Packing Analysis of Carbohydrates and Polysaccharides. 6. Molecular and Crystal Structure of Regenerated Cellulose II

Arthur J. Stipanovic and Anatole Sarko*

Department of Chemistry, SUNY College of Environmental Science and Forestry, Syracuse, New York 13210. *Received May 19, 1976*

ABSTRACT: The crystal structure of regenerated cellulose II (Fortisan) has been solved by a combined stereochemical packing refinement and x-ray diffraction analysis. The structure is based on a two-chain unit cell with antiparallel chain packing and an extensive hydrogen bond network. The latter links the chains in two directions into sheets and further by intersheet hydrogen bonds into a three-dimensionally stabilized structure. The hydrogen bonding is maximized by unequal rotations of the O(6) hydroxymethyl groups in the two chains of the unit cell. The most probable structure predicted by packing analysis, using the criteria of minimization of packing energy and maximization of hydrogen bonding, is identical in all details with the structure refined by x-ray analysis. The presence of weak odd-order meridionals in the x-ray diffraction diagram may be due to rotational disorder of the hydroxymethyl groups. The R index for the final structure is 0.15, which indicates the reliability that can be obtained with this method of structure refinement.

Cellulose, the major structural component of higher plants, is known to crystallize in at least four polymorphic forms, commonly known as celluloses I, II, III, and IV.[1] Of the four, celluloses I and II are the two industrially most utilized forms. As such, interest in their crystal structures has remained high for many decades but it was only recently that the crystal structure of the native cellulose I polymorph was determined.[2,3] The most interesting finding of these studies was that the biosynthetic cellulose crystallizes with parallel chain polarity.

Cellulose II apparently does not exist naturally but can only be obtained from the native material either by a solution regeneration process or by mercerization which involves a swelling treatment with sodium hydroxide. However, there is one report in the literature indicating that cellulose II may exist natively in one form of algae.[4]

Because cellulose II crystallizes upon regeneration from solution, it has been suggested that the resulting crystal structure is based on antiparallel chain packing, in analogy with synthetic polymers similarly crystallized. This view is in accord with the predictions that of the four cellulose polymorphs, antiparallel cellulose II is the lowest energy form, mainly because of an extensive hydrogen bonding network.[5] Recent electron microscopic experiments on the conversion of cellulose I to II appear to indicate that chain folding may occur during this process.[6]

As a part of our ongoing program in the characterization of cellulose polymorphy and in order to verify the predictions of stability, we undertook to determine the crystal structure of the best available crystalline cellulose II. The commercially regenerated cellulose II obtained from Celanese Corp., tradenamed Fortisan, was found to be the best sample for this purpose. An important objective of this study was to approach the structure determination by two parallel methods. One method was to be a stereochemical structure prediction based on known stereochemical features of polysaccharides and least energy packing principles, and the other method was to be structure refinement against x-ray diffraction intensities. The methods used for this purpose have been previously developed in this laboratory and used with success in detailed crystal structure analysis of polysaccharides.[7–10] While this study was in progress, a preliminary x-ray structure characterization of regenerated cellulose II was published by Kolpak and Blackwell.[11] The results of the present work and that published by the above authors agree in most respects, although not in all details.

Experimental Section

The x-ray diffraction diagrams of Fortisan were significantly improved by first treating the fibers with 4% H_2SO_4 for about 2–4 h at room temperature, followed by a rinsing sequence of very dilute NaOH and water. This treatment reduced the magnitude of amorphous scattering. The x-ray diagrams were then recorded on flat-film multiple-sheet packs of Kodak No-Screen film or Ilford Type G Industrial X-Ray film, either in an evacuated pin-hole camera or in a Searle x-ray camera with toroidal focusing. In both instances, nickel filtered Cu K_α radiation was used. Both types of films were developed in Kodak Liquid X-Ray Developer. The Ilford film was found more suitable for the measurement of diffraction intensities. A typical x-ray diffractogram is shown in Figure 1.

A total of 41 reflections were observed of which 37 were sufficiently resolved for use in a least-squares refinement procedure of the unit cell parameters. The diffractograms were scanned with a Joyce-Loebl recording microdensitometer and the resulting layer line tracings were resolved into individual integrated intensities with a least-squares curve resolution computer program. A total of 33 envelopes were resolved in this manner which included the contributions of 69 *hkl* planes. Not observed were five planes predicted to lie within the 33 intensity envelopes and these were arbitrarily given a value of one-half of the minimum resolvable intensity in the corresponding region of the diffraction angle. The integrated intensities were corrected for Lorentz[12] and polarization factors, reflection arcing, distance of diffracted ray to film, and tracing direction other than radial[9] and were then converted to relative structure amplitudes.

Results and Discussion

Diffraction Measurements. The least-squares refinement of unit cell parameters with 37 reflections resulted in a monoclinic cell of dimensions a = 9.09 Å, b = 7.96 Å, c (fiber axis) = 10.31 Å and the unique angle γ = 117.3°. These dimensions are in good agreement with those previously published.[1,11] The density calculated for this unit cell, 1.62 g/cm^3, is in satisfactory agreement with the density usually observed for regenerated cellulose which is generally less than 50% crystalline. A comparison of the observed and calculated d spacings is shown in Table I. The x-ray diffractogram revealed no obvious systematic extinctions and therefore a space group assignment could not be made. As explained below, the space group was assumed to be $P2_1$ despite the presence of weak odd-order meridional reflections.

