Packing Analysis of Carbohydrates and Polysaccharides. III. *Valonia* Cellulose and Cellulose II[1a]

Anatole Sarko* and Reto Muggli[1b]

Department of Chemistry, State University of New York College of Environmental Science and Forestry, Syracuse, New York 13210. *Received October 24, 1973*

ABSTRACT: An attempt to solve the crystal structure of naturally occurring crystalline cellulose I of the alga *Valonia ventricosa* through combined stereochemical packing analysis and X-ray intensity analysis is described. From an excellent correlation of results obtained separately by the two analysis methods it was found highly probable that the chains of *Valonia* cellulose packed with parallel polarity maintained a twofold screw axis for the rigid components of the glucose ring but showed some flexibility of the hydroxymethyl group rotations. The most probable positions of the latter were near the tg position which allowed extensive hydrogen bonding in at least two directions within the unit cell. Based on the results of both packing and X-ray analysis, a new triclinic, two-chain unit cell was proposed for *Valonia*. The applicability of this unit cell to other native celluloses is discussed. The stereochemical packing analysis was also extended to predict the crystal structure of cellulose II. A highly hydrogen-bonded, antiparallel structure with hydroxymethyl groups again near the tg position was shown to be the most probable structure for this polymorph. The interconversion of celluloses I and II is discussed in the light of stabilities of both structures as predicted by packing analysis.

Cellulose is known to crystallize in at least four different polymorphic forms: celluloses I, II, III, and IV.[2a] In addition, native crystalline celluloses found in some bacteria and marine algae resemble cellulose I but show evidence of larger unit cells.[2a] Of these polymorphs, cellulose I of ramie and cotton[2b-7] and the algal celluloses from *Valonia ventricosa*[7-9] and *Chaetomorpha melagonium*[10] have been studied more intensively than others, primarily because of their higher degree of crystallinity. Most such studies have been based on X-ray or electron diffraction, although significant information has been obtained from infrared and Raman spectroscopies.[7,11] In spite of the studies which have spanned a period of more than 30 years, the crystal structures of these celluloses are still unknown. The unit cells of the various polymorphs were determined quite some time ago,[2a] but doubts remain whether these cells are really correct or are merely the smallest elementary cells which satisfy most but not all X-ray reflections. This is particularly true for the less crystalline polymorphs where broadening of strong reflections may hide the weak reflections. Similarly, the observed fiber repeat is in the range 10.3–10.4 Å for all polymorphs. This information, coupled with stereochemical and density data, indicates that there are two glucose residues per fiber repeat. Nonetheless, claims have been made that this is not a true repeat and that the true conformation of cellulose is helical.[12] Certainly, the assumption of a crystallographic twofold screw axis which relates contiguous glucose residues along the chain is questionable[6] because of the appearance of weak odd order meridional reflections.

Along with questions on the size and shape of the unit cells and the chain conformation, the crystalline packing and the size and supermolecular structure of the crystallites in the various polymorphs remain unknown. Because the cellulose chain is stereochemically nonsymmetric, both parallel and antiparallel chain packing schemes are possible. Structures based on both polarities have been advanced.[7-13] Electron micrographic evidence indicates the existence of elementary fibrils 30–40 Å in diameter and less.[14-17] These elementary fibrils are presumed to be single crystals, although serious disagreement exists between the crystallite size measurements by X-ray diffraction and the diameters of the elementary fibrils obtained from electron micrographs. This once again points out our inadequate knowledge of the total cellulose structure.

In view of our recent success with the packing analysis of crystalline sugars[18] and polysaccharides[19] and the excellent diffraction diagrams obtained from *Valonia* cellulose, we began a thorough comparative investigation of the structure of cellulose polymorphs. In this communication we report on the crystalline structure of cellulose I of *Valonia* and some preliminary findings on cellulose II.

Experimental Section

Both electron and X-ray diffraction diagrams were obtained of fibers separated[16] from the cell walls of commercially available *Valonia ventricosa* cells. X-Ray diagrams were recorded either in an evacuated flat-film camera or helium-filled cylindrical camera with Cu $K\alpha$ radiation. Kodak No-Screen film was used. Electron diffraction diagrams were obtained with a JEM-7A microscope. The electron diffractograms were exceptionally rich in diffraction data and showed 16 layer lines with 160 clearly resolved reflections. The X-ray patterns showed a smaller number of layer lines allowing the measurement of 52 reflection intensities, not counting the unobservable intensities. The intensities were measured by integration with a Joyce-Loebl recording microdensitometer and converted to structure amplitudes corrected for Lorentz and polarization factors, reflection arcing, distance of diffracted ray to film, the Franklin artificial temperature factor, and film scanning direction other than the radius of the reflection.[20,21] The individual intensities of partially superimposed reflections were obtained by least-squares resolution of microdensitometer tracings.[22] No evidence for crystallite orientation other than parallel to the fiber axis was found.

Results and Discussion

Valonia **Cellulose. Diffraction Measurements.** Careful evaluation of the unit cell parameters from electron diffraction patterns showed that all reflections could be indexed with monoclinic unit cell with dimensions a = 15.76 Å, b = 16.42 Å, c (fiber axis) = 10.34 Å, γ = 96.8°, containing eight chains. These dimensions are in good agreement with those previously reported.[8,9] Good agreement also exists between our results and those reported for *Chaetomorpha melagonium*[10] cellulose. In both X-ray and electron diffraction patterns, reflections requiring the above, large unit cell were present only on the first and higher order layer lines; the equator could be indexed with a smaller unit cell (a = 7.85 Å, b = 8.14 Å, c = 10.34 Å, γ = 96.6°) almost identical with the two-chain Meyer and Misch cell.[2] On the other hand, the equatorial reflections could not be indexed on the basis of the smallest possible, one-chain cell, as could be done for *Chaetomorpha* cellulose,[10] because of the presence of a few weak reflections requiring the larger, two-chain cell.

These findings indicated, first, that at least some of the eight chains of the true unit cell were nonequivalent. The nonequivalence could be of several types, *viz.*, different

chain polarity, different rotational and translational positions of the chains relative to the c axis, or difference in the rotational positions of the hydroxymethyl groups (*i.e.*, the O(6) atoms). However, on the basis of reflection indexing it was considered likely that the true unit cell consisted of four Meyer-Misch-type subcells in which the two chains could be nonequivalent in several ways, yet the differences between the four subcells were only translational. This argument was based on the fact that such translational differences would not cause the appearance of new reflections on the equator. These conclusions were substantiated in subsequent packing and X-ray refinement analyses.

Chain Packing Analysis. The objectives of the packing analysis were twofold. The first was to provide suitable packing models for subsequent refinement of the structure against X-ray intensity data. The second was to compare the results of the packing analysis with independent X-ray refinement, in order to measure the effectiveness of the former procedure when used alone. The relative wealth of X-ray data for *Valonia* made such a comparison possible.

The chain backbone conformation used in the packing analysis remained fixed once its most probable conformation had been determined. The latter was obtained by the calculation of the conformational energy maps (ϕ-ψ maps) using the potential energy functions previously used for amylose.[23] The intramolecular hydrogen-bond energies, as previously defined,[23] were included in these calculations. The atomic coordinates of the glucose residue used in the calculations were obtained by averaging the coordinates determined by X-ray diffraction for the crystal structures of β-D-glucose[24] and cellobiose.[25] The glycosidic bond angle was 116°. The bond lengths, angles, and conformational angles of this residue were close to the corresponding average values determined by Arnott and Scott.[26]

The most probable chain conformation thus determined was characterized by a twofold screw relating the two residues along the fiber repeat. The justification for such a chain model can be seen in the ϕ-ψ maps calculated for the three most probable rotational positions of the hydroxymethyl groups, *viz.*, tg, gg, and gt[27] (*cf.* Figure 1).

