	1.601	15.243	−10.165
H(6B)	1.599	−2.446	1.695	H(6B)	0.654	15.408	−8.755
C(1′)	0.129	−0.387	9.085	C(1′)	0.130	13.026	−16.887
C(2′)	−0.154	−1.497	8.085	C(2′)	−0.148	14.138	−15.887
C(3′)	0.363	−1.089	6.718	C(3′)	0.367	13.727	−14.520
C(4′)	−0.197	0.271	6.326	C(4′)	−0.198	12.369	−14.128
C(5′)	0.061	1.300	7.419	C(5′)	0.056	11.339	−15.221
C(6′)	−0.608	2.626	7.137	C(6′)	−0.618	10.016	−14.939
O(2′)	0.476	−2.691	8.517	O(2′)	0.487	15.329	−16.319
O(3′)	−0.021	−2.073	5.762	O(3′)	−0.013	14.713	−13.564
O(4′)	0.458	0.728	5.145	O(4′)	0.455	11.910	−12.947
O(5′)	−0.481	0.831	8.655	O(5′)	−0.485	11.811	−16.457
O(6′)	0.077	3.357	6.123	O(6′)	−0.087	8.970	−15.749
H(1′)	1.164	−0.256	9.205	H(1′)	1.165	12.892	−17.007
H(2′)	−1.190	−1.655	8.029	H(2′)	−1.184	14.300	−15.831
H(3′)	1.411	−1.032	6.750	H(3′)	1.415	13.666	−14.552
H(4′)	−1.229	0.192	6.150	H(4′)	−1.230	12.453	−13.953
H(5′)	1.095	1.446	7.524	H(5′)	1.089	11.189	−15.327
H(6A′)	−0.641	3.196	8.018	H(6A′)	−1.650	10.105	−15.113
H(6B′)	−1.591	2.445	6.816	H(6B′)	−0.468	9.770	−13.930
	Chain 3				Chain 4		
C(1)	4.476	−0.404	−7.000	C(1)	4.340	13.040	−5.525
C(2)	5.224	−1.272	−6.000	C(2)	4.760	14.106	−6.525
C(3)	4.579	−1.136	−4.633	C(3)	4.196	13.766	−7.892
C(4)	4.474	0.331	−4.421	C(4)	4.580	12.346	−8.284
C(5)	3.785	1.137	−5.334	C(5)	4.195	11.358	−7.191
C(6)	3.793	2.622	−5.052	C(6)	4.691	9.958	−7.473
O(2)	5.192	−2.621	−6.432	O(2)	4.286	15.369	−6.093
O(3)	5.362	−1.846	−3.677	O(3)	4.701	14.694	−8.848
O(4)	3.685	0.448	−3.060	O(4)	3.873	11.976	−9.465
O(5)	4.480	0.958	−6.570	O(5)	4.792	11.755	−5.955
O(6)	3.216	3.366	−6.123	O(6)	3.917	9.317	−8.485
H(1)	3.491	−0.748	−7.120	H(1)	3.297	13.041	−5.405
H(2)	6.222	−0.951	−5.944	H(2)	5.808	14.132	−6.582
H(3)	3.616	−1.553	−4.666	H(3)	3.149	13.842	−7.860
H(4)	5.434	0.720	−4.066	H(4)	5.614	12.295	−8.460
H(5)	2.794	0.807	−5.440	H(5)	3.150	11.343	−7.086
H(6A)	4.780	2.943	−4.894	H(6A)	4.656	9.389	−6.591
H(6B)	3.237	2.802	−4.180	H(6B)	5.688	10.014	−7.798
C(1′)	4.362	0.404	−12.145	C(1′)	4.498	12.240	−0.380
C(2′)	3.614	1.272	−11.145	C(2′)	4.078	11.174	−1.380
C(3′)	4.259	1.136	−9.778	C(3′)	4.642	11.514	−2.747
C(4′)	4.364	−0.331	−9.386	C(4′)	4.258	12.934	−3.139
C(5′)	5.053	−1.137	−10.479	C(5′)	4.643	13.922	−2.046
C(6′)	5.045	−2.622	−10.197	C(6′)	4.147	15.322	−2.328
O(2′)	3.646	2.621	−11.577	O(2′)	4.552	9.911	−0.948
O(3′)	3.476	1.846	−8.822	O(3′)	4.137	10.586	−3.703
O(4′)	5.153	−0.448	−8.205	O(4′)	4.965	13.304	−4.320
O(5′)	4.358	−0.958	−11.715	O(5′)	4.046	13.525	−0.810
O(6′)	5.628	−3.365	−11.265	O(6′)	4.827	15.910	−3.435
H(1′)	5.347	0.748	−12.265	H(1′)	5.541	12.239	−0.260
H(2′)	2.616	0.951	−11.089	H(2′)	3.030	11.148	−1.437
H(3′)	5.222	1.553	−9.811	H(3′)	5.689	11.438	−2.715
H(4′)	3.404	−0.720	−9.211	H(4′)	3.224	12.985	−3.315
H(5′)	6.044	−0.807	−10.585	H(5′)	5.688	13.937	−1.941