Packing Analysis. The objectives of chain packing analysis were essentially twofold. The first was to determine the minimum energy conformation and packing models for subsequent refinement against x-ray data. The second objective was to determine how well the packing method would predict the structure that would be obtained after x-ray refinement.

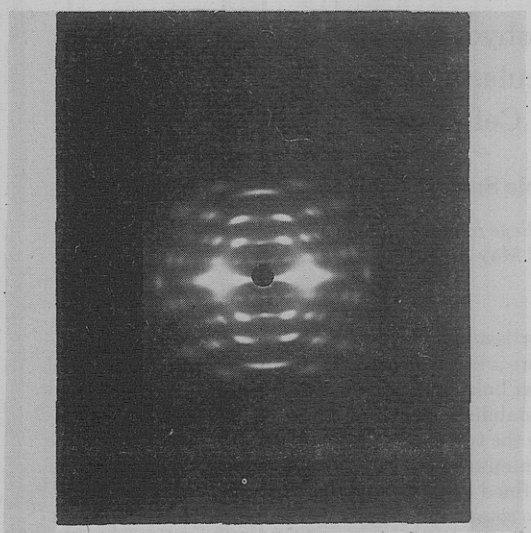


Figure 1. X-ray diffraction diagram of Fortisan. Fiber axis is vertical.

A packing method used with these objectives in mind has been under development in this laboratory and has previously been successfully applied to the structure determination of other polysaccharides.[7,8] Its main premise is that a crystal structure occupies an energy minimum and that the latter can be found by refining a suitable starting model of the structure against accepted stereochemical criteria. The latter consist of bond lengths, bond and conformation angles, and nonbonded contacts. This is accomplished by minimizing the function

$$Y = \sum_{i=1}^{l} \left(\frac{r_i - r_{0i}}{\mathrm{SD}_i{}^r}\right)^2 + \sum_{i=1}^{m} \left(\frac{\theta_i - \theta_{0i}}{\mathrm{SD}_i{}^\theta}\right)^2 + \sum_{i=1}^{n} \left(\frac{\phi_i - \phi_{0i}}{\mathrm{SD}_i{}^\phi}\right)^2 + \frac{1}{W^2} \sum_{\substack{i=1 \\ j=1}}^{N} w_{ij}(d_{ij} - d_{0ij})^2 \quad (1)$$

where r_i, θ_i, *and* ϕ_i *are the* bond lengths, bond angles, and conformation angles, respectively, for the l, m, and n corresponding variables in the monomer residue, while r_{0i}, θ_{0i}, ϕ_{0i} and $SD_i{}^r$, $SD_i{}^\theta$, $SD_i{}^\phi$ are their standard values and standard deviations, respectively. The first three terms of this function thus represent the bond length, bond angle, and conformation angle strains occurring during the adjustment of these variables. The remaining term approximates the nonbonded repulsion with d_{ij} the distance between the nonbonded atoms i and j, d_{0ij} the corresponding equilibrium distance, w_{ij} the weight assigned to the contact in the summation, and W the weight of the nonbonded contact term in the entire function Y. The values of the standard parameters r_{0i}, θ_{0i}, ϕ_{0i} and their standard deviations have been obtained by averaging over known carbohydrate crystal structures.[13] Similarly, the values for d_{0ij} and w_{ij} were previously determined from the prediction of known crystal structures of monosaccharides.[14] A complete discussion of the methods and strategy of packing analysis using the above energy function has been previously published.[7]

To begin the packing analysis, the conformation of a single, isolated chain of cellulose was examined. The restrictions imposed by a unit cell of finite dimensions were therefore removed and a probable starting chain model for cellulose was subjected to the refinement procedure. As previously determined, a probable model is characterized by a twofold screw axis in the fiber direction with a repeat distance of approximately 10.3 Å and with the cellobiose residue in the Hermans conformation.[2] The following parameters were allowed to vary during this stage of analysis: the rotation of the C(6) hydroxymethyl group, the virtual bond length of the residue (i.e., the distance between successive glycosidic oxygens), the glycosidic bridge angle, as well as all bond lengths, bond angles, and conformation angles within one glucose residue. The limits on these variations have been previously published.[7,13] Hydrogen bonds were defined when appropriate oxygen–oxygen contact distances were in the range 2.60–3.0 Å.

As was previously found in a ϕ–ψ type of chain conformation analysis, but with rigid residue rings, three reasonable chain models exist for cellulose, differing mainly in the rotational position of the hydroxymethyl group.[2] The three models are with the O(6) in the staggered tg, gt, and gg positions, respectively. Although all three models proved to be of reasonably low conformational energy, the O(6) tg model allowed the formation of an additional intramolecular O(2)–O(6′) hydrogen bond while the gt or gg models only permitted the usual O(3)–O(5′) intrachain hydrogen bond. (Some of the conformational features of these models will be compared later with the corresponding values found after packing and x-ray refinements.)