All three energy maps were similar to the extent that they contained two regions of allowed conformations, approximately 150° apart along the ϕ axis. Within the main allowed region, a narrow potential energy well appeared near one of the $n = 2$, $h = 5.15$ Å positions (marked Cl), at $(\phi,\psi) \simeq (140°, 210°)$. The intersection of these two contours corresponded to the twofold screw-related cellulose molecule with a fiber repeat of 10.3 Å and with the cellobiose residue in the Hermans conformation. Two of the three energy maps showed a minimum very near to the Cl position. The conformations corresponding to this minimum were all characterized by a strong O(5)--O(3′) hydrogen bond. An additional O(2)--O(6′) hydrogen between contiguous residues was possible for these conformations when the O(6) was in the tg position. The conformation of crystalline cellobiose (marked CB in the maps) was also close to the Cl position.

The foregoing analysis was approached from the point of view that all monomer residues of the cellulose chain were equivalent, or at least that their ring components were. The possibility still existed, however, that a cellulose crystal structure was based on the cellobiose unit as its monomer, with a consequence that there would be two nonequivalent β-(1→4) linkages in the chain. As the conformational energy maps showed, such structures were conceivable, although of higher energy, and could be arrived at by either moving along the $n = 2$ contour simultaneously in both directions from the $h = 5.15$ Å point, separately for each glucose residue, or by varying (ϕ,ψ) and the glycosidic bond angle separately for each residue. The consequence of such structures that argued against their existence was that the O(5)--O(3′) hydrogen bonds between successive residues would be of unequal strength. The experimental infrared spectroscopic evidence is, however, not consistent with this.

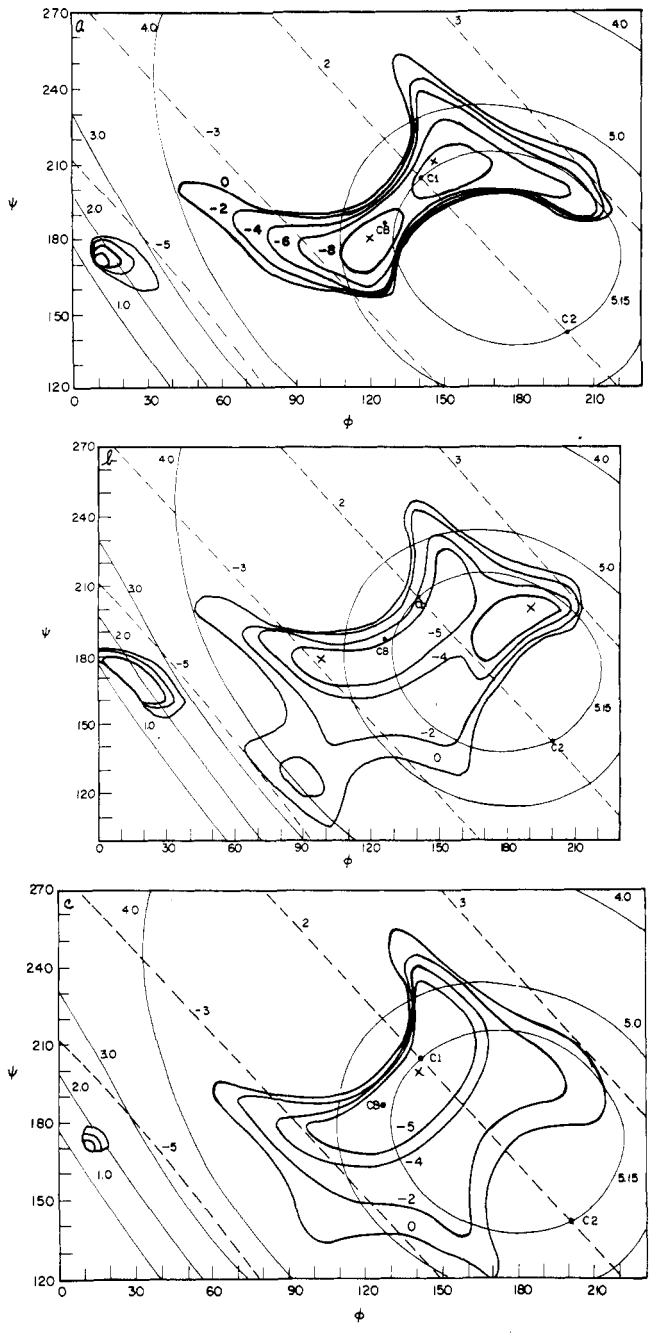

Figure 1. Conformational energy maps of the cellulose chain as a function of rotations about the C(1)—O(1) bond (ϕ) and the O(1)—C(4′) bond (ψ). The $(\phi,\psi) = (0,0)$ position is as previously defined.[23] The hydroxymethyl groups are (a) tg, (b) gg, and (c) gt. The energy contours are in kcal/mol of residue; the dashed lines denote constant n (number of residues/helix turn, with positive n indicating a right-handed helix); the thin solid lines denote constant h (rise per residue in Å along the helix axis). The conformations marked by C1 and C2 denote the two helix conformations with $n_h = 2_{5.15}$ which satisfy the observed fiber repeat. CB denotes the experimentally observed conformation of crystalline cellobiose and crosses mark energy minima.

The most probable conformation for an *isolated* cellulose chain should thus be very close to a position dictated

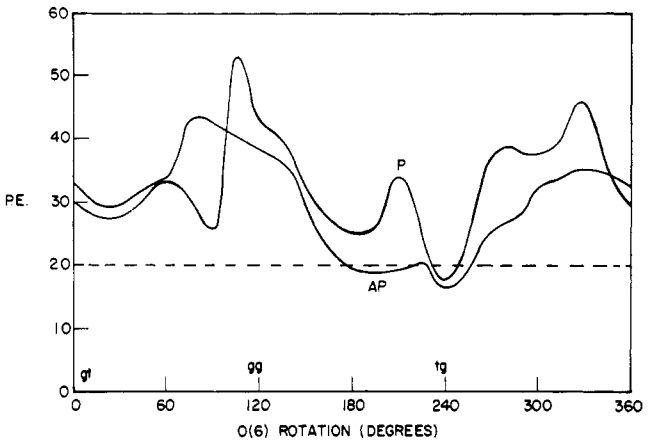

Figure 2. Packing energy (PE) of best models as a function of hydroxymethyl group rotations in parallel (P) and antiparallel (AP) packing models of *Valonia* cellulose. All four O(6) of the two-chain Meyer-Misch-type subcell were rotated in a synchronous fashion. The dashed line represents an arbitrary boundary separating probable and improbable models.