Table IX (Continued)

atom	x	y	z	atom	x	y	z
	Chain 3				Chain 4		
H(6A′)	4.057	−2.944	−10.044	H(6A′)	4.281	15.920	−1.476
H(6B′)	5.597	−2.801	−9.322	H(6B′)	3.121	15.277	−2.549
				Sodium Ions			
$Na^+(1)$	−0.135	5.084	−6.956	$Na^+(2)$	4.866	5.350	−6.592
$Na^+(3)$	3.905	7.834	−0.163	$Na^+(4)$	0.025	7.435	−1.734
$Na^+(5)$	1.026	17.572	−5.177	$Na^+(6)$	4.847	17.501	−5.220
$Na^+(7)$	4.147	20.253	−1.981	$Na^+(8)$	−0.414	20.429	−2.247

selection was made on the basis of packing energy and the CT index, which eliminated all but three antiparallel models (I, II, and III) and one parallel model (VII). These four models constituted the final selection from which the best two models to be subjected to a complete, final structure refinement would be chosen.

Three-Dimensional Refinement: Na^+ Ions. The z coordinate of each Na^+ ion was refined independently, using R'' as the refinement criterion, keeping the chain positions invariant. Generally, two or three minima were obtained for each Na^+ ion, as shown in Figure 6. The most interesting result was that the probable Na^+ positions were most often found to be near the O(2) and O(6) hydroxyls.

Final Refinement. On the basis of all previous results, one antiparallel and one parallel model were selected for final refinement (models II and VII in Table VI, respectively). In the initial cycles of this refinement, all chain rotations and the (x, y, z) coordinates of all Na^+ ions were refined, using the minimization of the function

$$\Phi = fR'' + (1 - f)E \qquad (5)$$

as the refinement criterion. The use of this function—in which E is the total conformational and packing energy, R'' is expressed as a percentage, and the fraction f is chosen to weight the E contribution just sufficiently (typically, f = 0.90–0.95)—is necessitated to ward off the possible development of nonbonded repulsions in the model. (When the number of intensities used in X-ray refinement is small, optimal stereochemistry and R'' factor minima may not always coincide.)

In this final refinement, hydrogen bonds were defined and included in the energy term when the distance between the oxygens fell in the range 2.6–3.0 Å. Likewise, a Na^+- - -O contact was considered a secondary ionic bond when the distance was less than 2.8 Å. (The $d°$ = 2.40 Å parameter for this contact was arbitrarily determined from the ionic crystal radius of 0.95 Å for Na^+ and the van der Waals radius of 1.40 Å for O.[13] The weight parameter w_{ij} for this contact was taken as half of that for a hydrogen bond.)

Because of an already large number of variables in this refinement, the ring components of the monomer residues remained invariant at their most probable conformations as determined from single-chain refinement. A twofold molecular screw axis thus remained in place controlling the positions of the anhydroglucopyranose rings along the c axis. It did not control the rotational positions of the O(6) atoms.

The results of the final refinement for the two models, in terms of structural characteristics, are given in Table VII. The comparisons of calculated and observed structure factor amplitudes are shown in Table VIII. The atomic coordinates of the final structure are listed in Table IX.

Discussion and Conclusions

As can be seen in Table VII, both the R and R'' factors for the two best models are virtually identical, thus not allowing a preference to be drawn for either packing polarity. In terms of stereochemical structural characteristics, however, there are several differences, all of which favor the antiparallel structure. For example, the packing energy for the antiparallel model is 22.2 (in arbitrary units), which is more acceptable than the value of 27.1 for the parallel structure. (Generally, it has been noted that for polysaccharide crystal structures an upper limit in the range 20–25 for the energy value describes the limit of acceptable stereochemistry.) The parallel structure also shows one marginally short hydrogen bond (O(6)- - -O(5) = 2.51 Å) while the antiparallel structure is free of short contacts. Perhaps more importantly, the Na^+ ion positioning is considerably better in the antiparallel model. In the latter, all Na^+ ions participate in secondary bond formation with O(2) and O(6) hydroxyls, while in the parallel model two of the eight Na^+ ions do not show such bond formation. In the antiparallel structure, the positions of Na^+ ions are also more closely described by $P2_1$ symmetry (CT = 5.51), which is not the case with the parallel model (CT = 5.92). The 2_1 symmetry also extends to the O(6) atom positions within each chain of the antiparallel model. As shown in Table VII, chains 1 and 4 possess nearly perfect tg positions for the O(6) atoms, while chains 2 and 3 show a similarly good gt positioning. In the parallel model, considerable deviations from both the most probable positions and 2_1 symmetry are evident. Therefore, even though on the basis of pure diffraction analysis it is not possible to identify the most probable model for the structure of Na-cellulose I, the antiparallel structure has a decided edge over the parallel model on the basis of more stable stereochemistry and better symmetry.