The packing analysis was thus performed with six initial models: the three conformations each packed with parallel and antiparallel chain polarity. In the first stage of the refinement the ring components of the monomer residues remained invariant (i.e., bond lengths, bond angles, and ring conformation angles) while the chains themselves were permitted to rotate about their respective helix axes as well as translate relative to one another along the c axis of the unit cell. Also, the glycosidic bridge angle was allowed to adjust by rotation of the residue about the virtual bond and the O(6) rotation was permitted to vary within limits near the tg, gt, or gg position. After suitable packing positions were found in this fashion, the resulting structures were refined with all bond lengths, bond angles, and conformation angles variable, in addition to the above parameters.

During the course of the packing refinement several other models were examined based on the suggestion by Jones that the O(6) rotational positions on the corner and center chains of the unit cell may differ.[15] A preliminary analysis indicated that the O(6) tg + gt (corner chain + center chain) models of both polarities were significantly lower in packing energy by a factor of 2–3 over any other possible O(6) rotationally "mixed" models. As a result, only this mixed model was subjected to complete refinement procedures.

The results of the packing analysis are shown in Table II. There are three criteria that can be applied to the evaluation of the probability of each model: (1) the "packing energy", PE; (2) the number and nature of short nonbonded contacts; (3) the number and characteristics of the hydrogen bonds. Clearly indicated is that the antiparallel models (except gg) are of lower packing enegy than the corresponding parallel models. In fact, none of the parallel models was acceptable in this regard. Of the four antiparallel models, the O(6) gg could similarly be eliminated by unfavorable packing energy. The remaining three antiparallel models, O(6) tg, O(6) gt, and O(6) tg + gt, were all of practically identical packing energy. But as shown in Table III, the first two models each had one very short nonbonded contact, whereas the mixed O(6) model was much more acceptable in this regard.

Similarly, applying the criterion of maximizing the hydrogen bond network, the O(6) tg model could be ruled out and the mixed O(6) model was preferred over the O(6) gt model.

At this stage of structure analysis, the two surviving best packing models would have been further refined against x-ray intensity data. However, since we were interested in a comparative analysis by both stereochemical and x-ray refinement methods, all eight models were subjected to x-ray refinement.

Table II
Probable Models of Cellulose II Predicted by Packing Analysis

Model[a]	Bridge angle, deg	Chain rotations,[b] deg		Center chain translation,[b] Å	PE[c]	Hydrogen bonds and lengths,[d] Å
		Corner	Center			
			Parallel			
O(6) tg (165°)	117.3	34	36	−2.53	28.3	$O(3)_1$–$O(6)_1$, $O(3)_2$–$O(6)_2$ 2.54 $O(3)_3$–$O(6)_3$, $O(3)_4$–$O(6)_4$ 2.52 O(5)–O(3′) 2.75 O(2)–O(6′) 2.70
O(6) gt (50°)	116.9	34	33.5	−2.08	40.2	$O(2)_1$–$O(6)_1$, $O(2)_2$–$O(6)_2$ 2.66 $O(2)_3$–$O(6)_3$, $O(2)_4$–$O(6)_4$ 2.67 $O(6)_1$–$O(2)_3$, $O(6)_2$–$O(2)_4$ 2.77 $O(3)_1$–$O(6)_3$, $O(3)_2$–$O(6)_4$ 2.85 O(5)–O(3′) 2.81
O(6) gg (−60°)	116.6	48	50	1.98	42.8	$O(2)_2$–$O(6)_4$, $O(2)_1$–$O(6)_3$ 2.55 O(5)–O(3′) 2.75
O(6) tg (162°) + gt (83°)	115.9	38	33	−2.79	34.3	$O(3)_1$–$O(6)_1$, $O(3)_2$–$O(6)_2$ 2.54 $O(3)_3$–$O(6)_3$, $O(3)_4$–$O(6)_4$ 2.90 $O(2)_1$–$O(6)_1$, $O(2)_2$–$O(6)_2$ 2.55 $O(5)_1$–$O(3')_2$ 2.81 O(2)–O(6′) 2.64
			Antiparallel			
O(6) tg (160°)	116.7	30	50	−4.56	22.0	$O(3)_1$–$O(6)_1$, $O(3)_2$–$O(6)_2$ 2.62 $O(3)_3$–$O(6)_3$, $O(3)_4$–$O(6)_4$ 2.84 O(5)–O(3′) 2.77 O(2)–O(6′) 2.74
O(6) gt (36°)	115.4	41	72	−3.16	22.6	$O(2)_1$–$O(6)_1$, $O(2)_2$–$O(6)_2$ 2.66 $O(2)_3$–$O(6)_3$, $O(2)_4$–$O(6)_4$ 2.61 $O(2)_1$–$O(2)_3$, $O(2)_2$–$O(2)_4$ 2.79 $O(3)_2$–$O(6)_3$, $O(3)_1$–$O(6)_4$ 2.64 $O(6)_1$–$O(6)_3$, $O(6)_2$–$O(6)_4$ 2.69 O(5)–O(3′) 2.80
O(6) gg (−45°)	116.5	48	48	−3.98	44.2	$O(6)_1$–$O(6)_4$, $O(6)_2$–$O(6)_3$ 3.00 $O(6)_1$–$O(2)_3$, $O(6)_2$–$O(2)_4$ 2.54 $O(2)_1$–$O(6)_3$, $O(2)_2$–$O(6)_4$ 2.54 O(5)–O(3′) 2.75
O(6) tg (166°) + gt (65°)	116.1	31	64	−2.97	22.1	$O(3)_1$–$O(6)_1$, $O(3)_2$–$O(6)_2$ 2.65 $O(2)_3$–$O(6)_3$, $O(2)_4$–$O(6)_4$ 2.77 $O(2)_1$–$O(2)_3$, $O(2)_2$–$O(2)_4$ 2.72 $O(3)_1$–$O(6)_4$, $O(3)_2$–$O(6)_3$ 2.85 $O(6)_1$–$O(3)_4$, $O(6)_2$–$O(3)_3$ 2.93 O(5)–O(3′) 2.74 O(2)–O(6′) 2.71