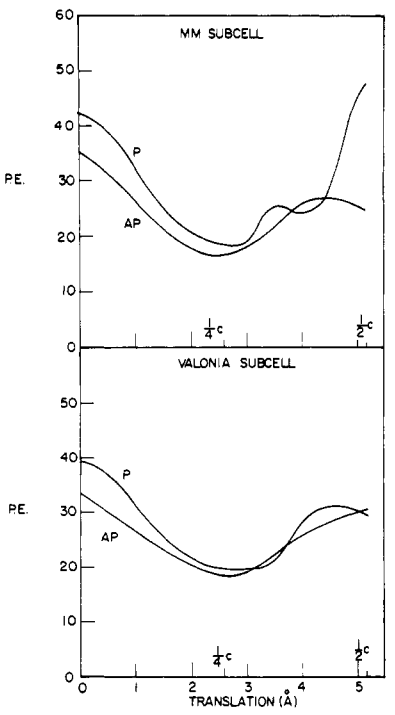

Figure 3. Packing energy (PE) of best models as a function of center chain translation in parallel (P) and antiparallel (AP) packing models of *Valonia* cellulose. Both Meyer-Misch (MM) and the new *Valonia* two-chain subcells are shown. All O(6) are tg.

by a molecular twofold screw axis and characterized by the same O(5)--O(3′) hydrogen bond as found in cellobiose. There is little likelihood for any other, comparable energy conformation, and it is reasonable to assume that the packing of chains into a crystal lattice will provide the small amount of energy needed for the small adjustment of the conformation to the twofold screw axis. As shown below, the screw axis need only apply to the rigid ring structure of the sugar residues, and not to their hydroxymethyl groups.

Even with the chain backbone conformation remaining invariant, the chain packing analysis required a relatively large number of variables. First, a decision had to be made how many of the eight chains of the unit cell were to be considered unique. As previously indicated, it appeared from the X-ray data that only two chains of one subcell were unique and all others were related to these two chains in a simple fashion. This restricted the number of variables within each packing polarity to seven, *viz.*, independent rotations of both chains about the *c* axis, translation of the second chain relative to the first chain along the same axis, and the four independent rotations of the hydroxymethyl groups, two on each chain. In our analysis procedure, the first three variables could be handled simultaneously, while the hydroxymethyl rotations could only be handled by fixing the rotational positions at predetermined values while the former three variables were being refined. Even though a concurrent refining of all variables would have been preferable, it turned out that our scheme did not overlook any energy minima nor lead to false minima. The refining was performed with a modified PCK5 least-squares program of Williams[28] as previously described.[19] This program determined the least-squares minimum repulsive energy position of the chains in the unit cell, in accordance with the supplied potential energy functions. Possibility of intermolecular hydrogen-bond formation between all oxygens was included in the refinement.

Rather than consider a large number of combinations of the hydroxymethyl rotational positions, an attempt was initially made to restrict the range of rotations to the most probable. This was done for both packing polarities by rotating all four hydroxymethyls to predetermined and equal positions and letting the PCK5 program find the minimum energy packing position. The complete 0–360° range of rotations was explored in 10° steps. The variable parameters were the chain rotations and the second chain translation. The results are shown in Figure 2 in the form of a plot of minimum packing energy (PE) as a function of the hydroxymethyl group rotation. Using the criterion of PE = 20 or lower as indicative of a probable model,[18] it was clear that for both parallel and antiparallel chain packing polarities the vicinity of the tg hydroxymethyl position was highly favored. On the other hand, the gg position was definitely excluded while the gt position was again not favored. These results were in good agreement with the conformational analysis of the isolated chain (*cf.* Figure 1) which indicated that the tg conformation was one of those favored by virtue of the formation of the intramolecular O(2)--O(6′) hydrogen bond.

Similarly, an attempt was made to find the most probable range of translational positions of the second (center) chain of the two-chain subcell relative to the first (corner) chain. This was accomplished for both packing polarities by finding the minimum packing energy position of the two chains for fixed, predetermined translational positions (and fixed hydroxymethyl rotations) while letting the chain rotations vary. The results, shown in Figure 3 as plots of minimum PE as a function of center chain translation, clearly showed preference for a relative translation of approximately $\frac{1}{4}c$. Both parallel and antiparallel packing models were identical in this respect.

These results led to the following test. If, it was argued, the minimum energy positions of two neighboring chains on the subcell diagonal (*i.e.*, corner and center chains) was $\frac{1}{4}c$, would this relative translational difference be the same for all pairs of chains in that diagonal direction? In other words, would there be a preference for an alternation of up-down-up-down, etc., by $\frac{1}{4}c$ for successive chains on the diagonal, or would there be a preference of a $\frac{1}{4}c$ step in only one direction? The latter situation would result in every fourth chain ($4 \times \frac{1}{4}c = c$) being translationally identical, thus explaining the large *Valonia* unit cell. The results of this test, also shown in Figure 3 as the *Valonia* subcell curves, were virtually identical with the results of the former (*i.e.*, Meyer–Misch subcell) case.

Table I
Minimum Energy Packing Models for *Valonia* Cellulose

PE (Arbitrary Units)	Chain Rot. (deg)[a] Corner	Chain Rot. (deg)[a] Center	Center Chain Translation (Å)	O(6) Rot. Positions[b]	Hydrogen Bonds and Lengths (Å)[c]	
				Model 1: MM subcell, parallel packing		
19.7	37	32	−2.43	tg, tg, tg, tg	$O(6)_1$—$O(3)_1$	2.86
					$O(6)_2$—$O(3)_2$	2.88
					($O(6)_2$—$O(5)_1$	3.19)
19.1	33	37	2.36	tg, tg, tg, tg	$O(6)_1$—$O(3)_1$	2.86
					$O(6)_2$—$O(3)_2$	2.86
19.1	38	29	−2.79	tg+15°, tg+15°, tg+15°, tg+15°	$O(6)_1$—$O(3)_1$	2.77
					$O(6)_2$—$O(3)_2$	2.75
18.6	30	38	3.00	tg+15°, tg+15°, tg+15°, tg+15°	$O(6)_1$—$O(3)_1$	2.73
					$O(6)_2$—$O(3)_2$	2.80
19.1	38	28	−2.50	tg, tg, tg−15°, tg−15°	$O(6)_1$—$O(3)_1$	2.87
					$O(6)_2$—$O(5)_1$	2.79
					($O(6)_2$—$O(3)_2$	3.22)
19.0	38	27	−2.58	tg, tg, tg+15°, tg+15°	$O(6)_1$—$O(3)_1$	2.87
					$O(6)_2$—$O(3)_2$	2.78
					($O(6)_2$—$O(5)_1$	3.15)
18.8	32	37	2.49	tg, tg, tg+15°, tg+15°	$O(6)_1$—$O(3)_1$	2.87
					$O(6)_2$—$O(3)_2$	2.76
					($O(6)_1$—$O(5)_2$	3.21)
19.6	37	31	−2.58	tg+15°, tg, tg+15°, tg	$O(6)_1$—$O(3)_1$	2.75, 2.85
					$O(6)_2$—$O(3)_2$	2.72, 2.82
					($O(6)_2$—$O(5)_1$	3.17)
18.8	31	38	2.87	tg+15°, tg, tg+15°, tg	$O(6)_1$—$O(3)_1$	2.72, 2.89
					$O(6)_2$—$O(3)_2$	2.78, 2.87
					($O(6)_2$—$O(5)_1$	3.18)
				Model 2: MM subcell, antiparallel packing		
16.5	36	38	2.37	tg, tg, tg, tg	$O(6)_1$—$O(3)_1$	2.85
					$O(6)_2$—$O(3)_2$	2.87
19.4	35	36	2.68	tg+15°, tg+15°, tg+15°, tg+15°	$O(6)_1$—$O(3)_1$	2.72
					$O(6)_2$—$O(3)_2$	2.72
17.9	35	37	2.74	tg+15°, tg+15°, tg+15°, tg+15°	$O(6)_1$—$O(3)_1$	2.73
					$O(6)_2$—$O(3)_2$	2.86
17.8	36	38	2.64	tg, tg+15°, tg, tg+15°	$O(6)_1$—$O(3)_1$	2.85, 2.74
					$O(6)_2$—$O(3)_2$	2.87, 2.78
				Model 3: *Valonia* subcell, parallel packing		
19.5	35	35	2.68	tg, tg, tg, tg	$O(6)_1$—$O(3)_1$	2.85
					$O(6)_2$—$O(3)_2$	2.85
20.2	33	36	2.91	tg+15°, tg+15°, tg, tg	$O(6)_1$—$O(3)_1$	2.71
					$O(6)_2$—$O(3)_2$	2.85
20.0	36	36	2.79	tg−15°, tg, tg−15°, tg	$O(6)_1$—$O(3)_1$	2.85
					$O(6)_2$—$O(3)_2$	2.85
					($O(6)_2$—$O(5)_1$	3.16)
20.4	35	36	2.87	tg+15°, tg, tg+15°, tg	$O(6)_1$—$O(3)_1$	2.72, 2.85
					$O(6)_2$—$O(3)_2$	2.73, 2.85
				Model 4: *Valonia* subcell, antiparallel packing		
18.3	38	39	2.77	tg, tg, tg, tg	$O(6)_1$—$O(3)_1$	2.86
					$O(6)_2$—$O(3)_2$	2.89
19.3	36	37	2.60	tg+15°, tg+15°, tg, tg	$O(6)_1$—$O(3)_1$	2.74
					$O(6)_2$—$O(3)_2$	2.86
19.3	37	38	2.56	tg+15°, tg, tg+15°, tg	$O(6)_1$—$O(3)_1$	2.75, 2.85
					$O(6)_2$—$O(3)_2$	2.79, 2.87