We thus conclude that Na-cellulose I is most likely an antiparallel chain crystal structure described (very closely) by the space group $P2_1$. An antiparallel structure is also more probable on the basis of observations that when a mixed cellulose I + Na-cellulose I structure (i.e., both crystal structures present simultaneously) is washed free of alkali and dried, a mixture of cellulose I and cellulose II structures is obtained.[2] The mechanism proposed by us earlier[2,3] for cellulose mercerization, in which the conversion from a parallel to an antiparallel structure occurs in the very first step, is thus supported by the findings of the present study. This is not to say that a parallel Na-cellulose structure could not form. In fact, it is likely that it may form under conditions when the conversion is rapid. However, given time to "anneal", the partially (or even substantially) parallel structure should spontaneously convert to the more stable antiparallel form.

The proposed crystal structure of Na-cellulose I (cf. Figure 7) differs considerably from that of cellulose I.[10] The hydrogen-bonded sheet structure of the latter is

disrupted—there are no intermolecular hydrogen bonds in Na-cellulose I—and contacts between chains are stabilized by interactions through Na^+ ions (and, presumably, water). A different type of sheet structure—in the 200 plane—is thus formed, having chains of both polarity and the Na^+ ions present in that sheet. In turn, these sheets are close packed in the *a* direction of the unit cell, as in cellulose I, with stabilization provided only through van der Waals forces.

It is interesting to note that the Na^+ ions apparently seek out the O(2) and O(6) hydroxyls only. This is in agreement with the results of studies on the distribution of substituents in partially derivatized cellulose ethers prepared from alkali–cellulose which show that the reactivity order of the three hydroxyls in cellulose is O(2) > O(6) > O(3).[14] The higher reactivity of O(2) relative to O(3) has been ascribed to the higher acidity of O(2) caused by its proximity to the anomeric center at C(1).[15] However, the higher reactivity of O(2) relative to O(6) has not been explained.

Another interesting feature of the structure of Na-cellulose I concerns its two different chain conformations. Two of the four chains have O(6) atoms in the tg position, while the other two chains have O(6) gt. In this respect, the antiparallel Na-cellulose I structure is similar to cellulose II, in which the same O(6) mixture is found.[16]

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

Appendix: Definition of the CT Index

In the case of $P2_1$ symmetry (c axis unique), there are two general equivalent positions with coordinates (x, y, z) and $(-x, -y, {}^1/_2 + z)$. Therefore, for the two equivalent atoms

$$\sum x_i = \sum y_i = 0 \quad \text{or} \quad \sum x_i + \sum y_i = 0 \qquad \text{(A.1)}$$

The eight Na^+ ions in the unit cell can be regarded as being made up of four pairs of equivalent ions:

$$(x_1, y_1, z_1) \text{ and } (-x_1 + \Delta x_1, -y_1 + \Delta y_1, {}^1/_2 + z_1 + \Delta z_1)$$

$$(x_2, y_2, z_2) \text{ and } (-x_2 + \Delta x_2, -y_2 + \Delta y_2, {}^1/_2 + z_2 + \Delta z_2)$$

$$(x_3, y_3, z_3) \text{ and } (-x_3 + \Delta x_3, -y_3 + \Delta y_3, {}^1/_2 + z_3 + \Delta z_3)$$

$$(x_4, y_4, z_4) \text{ and } (-x_4 + \Delta x_4, -y_4, + \Delta y_4, {}^1/_2 + z_4 + \Delta z_4) \qquad \text{(A.2)}$$

where Δx, Δy, and Δz are deviations from the exact equivalent position. Therefore, CT can be defined as follows:

$$\text{CT} = \sum\{|\sum x_{ij}| + |\sum y_{ij}|\} = \sum|\Delta x_j| + \sum|\Delta y_j| \qquad \text{(A.3)}$$

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) Sarko, A. FORTRAN computer program LSQ based on: Fraser, R. D. B.; Suzuki, E. *Anal. Chem.* **1966**, *38*, 1770.
(6) Cella, R. J.; Lee, B.; Hughes, R. E. *Acta Crystallogr., Sect. A* **1970**, *26*, 118.
(7) Franklin, R. E.; Gosling, R. G. *Acta Crystallogr.* **1953**, *6*, 678.
(8) Sarko, A. FORTRAN computer program FIBRXRAY.
(9) Hayashi, J.; Yamada, T.; Kimura, K. *Appl. Polym. Symp.* **1976**, *28*, 713.
(10) Woodcock, C.; Sarko, A. *Macromolecules* **1980**, *13*, 1183.
(11) Sarko, A.; Marchessault, R. H. *J. Polym. Sci., Part C* **1969**, *28*, 317.
(12) Fraser, R. D. B.; MacRae, T. P.; Suzuki, E. *J. Appl. Crystallogr.* **1978**, *11*, 693.
(13) Pauling, L. The Nature of the Chemical Bond; Cornell University Press: Ithaca, NY, 1948.
(14) Isogai, A.; Ishizu, A.; Nakano, J. *Sen-i Gakkaishi* **1984**, *40*, T-504.
(15) Haines, A. H. *Adv. Carbohydr. Chem.* **1976**, *33*, 11.
(16) Stipanovic, A. J.; Sarko, A. *Macromolecules* **1976**, *9*, 851.

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