[a] 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) to C(6), and pure gt = 60°, tg = 180°, gg = −60°. [b] Corner chain 0° position is with $O(4)_1$ at 0, $-y$, z; for center chain with $O(4)_3$ at $a/2$, $b/2 - y$, z. Positive rotation is clockwise looking down the c axis. Translation is center chain relative to the corner chain along c. [c] PE = packing energy is the nonbonded term of eq 1. [d] Atom subscripts indicate residue numbers: 1 and 2 for corner chain, 3 and 4 for center chain. Unsubscripted hydrogen bonds are intramolecular.

Table III
Short Nonbonded Contacts in the Three Most Probable Antiparallel Models Predicted by Packing Analysis

Model	Contact	Distance, Å	Limits,[16] Å	
			Extreme	Normal
O(6) tg	O(2)–HC(6) (center–center)	2.16	2.20	2.40
O(6) gt	O(6)–O(5) (intrachain)	2.62	2.70	2.80
	C(6)H–O(3) (corner–center)	2.32	2.20	2.40
O(6) tg + gt	C(6)–O(3) (corner–center)	2.73	2.70	2.80
	C(6)H–O(3) (corner–center)	2.30	2.20	2.40

X-Ray Analysis. The method of x-ray refinement was very similar to the procedure followed for the packing analysis. (Typical strategy and method details have previously been published.[7]) The same structural parameters were allowed to vary as in the packing refinement and the criterion of success was the minimization of the crystallographic reliability index, R. (The latter is defined by the expression $R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c represent the observed and calculated structure factors, respectively.) An isotropic temperature factor of 3.0 was used, unless otherwise indicated. In the final refinement runs of each model, the function

$$\Phi = fR + (1 - f)Y \tag{2}$$

where Y is the energy function of eq 1, was minimized in order to remove any bad contacts that developed in the refinement based on R alone. Typical values of f were in the range 0.8–0.9, which allowed approximately equal weight to both terms in eq 2. In addition, in the final runs the components of an anisotropic temperature factor

$$B = \frac{h^2a^{*2}B_x}{4} + \frac{k^2b^{*2}B_y}{4} + \frac{l^2c^{*2}B_z}{4} \tag{3}$$

were refined.

The results of this refinement are shown in Table IV. (The gg models were not refined to completion because of very high R values for the initial models.) Strongly indicated was that antiparallel chain packing polarity was again preferred over parallel, confirming the results of the packing analysis. Of greater significance, however, was that the best predicted packing model, with the mixture of tg and gt O(6) rotations, was the correct model for the structure of cellulose II. This was indicated both by the final results and the initial R factors calculated before refinement. The packing energy corresponding to this structure, 26.0 units, represented only a slight increase above the minimum packing energy of 22.1 obtained for the best packing model. This was a further indication of the near identity of the best packing model and the true structure, as was also indicated by the nearly coincident chain rotations and translation, O(6) rotations, and bridge angles (compare Tables II and IV). It is interesting to note that the O(6) gt model experienced only a 1.8 unit rise in packing energy during the course of the x-ray refinement. However, the refined R factor for this model, 0.22, was not sufficiently low to render it a possible alternative to the mixed O(6) structure.

The antiparallel mixed O(6) model was further subjected to a refinement sequence in which the O(6) rotational positions were allowed to vary independently on each of the four glucose residues within the unit cell. No significant changes in O(6) positions from those obtained previously were observed and the tg + gt structure with equal rotations within the chains was maintained.

A comparison of the observed and calculated structure amplitudes for this structure is shown in Table V. The atomic coordinates are given in Table VI and the unit cell structure is shown in projection in Figure 2.

Odd-Order Meridionals. The existence of weak odd-order meridional reflections in the x-ray diagram of cellulose II has long been a subject of debate. After careful examination of numerous Fortisan diffractograms, the presence of these reflections was confirmed. One possible explanation for their presence may be the disorder brought about by different O(6) rotational positions in the same cellulose chain. To test this hypothesis, a cellulose II model was constructed in which successive residues contained tg and gt O(6) rotational positions in an alternating fashion. The actual model constructed in this manner refined to a packing energy of 26.4 units and gave an R factor of 0.24 (cf. Table IV). The odd-order meridionals predicted by this structure were $F_{001} = 1.26$ and $F_{003} = 0.21$ (compare Table V). Based on the reasonably low