[a] The 0° position is by definition one where the glycosidic oxygens lie in the *yz* plane. Positive rotation is clockwise looking down the *z* axis toward the *xy* plane. [b] The O(6) sequence for parallel packing models is: lower residue of corner chain, upper residue of the corner chain, lower residue of center chain, upper residue of center chain. For antiparallel packing models, the last two are reversed. [c] Subscripts 1 and 2 denote, respectively, the corner and center chains.

With the most probable ranges for the O(6) rotation and chain translation thus determined, the strategy for finding the minimum packing energy models was considerably simplified. Because the least-squares energy surface as a function of the model parameters is usually not smooth but may exhibit a number of minima and maxima, the search for the lowest minimum should involve starting from as many different combinations of the parameter values as is possible. The limited range of O(6) rotations and of chain translation made the number of starting positions manageable and allowed the testing of a large number of models. For example, the four hydroxymethyl groups in one subcell could be considered independently rotatable (near the most probable position) and not necessarily restricted by any symmetry relations such as the twofold screw axis. A large number of starting chain rotations could also be used for every combination of O(6) rotations.

On the average, 185 optimization runs/model were made in this fashion. Of this number, 27 on the average resulted in a minimum energy structure with PE $\cong$ 20 or less and 24 of these structures were within the same minimum. As previously shown in a packing analysis of known monosaccharide crystal structures,[18] the real structure

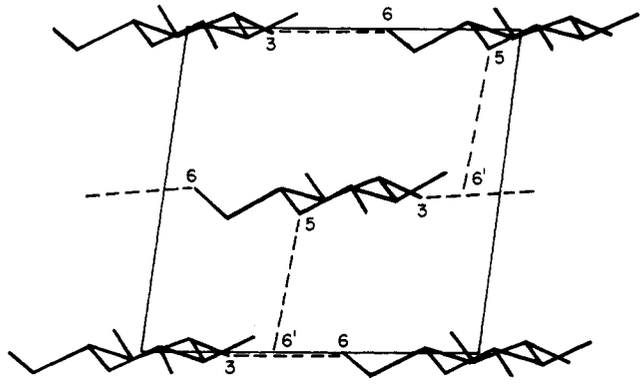

Figure 4. An approximation of the most probable *Valonia* subcell structure in *ab* projection showing the possible hydrogen bonds (dashed lines). Only one residue of each chain is shown for clarity; hydrogen atoms are omitted. Numbers indicate hydrogen bonded oxygens (primed oxygens are in the symmetry-related residue which is not shown).

was approached only by those predicted minimum energy structures that contained the largest number of possible hydrogen bonds. Not all predicted cellulose structures exhibited the maximum number of hydrogen bonds. Those that did are shown in Table I.

The predicted minimum energy structures showed several interesting features. First, in all four packing models of Table I the rotational positions of both chains of the subcell were severely restricted to an average value of 35°, with a range of 27–39°. In models 2, 3, and 4 the rotational positions of the two chains were nearly identical, in model 1 (MM subcell, parallel) there was a consistent difference of up to 11°. Because the cellulose chain is roughly ribbon like, the above rotational position resulted in a sheet-like structure with the plane of the sheet nearly coincident with the *bc* plane of the unit cell. This can clearly be seen in Figure 4. Secondly, the relative translational difference between the corner and center chains was similarly restricted, with an average value of almost exactly $\pm\frac{1}{4}c$. Maximum deviation from this figure was ~0.4 Å. No systematic differences between the four models were apparent in these respects.

The O(6) rotational positions were also restricted to a narrow range of tg ±15°, primarily because of the possibility of forming O(6)--O(3) hydrogen bonds between neighboring chains in the plane of the sheet (*i.e.*, between corner–corner and center–center chains). An interesting feature was that identical rotational positions of the four hydroxymethyls were not required and almost any mix of positions within the 30° range refined to a minimum energy structure.

The general features of the predicted minimum packing energy structure of *Valonia* cellulose were thus the following. The ribbon-like chains were arranged in sheets stabilized by hydrogen bonds in two directions, *viz.*, the O(5)--O(3′) and O(2)--O(6′) intramolecular bonds along the chain and the intermolecular O(6)--O(3) bonds roughly perpendicular to the chain axes. The latter bonds necessarily restricted the translational positions of the chains *within* the sheet to nearly identical. The two sheets that formed the subcell also were nearly identical in their internal structure, but were translated relative to one another by $\sim\frac{1}{4}c$ in either direction. From a comparison of models 1 and 2 on the one hand with models 3 and 4 on the other, it appeared that either a regular alternation of + and − shifts of $\frac{1}{4}c$ between neighboring sheets or a one-directional shift of either all + or all − were equally likely. The latter possibility suggested a new unit cell for the *Valonia* cellulose—one which was triclinic with approxi-

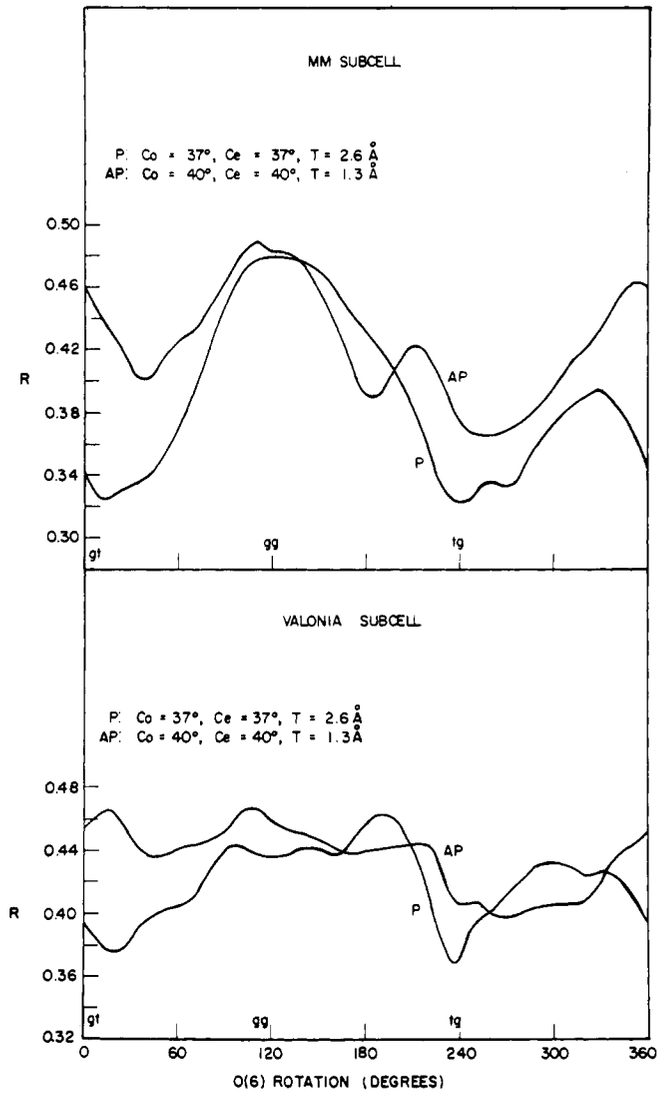

Figure 5. Crystallographic *R* index as a function of O(6) rotation for parallel (P) and antiparallel (AP) packing models in both Meyer-Misch (MM) and the new *Valonia* two-chain subcells. Rotations of corner (Co) and center (Ce) chains and translation of center chain (T) are indicated.

mate parameters of a = 9.41 Å, b = 8.15 Å, c = 10.34 Å, α = 90°, β = 57.5°, and γ = 96.2°. The main difference between this cell and the usual Meyer-Misch type cell was that the two nearest like sheets (*e.g.*, corner-corner sheets) were translated by $\frac{1}{2}c$ relative to one another. (The probability of this cell is further discussed under X-ray results.)

Thus far, the polarity of the sheets had not been considered. Based on a comparison of packing energy values, either between models 1 and 2 or models 3 and 4, the probability for either packing polarity was the same. In the light of previous findings with known monosaccharide structures,[18] the differences in packing energies between parallel and antiparallel models were considered insignificant. However, based on the possibility of forming the maximum number of hydrogen bonds—an important point also found in previous work[18]—the parallel packing mode was slightly favored. As shown in Table I, both parallel models (1 and 3) may in some cases form a weak intersheet hydrogen bond between atoms O(6)--O(5), 3.1–3.2 Å long. Although at this length the hydrogen bond will be very weak, a slight adjustment of bond and torsional angles in the chain molecular structure will undoubtedly decrease its length by at least 0.1–0.2 Å. No

Table II
Best Crystallographic Packing Models for *Valonia* Cellulose

R Index	Chain Rot. (deg) Corner	Chain Rot. (deg) Center	Center Chain Translation (Å)	O(6) Rot. Positions
		Model 1: MM subcell, parallel packing		
0.30	37	40	2.6	tg−15°, tg−15°, tg−15°, tg−15°
0.29	37	40	2.6	tg, tg, tg, tg
0.30	34	37	2.6	tg+15°, tg+15°, tg+15°, tg+15°
0.30	34	40	2.6	tg, tg, tg−15°, tg−15°
0.29	37	40	2.6	tg, tg−15°, tg, tg−15°
0.30	37–40	34–40	2.6	tg, tg, tg+15°, tg+15°
0.30	37	34	2.6	tg, tg+15°, tg, tg+15°
		Model 2: MM subcell, antiparallel packing		
0.37	40	43	1.3	tg−15°, tg−15°, tg−15°, tg−15°
0.36	40	40	1.3	tg, tg, tg, tg
0.36	40	40	1.3	tg+15°, tg+15°, tg+15°, tg+15°
0.38	40	43	1.3	tg, tg, tg−15°, tg−15°
0.38	37–40	40–43	1.3	tg, tg−15°, tg, tg−15°
0.36	40	40	1.3	tg, tg, tg+15°, tg+15°
0.36	40	40	1.3	tg, tg+15°, tg, tg+15°
		Model 3: *Valonia* subcell, parallel packing		
0.32	37	37	2.6	tg−15°, tg−15°, tg−15°, tg−15°
0.29	34	37–40	2.6	tg, tg, tg, tg
0.31	34	37	2.6	tg+15°, tg+15°, tg+15°, tg+15°
0.31	34–37	37–40	2.6	tg, tg, tg−15°, tg−15°
0.30	37	40	2.6	tg, tg−15°, tg, tg−15°
0.30	34	37	2.6	tg, tg, tg+15°, tg+15°
0.30	34	37	2.6	tg, tg+15°, tg, tg+15°
		Model 4: *Valonia* subcell, antiparallel packing		
0.37	37–40	40–43	1.3	tg−15°, tg−15°, tg−15°, tg−15°
0.36	37–40	40–43	1.3	tg, tg, tg, tg
0.36	37	40	1.3	tg+15°, tg+15°, tg+15°, tg+15°
0.37	34–37	40	1.3–1.6	tg, tg, tg−15°, tg−15°
0.37	40	43	1.3–1.6	tg, tg−15°, tg, tg−15°
0.36	37–40	40–43	1.3–1.6	tg, tg, tg+15°, tg+15°
0.36	37–40	40–43	1.3–1.6	tg, tg+15°, tg, tg+15°

such possibility existed for the antiparallel packing mode. The presence of an intersheet hydrogen bond will further stabilize the crystal structure. The direction of this bond is shown in Figure 4. Whether the possibility of such a bond could cause a truly bifurcated hydrogen bond (not an unknown phenomenon) or a statistical mix of two independent hydrogen bonds, cannot at present be decided.

X-Ray Intensity Analysis. For reasons stated earlier, chain packing and X-ray intensity analysis were performed in parallel and using essentially the same strategy and variables. It was more convenient not to use least-squares methods in the X-ray refinement (because of their relatively small radius of convergence); rather, all parameters were varied in systematic fashion. This ensured the widest possible coverage of all parameters. Both Meyer-Misch and *Valonia* subcells were investigated.

The criterion of success was the minimization of the crystallographic reliability index

$$R = (\Sigma||F_0| - |F_c||)/\Sigma|F_0|$$

where F_0 and F_c are the observed and calculated structure factors, respectively. The structure factors were calculated in the usual manner, using an isotropic temperature factor of 32.5. The latter is so large because an artificial temperature factor[20] had been applied to the observed structure amplitudes. Without it the temperature factor used in the calculation would be 6.5.

The separate effects of O(6) rotation and the center chain translation on the R index are shown in Figures 5 and 6. For the parallel packing models the results were identical with those obtained in packing analysis with a minimum in R near the tg position for O(6) and a $\frac{1}{4}c$ translation of the chains. This was true for both Meyer-Misch and *Valonia* subcells. A minimum was also observed for the same packing models near the gt O(6) position (~15°), but the packing analysis ruled out this conformation (*cf.* Figure 2).