Table IV
X-Ray Intensity Refinement of Cellulose II Models

Model[a]	Bridge angle, deg	Chain rotations,[b] deg		Center chain translation,[b] Å	Initial R (B = 3.0)	Refined R	PE[c]	Temp factor (B)	f (eq 2)
		Corner	Center						
				Parallel					
O(6) tg (185°)	117.6	23	31	−2.24	0.245	0.232	35.5	3.0	0.8
O(6) gt (45°)	117.8	31	29	−2.23	0.282	0.276	41.6	3.0	0.9
O(6) gg (−60°)	116.6	48	50	1.98	0.490		42.8		
O(6) tg (188°) + gt (80°)	115.2	38	31	−2.58	0.286	0.276	37.8	3.0	0.9
				Antiparallel					
O(6) tg (178°)	116.9	25	57	−3.55	0.347	0.200	26.5	3.0	0.9
O(6) gt (31°)	116.4	37	69	−3.23	0.225	0.222	24.4	3.0	0.8
O(6) gg (−45°)	116.5	48	48	−3.98	0.565		44.2		
O(6) tg (166°) + gt (64°)	116.8	29	64	−3.02	0.195	0.159	26.0	3.0	0.9
	116.8	29	64	−3.02	0.195	0.151	26.0	B_x = 1.0 B_y = 2.2 B_z = 2.8	0.9
O(6) tg (166°) and gt (50°) alternating in same chain	117.4	29	61	−2.93	0.243	0.238	26.4	3.0	0.9

[a] 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) to C(6), and pure gt = 60°, tg = 180°, gg = −60°. [b] Corner chain 0° position is with $O(4)_1$ at 0, $-y$, z; for center chain with $O(4)_3$ at $a/2$, $b/2 - y$, z. Positive rotation is clockwise looking down the c axis. Translation is center chain relative to the corner chain along c. [c] PE = packing energy is the nonbonded term of eq 1.

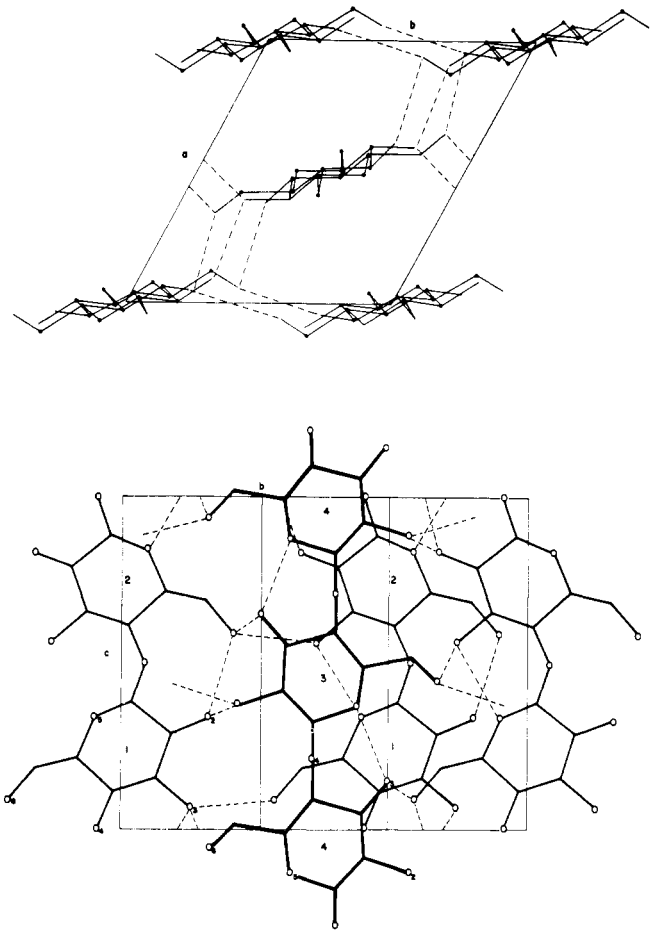


Figure 2. Projection of the cellulose II unit cell on the (top) x–y plane and the (bottom) y–z plane. (In the y–z projection, the left rear corner chain is not shown for clarity.) The hydrogen atoms are not shown and the hydrogen bonds are indicated by dashed lines. The residues are numbered in accordance with the notation used in Tables II, IV, and VI.

packing energy and the similarities between this and the best structure, this type of disorder does not appear to be improbable. However, because the observed intensities of the odd-order meridionals are much weaker than the predicted values, the fraction of alternating O(6) tg–gt sequences is likely to be quite small.

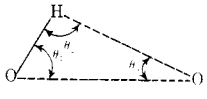
Hydrogen Bonding Network. Emphasized previously was the concept that the additional stability of cellulose II over the native polymorph was due to an extensive hydrogen bond network. The final structure of cellulose II is characterized by a total of five intermolecular hydrogen bond pairs and two intrachain hydrogen bonds. The intermolecular hydrogen bonds may be further classified into intrasheet and intersheet bonds, the latter contributing more significantly to the stability of cellulose II over cellulose I. The O(3)–O(6) (corner–corner) and O(2)–O(6) (center–center) hydrogen bonds represent the intrasheet variety while the O(2)–O(2), O(6)–O(3), and O(3)–O(6) hydrogen bonds represent the intersheet interactions.