For the antiparallel case, a single minimum near the tg O(6) position was again found, but the R index was approximately 0.05 unit greater than for the parallel case. Similarly, the results of chain translation (*cf.* Figure 6) were completely in favor of the parallel model, not only because of a better R index but also because of a lack of agreement between the minima of the R index and the packing energy for the antiparallel model (*cf.* Figure 3).

The results of the complete analysis of chain rotations, translations and O(6) rotations bore out the preliminary results and are shown in Table II. As in the case of the packing analysis, a ±15° range about the tg O(6) positions was found and a mixture of rotational positions was as probable as a single fixed rotation. The most probable rotational positions of the chains were nearly identical with those found in packing analysis, especially for parallel models. For the antiparallel models, the serious disagreement in the best translational positions of the center chains noted in preliminary runs persisted, thus eliminating the antiparallel models as a likely packing polarity for *Valonia* cellulose.

Although a range of chain rotations for the parallel models is in some cases given in Table II, the minimum in the R index always occurred at unequal rotational positions of the two chains of the subcell. This was consistent with the observation that in projection down the fiber axis a single chain subcell did not describe the *Valonia* structure.

The comparisons of observed and calculated structure amplitudes for the best parallel and antiparallel Meyer-Misch and *Valonia* models are shown in Table III. It is

Table V
Minimum Energy Packing Models for Cellulose II

PE	Chain Rot. (deg) Corner	Chain Rot. (deg) Center	Center Chain Translation (Å)	O(6) Rot. Positions	Hydrogen Bonds and Lengths (Å)	
			Model 1: Parallel packing			
18.5	26	28	2.62	tg, tg, tg, tg	$O(6)_1$—$O(3)_1$	2.80
					$O(6)_2$—$O(3)_2$	2.76
					$O(6)_2$—$O(2)_1$	2.78
19.2	23	30	2.74	tg+15°, tg+15°, tg, tg	$O(6)_1$—$O(3)_1$	2.74
					$O(6)_2$—$O(3)_2$	2.69
					$O(6)_2$—$O(2)_1$	2.74
19.6	27	32	2.72	tg, tg, tg−15°, tg−15°	$O(6)_1$—$O(3)_1$	2.80
					$O(6)_2$—$O(3)_2$	2.86
					$O(6)_2$—$O(2)_1$	2.78
			Model 2: Antiparallel packing			
18.5	26	43	−4.67	tg−45°, tg−45°, tg+30°, tg+30°	$O(6)_2$—$O(2)_2$	2.88
					$O(6)_2$—$O(3)_2$	2.77
					$O(6)_1$—$O(3)_2$	2.78
16.9	25	55	−3.43	tg, tg, gt−30°, gt−30°	$O(6)_2$—$O(5)_1$	2.85
					$O(2)_1$—$O(2)_2$	2.80
					$O(6)_1$—$O(3)_1$	2.84
					$O(6)_2$—$O(3)_1$	2.82
					$O(6)_2$—$O(2)_2$	2.83
15.1	22	54	−4.12	tg+15°, tg+15°, tg−45°, tg−45°	$O(2)_1$—$O(2)_2$	2.75
					$O(6)_1$—$O(3)_1$	2.79
					$O(6)_1$—$O(3)_2$	2.81
18.5	26	43	−4.54	tg, tg, tg+30°, tg+30°	$O(6)_1$—$O(3)_1$	2.81
					$O(6)_2$—$O(3)_2$	2.77
					$O(6)_2$—$O(2)_2$	2.88

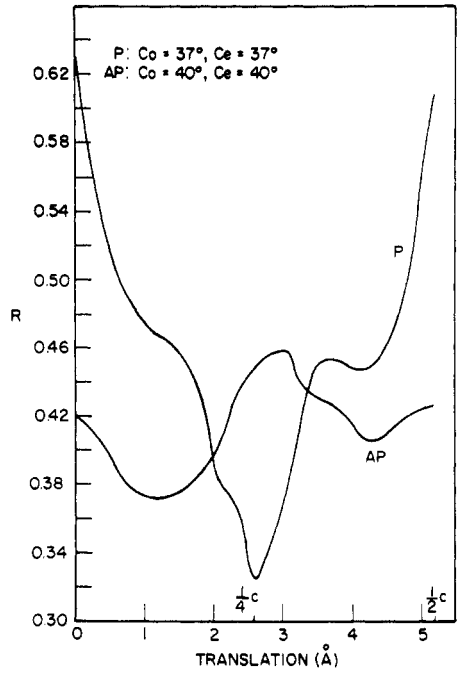

Figure 6. Crystallographic *R* index as a function of center chain translation for parallel (P) and antiparallel (AP) packing models of *Valonia* cellulose. Rotations of corner (Co) and center (Ce) chains are indicated. All O(6) are tg.

readily apparent that the parallel models of both unit cells show no major discrepancies between the observed and calculated values. On the other hand, several gross inconsistencies—such as medium to strong predicted intensities where no intensities were observed—appear in the antiparallel models. On this basis it was clear that the antiparallel packing of *Valonia* must be ruled out.

Concerning the applicability of the triclinic cell for *Valonia*, it was thought highly significant that the results obtained for the parallel packing with both subcells turned out to be identical (*cf.* Table II). This clearly pointed to the suitability of the new triclinic cell. A final test to judge its validity was performed by comparing the fit of observed reflection spacings with those calculated for both the Meyer-Misch subcell and the triclinic cell. The fit with both cells is shown in Table IV and it is again readily apparent that in both cases the fit is equally good. More importantly, the triclinic cell accounts for all reflections which previously required the large, eight-chain *Valonia* unit cell.

Cellulose II. On the strength of the excellent correlation of packing and X-ray analysis results obtained with *Valonia*, a similar packing analysis was performed for cellulose II. The unit cell of cellulose II is well established[2a] (a = 9.08 Å, b = 7.92 Å, c (fiber axis) = 10.34 Å, γ = 117.3°) but again the crystal structure is unknown. Typical X-ray diffraction patterns of cellulose II are not sufficiently well resolved to permit a structure determination on the basis of diffraction data alone.

The strategy used with cellulose II was essentially the same as was used for *Valonia*. The same chain model was again employed. The preliminary results for O(6) rotations and center chain translation again paralled those for *Valonia*, indicating clear preferences for particular positions. A selection of the final, most probable packing models obtained for both parallel and antiparallel packing polarities are shown in Table V.

As was true of cellulose I, the packing analysis of cellulose II resulted in only a small number of most probable models. For example, the chain rotations were in reasonably narrow ranges, as were the chain translations. In general, these ranges were not the same as in cellulose I. Similarly, although at least two of the O(6) were close to the tg position, considerably more rotational latitude was evident, including even the vicinity of the gt position, forbidden for cellulose I. The most notable differences between cellulose I and cellulose II were in the hydrogen bonds—both parallel and antiparallel packing polarities of cellulose II were able to form more and stronger hydrogen bonds. In most models the same intrasheet bonds were present as in cellulose I, but in addition a variety of intersheet hydrogen bonds could form. This added stability of

cellulose II over cellulose I may explain why it is possible to convert cellulose I into cellulose II but not in reverse.