Locating the hydroxyl hydrogen atomic positions as well as defining the hydrogen bond direction in terms of donors and acceptors was accomplished in a very simple manner. The hydroxyl hydrogen atoms, not included in the packing and x-ray analysis, were placed on the cellulose structure with appropriate bond angles and distances (O–H bond length of 1.0 Å and C–O–H bond angle of 110°). The conformation angles about the C–O bond were then refined using packing refinement until best hydrogen bonding was obtained. The optimum O–O distance was considered to be 2.8 Å and the optimum H–O distance of a hydrogen bonded hydroxyl hy-

Table VII
Hydrogen Bond Lengths and Angles

Hydrogen bond[a]	Length, Å		Angles, deg[b]		
	O–O	O–H	θ_1	θ_2	θ_3
Intramolecular					
$O(3)_1H–O(5)_2$	2.70	2.06	41.5	119.8	18.7
$O(3)_3H–O(5)_4$	2.70	2.06	41.5	119.8	18.7
$O(2)_1H–O(6)_2$	2.76	2.12	41.7	120.0	18.3
Intermolecular					
$O(2)_3H–O(2)_1$ $O(2)_4H–O(2)_2$	2.62	1.66	12.8	159.5	7.7
$O(6)_1H–O(3)_4$ $O(6)_2H–O(3)_3$	2.84	2.34	50.8	109.9	19.3
$O(3)_1H–O(6)_4$ $O(3)_2H–O(6)_3$	2.80	1.94	25.0	142.5	12.5
$O(6)_1H–O(3)_1$ $O(6)_2H–O(3)_2$	2.65	2.00	40.6	120.4	19.0
$O(6)_3H–O(2)_3$ $O(6)_4H–O(2)_4$	2.97	1.99	9.4	165.9	4.7

[a] Atom subscripts indicate residue numbers: 1 and 2 for corner chains and 3 and 4 for center chains, as shown in Figure 2. [b] Angle definition:



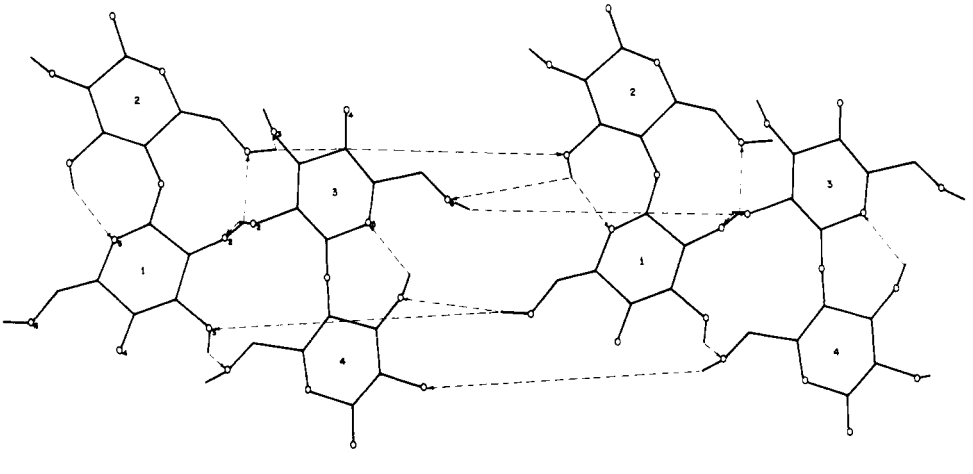


Figure 3. Hydrogen bond directions in cellulose II. Hydroxyl hydrogens are included. The chain separation is increased in the horizontal direction for better clarity.

Table VIII
Comparison of Conformational Features for both Chains of the Cellulose II Unit Cell Before and After Packing[a]

Chain	Bridge angle, deg	Rotation of residue about virtual bond[b]	Conformational angles (deg) ϕ, ψ[c]	O(6) conformational angle, deg	Hydrogen bond lengths, Å		Conformational energy (arbitrary units)
					O(5)–O(3′)	O(2)–O(6′)	
Corner							
Before	118.3	44.3	148, 202	165.1	2.76	2.81	12.5
After	116.8	41.4	148, 199	165.4	2.70	2.76	12.7
Center							
Before	118.0	41.8	148, 200	59.3	2.76		12.7
After	116.8	41.4	148, 199	64.0	2.70		12.9

[a] Before packing: minimum energy conformation for isolated chain. After packing: the final refined structure. [b] Rotation is relative to an arbitrary origin in the y–z plane. [c] Using the convention of ref 2.

Table IX
Comparison of Bond Angles and Conformation Angles Obtained in Single Chain Conformation Refinement and in the Final Refined Structure