Although it was not possible with any degree of certainty to pick any model of cellulose II as the most probable, one model stood out with regard to maximum hydrogen bond formation—the second antiparallel model of Table V, with PE = 16.9. This model, shown in Figure 7 in projection, was the most extensively hydrogen-bonded structure and should, therefore, be considered the most probable for cellulose II. It was characterized by two sheetlike hydrogen-bond formations, one between parallel chains as in cellulose I and the other between antiparallel chains in the long diagonal direction of the unit cell. The staggering of antiparallel chains was more pronounced—approximately $\frac{1}{3}c$—than in cellulose I.

It remains to be seen whether the X-ray data will substantiate this model when sufficiently resolved diffractograms become available.

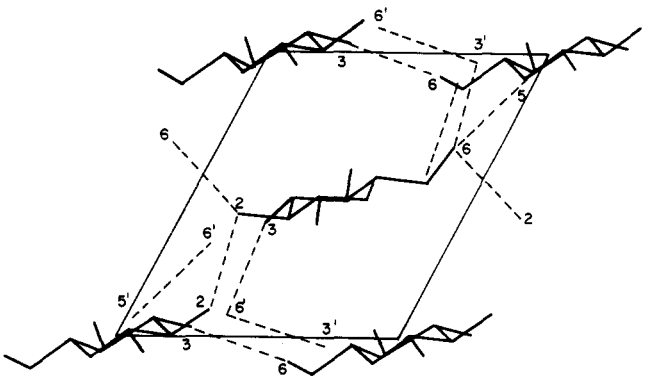

Figure 7. An approximation of the most probable cellulose II packing structure shown in *ab* projection. The possible hydrogen bonds between numbered oxygens are indicated with dashed lines. Only one residue of each chain is shown for clarity and hydrogen atoms are omitted. (Primed oxygens are in the symmetry-related residue which is not shown.)

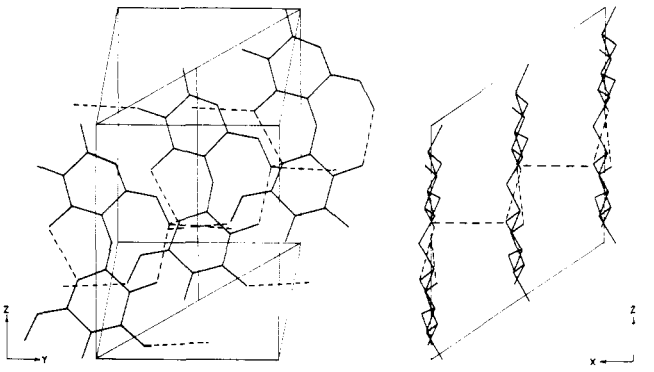

Figure 8. The *yz* and *xz* projections of the new proposed *Valonia* unit cell. Only the chains on the long *ab* diagonal of the cell are shown for clarity and hydrogen atoms are omitted. Possible hydrogen bonds are shown by dashed lines.

Conclusions

The combined results of both packing and X-ray analysis clearly pointed out that the crystal structure of *Valonia* is based on a parallel packing of chains. It was not as clear what the arrangement of subcells was in the large unit cell although a unidirectional stepping of chains in the *c* direction was strongly indicated. It did not appear probable that successive sheets of chains should alternate in their translation to the extent that the four sheets which made up the large, observed unit cell of *Valonia* showed a regular series of translations such as $+\frac{1}{4}c$, $-\frac{1}{4}c$, $-\frac{1}{4}c$, $+\frac{1}{4}c$. There was no advantage in this scheme over one with regular $+\frac{1}{4}c$ translation from any point of view. It was also unlikely that a mix of subcells was involved, *i.e.*, a measure of translational disorder, because reflection line width measurements have indicated crystallite diameters of the order of 140 Å,[14] and there was no evidence of translational disorder in X-ray diagrams obtained with well monochromatized and focussed radiation. Furthermore, the suitability of the two-chain triclinic cell also supported the regular translation of adjacent sheets.

Similarly, both packing and X-ray analysis indicated that a mix of O(6) rotations in the vicinity of the tg position may be present. This was in agreement with the presence of odd order meridional reflections in the X-ray diagram of *Valonia*. A statistical mix seemed again unlikely because of no indications of disorder in the X-ray diagram. On the other hand, a definite mix of O(6) positions was likely, in order to accommodate unequal chain rotations and different hydrogen bonds. But it is not possible, at this time, to determine precisely the nature of this mix.

We therefore conclude that the most probable *Valonia* structure is characterized by the triclinic unit cell of the dimensions given above, with a unidirectional translation of sheets of $\sim\frac{1}{4}c$. The O(6) atoms of this cell, whether they are rotated by different degrees or not, could participate in a maximum number of hydrogen bonds. Some of the latter could be bifurcated thus providing a reason for the slightly different rotations of the two chains of the unit cell. Since all of the X-ray reflections requiring the large, eight-chain *Valonia* cell are accounted for by this cell, the necessity for the large cell seems to have disappeared.

Two projections of the proposed structure are shown in Figure 8 and the atomic coordinates are listed in Table VI.

Although the predicted structure for cellulose II is not as yet substantiated by experiment, the high degree of correlation obtained between the packing analysis and X-ray analysis demonstrated the efficacy of the predictive method and hence, lent considerable weight to the predicted structure. Because cellulose II is formed from solution where during crystallization a chain folding mechanism is expected to operate, the antiparallel packing feature of the predicted structure appeared even more probable. Its stability is evident in the number of hydrogen bonds it could form, whether again through a mix of O(6) rotations, or hydrogen-bond bifurcations, or through a measure of disorder. This remains to be determined. Again, the presence of weak odd-order medidional reflections in the X-ray diagram of cellulose II suggested different O(6) positions. Similarly, there was some evidence of disorder in the apparent smearing out of some reflections.

It was interesting to compare the published X-ray diagrams of cellulose I from other sources such as ramie and cotton with that of *Valonia*. In general, all cellulose I X-ray diagrams are alike except that *Valonia* shows a much larger crystallite size and a higher degree of crystallinity, therefore, much improved resolution in the diagram. Because of this, the few generally weak reflections which require a different unit cell for *Valonia* did not appear in the diagrams for ramie, cotton, etc., celluloses. Their diagrams could be indexed on the basis of the Meyer-Misch unit cell. Similarly, the *Valonia* diagram clearly showed the absence of the few reflections which were predicted for antiparallel packing, but this distinction could not be made for other native celluloses. The question now arose whether all native celluloses were identical, including that of *Valonia*? On the basis of what we predicted for the structure of cellulose II with its antiparallel packing giving it increased stability over cellulose

Table VI
Cartesian Atomic Coordinates of the Best Parallel Model, Triclinic *Valonia* Cell (Chain Rotations and O(6) Positions as in Table IIIB)