Bond angle	Angle, deg (difference from standard value[13])		Conformation angle	Angle, deg (difference from standard value[13])	
	Best conformation	Final structure		Best conformation	Final structure
C(1)–C(2)–C(3)	109.2 (-1.3)[a]	108.3 (-2.2)[a]	O(5)–C(1)–C(2)–C(3)	57.9 (1.9)	59.6 (3.6)[a]
C(2)–C(3)–C(4)	109.6 (-0.9)	109.9 (-0.6)	C(1)–C(2)–C(3)–C(4)	-52.1 (1.1)	-55.5 (-2.3)
C(3)–C(4)–C(5)	113.2 (2.9)[a]	111.1 (0.8)	C(2)–C(3)–C(4)–C(5)	52.9 (-0.1)	54.4 (1.4)
C(4)–C(5)–O(5)	106.5 (-3.5)[a]	108.6 (-1.4)[a]	C(3)–C(4)–C(5)–O(5)	-55.5 (-0.1)	-55.2 (0.2)
C(5)–O(5)–C(1)	113.7 (1.7)[a]	113.2 (1.2)[a]	C(4)–C(5)–O(5)–C(1)	62.1 (1.0)	61.0 (-0.1)
O(5)–C(1)–C(2)	109.7 (0.5)	109.5 (0.3)	C(5)–O(5)–C(1)–C(2)	-65.5 (-3.3)[a]	-64.3 (-2.1)
O(5)–C(1)–O(1)	109.2 (1.9)[a]	107.1 (0.2)			
C(2)–C(1)–O(1)	107.9 (-0.5)	108.3 (-0.1)			
C(1)–C(2)–O(2)	110.3 (1.0)	110.2 (0.9)			
C(3)–C(2)–O(2)	109.4 (-1.4)	110.4 (-0.4)			
C(2)–C(3)–O(3)	108.4 (-1.2)	110.1 (0.5)			
C(4)–C(3)–O(3)	110.5 (0.8)	110.9 (1.2)			
C(3)–C(4)–O(4)	107.6 (-2.8)[a]	107.3 (-3.1)[a]			
C(5)–C(4)–O(4)	105.6 (-3.0)[a]	106.7 (-1.9)[a]			
C(4)–C(5)–C(6)	114.6 (1.9)[a]	114.5 (1.8)[a]			
O(5)–C(5)–C(6)	106.0 (-0.9)	106.1 (-0.8)			
C(5)–C(6)–O(6)	111.5 (-0.3)	113.5 (1.7)[a]			

[a] One or more standard deviations different from standard value.[13]

drogen atom to be 1.8 Å. The best hydrogen bond scheme determined in this fashion is shown in Figure 3. Characteristic of this scheme is that the O(6) and O(3) atoms of the corner chains are involved in bifurcated hydrogen bonds as they donate a single hydrogen atom to two other oxygen atoms. Table VII outlines the H–O and O–O contact distances for this network as well as the corresponding bond angles.

Conformational Changes Due to Packing. Crystal structure analysis of polymers has long been done by packing of chains that have been fixed into a minimum energy conformation by single chain conformational analysis. In this work, it was suspected that packing forces would alter such a minimum energy conformation but it was not known by how much. A comparison of selected conformational features for both chains of the cellulose II unit cell is shown in Table VIII, as determined before packing (i.e., by isolated chain conformational refinement) and after (i.e., in the final structure). The differences are relatively small, but they reflect a systematic effect on all atoms (i.e., residue rotation), or those in the hydroxymethyl groups. Even small differences here may cause significant differences in the R index. However, as the conformational angles ϕ and ψ show, packing has not caused the chains to move away from the conformational energy minimum determined by single chain conformational analysis.

More significant differences were found in the bond and conformation angles, relative to their standard values which were used to establish the original model. Both in minimizing the isolated chain conformation and in final structure refinement several of the angles were found to exceed one or more standard deviations of the Arnott and Scott standard values. Some differences between the best conformation values and the final structure are also evident. Bond lengths remained essentially unchanged from the average values. The bond angles and conformation angles of the best conformation and the refined cellulose II structure are shown in Table IX, along with their differences from the Arnott–Scott average values.

These differences in structural parameters shown in Tables VIII and IX may become significant when attempts are made to refine structures against x-ray data with fixed, average conformations. Under such conditions, minimization of R may not be possible and one may not be able to distinguish between different models.

Conclusions

The extensive hydrogen bond system present in cellulose II is clearly the reason for the stability of its structure in comparison with other polymorphs. For example, in cellulose I only intrasheet hydrogen bonds are present[2,3] and the same has recently been found in cellulose III prepared from cellulose I.[17] In that respect, the predictions made previously about relative stabilities of cellulose polymorphs have been confirmed.[5]

A very important result of the present study, in our estimation, was the success of the packing analysis in predicting the correct crystal structure of cellulose II. All stereochemical features of the structure, complete with chain rotations and

translation, bridge angle, O(6) rotations, and hydrogen bonds were nearly perfectly predicted by the packing analysis. This was reflected in a very low R index for the predicted structure even prior to refinement against x-ray data. The even lower R index obtained after x-ray refinement points out the degree of reliability that can be obtained with this refinement method.

It is also interesting to compare the results obtained in this study with the previously made prediction of the cellulose II structure which made use of a rigid chain packing method.[2] In the latter, a probable conformation of the chain was established by ϕ, ψ rotations and then packed into the unit cell without further refinement of the conformation. Only helix rotation and translation, as well as limited O(6) rotations, were allowed as variables. The results were only partially conclusive, in that antiparallel polarity was correctly predicted but a choice of optimum hydrogen bonding could not be made. In addition, the chain rotations were off by nearly 10° and chain translation was almost 0.5 Å in error. This illustrates very clearly the utility of stereochemical packing analysis with fully variable models. In this connection, it is also noteworthy that the relatively small deviations shown in Tables VIII and IX can be significant in terms of final structure analysis.

Acknowledgment. This work has been supported by the National Science Foundation Grant No. MPS7501560.