Atom	x (Å)	y (Å)	z (Å)
	1. Corner Chain		
C(1)	−0.10	0.38	3.95
C(2)	0.19	1.47	2.95
C(3)	−0.32	1.08	1.58
C(4)	0.21	−0.25	1.17
C(5)	−0.05	−1.31	2.27
C(6)	0.60	−2.64	2.00
O(2)	−0.40	2.67	3.37
O(3)	0.05	2.08	0.64
O(4)	−0.47	−0.70	0
O(5)	0.50	−0.83	3.50
O(6)	0.15	−3.21	0.78
H(1)	−1.13	0.25	4.06
H(2)	1.23	1.60	2.90
H(3)	−1.37	1.03	1.61
H(4)	1.24	−0.19	0.98
H(5)	−1.08	−1.44	2.38
C(1′)	0.10	−0.38	9.12
C(2′)	−0.19	−1.47	8.12
C(3′)	0.32	−1.08	6.75
C(4′)	−0.21	0.25	6.34
C(5′)	0.05	1.31	7.44
C(6′)	−0.60	2.64	7.17
O(2′)	0.40	−2.67	8.54
O(3′)	−0.05	−2.08	5.81
O(4′)	0.47	0.70	5.17
O(5′)	−0.50	0.83	8.67
O(6′)	−0.15	3.21	5.95
H(1′)	1.13	−0.25	9.23
H(2′)	−1.23	−1.60	8.07
H(3′)	1.37	−1.03	6.78
H(4′)	−1.24	0.19	6.15
H(5′)	1.08	1.44	7.55
	2. Center Chain		
C(1)	3.85	4.00	6.48
C(2)	4.21	5.07	5.48
C(3)	3.67	4.72	4.11
C(4)	4.13	3.34	3.70
C(5)	3.81	2.31	4.80
C(6)	4.40	0.94	4.53
O(2)	3.67	6.30	5.90
O(3)	4.10	5.69	3.17
O(4)	3.42	2.93	2.53
O(5)	4.39	2.76	6.03
O(6)	3.92	0.39	3.31
H(1)	2.81	3.93	6.59
H(2)	5.25	5.16	5.43
H(3)	2.62	4.72	4.14
H(4)	5.17	3.36	3.51
H(5)	2.77	2.23	4.91
C(1′)	4.02	3.23	11.65
C(2′)	3.66	2.15	10.65
C(3′)	4.20	2.51	9.28
C(4′)	3.73	3.88	8.87
C(5′)	4.06	4.92	9.97
C(6′)	3.47	6.28	9.70
O(2′)	4.20	0.92	11.07
O(3′)	3.77	1.54	8.34
O(4′)	4.45	4.30	7.70
O(5′)	3.47	4.47	11.20
O(6′)	3.95	6.83	8.48
H(1′)	5.06	3.30	11.76
H(2′)	2.62	2.07	10.60
H(3′)	5.25	2.50	9.31
H(4′)	2.70	3.87	8.68
H(5′)	5.10	5.00	10.08

I, as well as the argument that the biosynthetic mechanisms of cellulose in different plants should be identical, we conclude that all native cellulose I's should be identical. The differences in crystallinity that arise may be caused by varying degrees of disorder, in turn caused by small differences in the biosynthetic mechanisms. The presence of other polysaccharides may be a factor in this. It stands to reason that a structure of lower stability (*i.e.*, higher energy) results because of kinetic control—such as during the biosynthesis—and not because of thermodynamic control presumably in operation during solution crystallization.

Even if cellulose II could be converted to a cellulose I structure during a controlled solid-state transformation, the negligible packing energy differences between parallel and antiparallel cellulose I's suggest that an antiparallel structure would result. Considering that the crystallite sizes would probably not change during such a transformation, the X-ray diagram of this hypothetical cellulose I would be indistinguishable from that of native ramie or cotton cellulose I. The same argument is, of course, valid for a reverse, I → II transformation. What this points out is that in a polysaccharide such as cellulose whose parallel and antiparallel packing polarities are nearly identical, the diffraction data may not yield information even on such a gross feature of the crystal structure as chain polarity. The remedy, of course, lies in the improvement of the crystallinity of the sample.

Acknowledgment. This work was supported by National Science Foundation, Grant No. GP-27997.

Supplementary Material Available. Tables III and IV, listing structure factor amplitudes, will appear following these pages in the microfilm edition of this volume of the journal. Photocopies of the supplementary material from this paper only or microfiche (105 × 148 mm, 24X reduction, negatives) containing all of the supplementary material for the papers in this issue may be obtained from the Journals Department, American Chemical Society, 1155 16th St., N.W., Washington, D. C. 20036. Remit check or money order for $3.00 for photocopy or $2.00 for microfiche, referring to code number MACRO-74-486.

References and Notes

(1) (a) Presented in part at the 4th Canadian Wood Chemistry Symposium, Quebec, Canada, July 1973. Part II of this series is by P. Zugenmaier and A. Sarko. (b) F. Hoffmann-LaRoche & Co., Ltd., 4002 Basel, Switzerland.
(2) (a) R. H. Marchessault and A. Sarko, *Advan. Carbohyd. Chem.*, 22, 421 (1967). (b) K. H. Meyer and L. Misch, *Helv. Chim. Acta*, 20, 232 (1937).
(3) N. M. Bikales and L. Segal, "Cellulose and Cellulose Derivatives," Wiley-Interscience, New York, N. Y., Vol. V, Part IV, 1971.
(4) D. W. Jones, *J. Polym. Sci.*, 32, 371 (1958).
(5) D. W. Jones, *J. Polym. Sci., Part A*, 42, 173 (1968).
(6) K. C. Ellis and J. O. Warwicker, *J. Polym. Sci.*, 56, 339 (1962).
(7) C. Y. Liang and R. H. Marchessault, *J. Polym. Sci.*, 37, 385 (1959).
(8) G. Honjo and M. Watanabe, *Nature (London)*, 181, 326 (1958).
(9) D. G. Fisher and J. Mann, *J. Polym. Sci.*, 42, 189 (1960).
(10) I. Nieduszynski and E. D. T. Atkins, *Biochim. Biophys. Acta*, 222, 109 (1970).
(11) J. Blackwell, P. D. Vasko, and J. L. Koenig, *J. Appl. Phys.*, 41, 4375 (1970).
(12) A. Viswanathan and S. G. Shenouda, *J. Appl. Polym. Sci.*, 15, 519 (1971).
(13) R. Muggli, *Cellulose Chem. Technol.*, 2, 549 (1968).
(14) D. F. Caulfield, *Textile Res. J.*, 41, 267 (1971).
(15) I. Nieduszynski and R. D. Preston, *Nature (London)*, 225, 5229 (1970).
(16) K. H. Gardner and J. Blackwell, *J. Ultrastruct. Res.*, 36, 725 (1971).
(17) R. B. Hanna, Ph.D. Thesis, SUNY College of Environmental Science and Forestry, Syracuse, N. Y., 1973.
(18) P. Zugenmaier and A. Sarko, *Acta Crystallogr., Sect. B*, 28, 3158 (1972).
(19) P. Zugenmaier and A. Sarko, *Biopolymers*, 12, 435 (1973).
(20) R. E. Franklin and R. G. Gosling, *Acta Crystallogr.*, 6, 678 (1953).
(21) A. Sarko, Fortran computer program FIBRXRAY.
(22) R. D. B. Fraser and E. Suzuki, *Anal. Chem.*, 38, 1770 (1966).
(23) J. Blackwell, A. Sarko, and R. H. Marchessault, *J. Mol. Biol.*, 42, 379 (1969).
(24) S. C. Chu and G. A. Jeffrey, *Acta Crystallogr., Sect. B*, 24, 830 (1968).
(25) R. A. Jacobson, J. A. Wunderlich and W. N. Lipscomb, *Acta Crystallogr.*, 14, 598 (1961).
(26) S. Arnott and W. E. Scott, *J. Chem. Soc., Perkin Trans. 2*, 324 (1972).
(27) M. Sundaralingam, *Biopolymers*, 6, 189 (1968).
(28) D. E. Williams, *Acta Crystallogr., Sect. A*, 25, 464 (1969).