Supplementary Material Available: Tables I, V, and VI (6 pages). Ordering information is given on any current masthead page.

References and Notes

(1) R. H. Marchessault and A. Sarko, *Adv. Carbohydr. Chem.*, **22**, 421–482 (1967).
(2) A. Sarko and R. Muggli, *Macromolecules*, **7**, 486–494 (1974).
(3) K. H. Gardner and J. Blackwell, *Biopolymers*, **13**, 1975–2001 (1974).
(4) W. A. Sisson, *Contrib. Boyce Thompson Inst.*, **12**, 31–44 (1941).
(5) A. Sarko, *J. Appl. Polym. Sci., Appl. Polym. Symp.*, **28**, 729–742 (1976).
(6) H. D. Chanzy, "Structure of Fibrous Biopolymers", Colston Papers No. 26, E. D. T. Atkins and A. Keller, Ed., Butterworths, London, 1975, pp 417–434.
(7) P. Zugenmaier and A. Sarko, *Biopolymers*, in press.
(8) T. Bluhm and A. Sarko, to be published.
(9) W. T. Winter and A. Sarko, *Biopolymers*, **13**, 1447–1460 (1974).
(10) W. T. Winter and A. Sarko, *Biopolymers*, **13**, 1461–1482 (1974).
(11) F. J. Kolpak and J. Blackwell, *Macromolecules*, **8**, 563–564 (1975).
(12) R. E. Franklin and R. G. Gosling, *Acta Crystallogr.*, **6**, 678–685 (1953).
(13) S. Arnott and W. E. Scott, *J. Chem. Soc., Perkin Trans. 2*, 324–335, (1972).
(14) P. Zugenmaier and A. Sarko, *Acta Crystallogr., Sect. B*, **28**, 3158–3166 (1972).
(15) D. W. Jones, *J. Polym. Sci.*, **42**, 173–188 (1960).
(16) P. R. Sundararajan, Ph.D. Thesis, University of Madras, Madras, India, 1969; A. G. Walton and J. Blackwell. "Biopolymers", Academic Press, New York, N.Y., 1973, p 34.
(17) A. Sarko, J. Southwick, and J. Hayashi, the following paper in this issue.

Packing Analysis of Carbohydrates and Polysaccharides. 7. Crystal Structure of Cellulose III_I and Its Relationship to Other Cellulose Polymorphs

Anatole Sarko,* Jeffrey Southwick, and Jisuke Hayashi

Department of Chemistry, SUNNY College of Environmental Science and Forestry, Syracuse, New York 13210, and Department of Applied Chemistry, Faculty of Engineering, Hokkaido University, Sappora, Japan. *Received May 19, 1976*

ABSTRACT: The crystal structure of cellulose polymorph III_I, obtained by treatment of native ramie cellulose with liquid ammonia, was solved through a combined stereochemical structure refinement and x-ray diffraction analysis. The structure is based on a parallel packing of chains very similar to that of cellulose I. The characteristics of the structure include the same intramolecular O(5)–O(3′) and O(2)–O(6′) hydrogen bonds and C(3)–O(6) intermolecular ones that bond the chains into sheets in cellulose I. The liquid ammonia apparently shifts adjacent sheets relative to one another into a metastable, higher energy structure which reverts back to cellulose I upon heating in water. A very similar structure for cellulose III_{II}, obtained by liquid ammonia treatment of cellulose II, was also predicted by the stereochemical analysis. This structure is antiparallel as is cellulose II and it resembles the latter in chain conformation and hydrogen bonds, although again it is metastable in comparison with cellulose II. The weak odd-order meridional reflections observed in the diffractograms of cellulose III are correctly predicted by relaxing the strict $P2_1$ chain symmetry, particularly with respect to the rotation of hydroxymethyl groups. This is consistent with the presence of a small amount of substituent group rotational disorder in the structure of cellulose III. As was found to be the case with celluloses I and II, the stereochemical and x-ray refinement techniques complemented each other well and both were necessary in order to solve the structure.

Native crystalline cellulose, commonly known as cellulose I, gives rise to at least three polymorphic structures upon appropriate treatment.[1] For example, its regeneration or mercerization result in cellulose II, and both the latter as well as cellulose I can be converted to cellulose III through the use of liquid ammonia. Cellulose III, in turn, can be converted to cellulose IV by heat treatment. All four polymorphs crystallize well and their structures apparently differ only in the crystalline packing of chains with nearly the same conformation, because all four polymorphs show the same fiber repeat of ~10.3 Å.[1]

* SUNY College of Environmental Science and Forestry.

The highly crystalline cellulose I of the alga *Valonia ventricosa* has been previously shown to crystallize with a parallel packing of chains.[2,3] Whether the same is true for the less crystalline native celluloses of ramie, cotton, etc., is presently not known, but their x-ray diffraction patterns are nearly identical with that of *Valonia*, although less well resolved. Conversion of ramie or cotton celluloses into cellulose II results in a structure that is based on antiparallel packing of chains.[4,5] When, for example, ramie cellulose I or cellulose II are treated with liquid ammonia, two different cellulose III diffraction diagrams are obtained: the so-called III_I from cellulose I and the III_{II} from cellulose II.[6] Both are nearly, but