Crystal Structure of *Valonia* Cellulose Iβ

V. L. Finkenstadt[†] and R. P. Millane*

Whistler Center for Carbohydrate Research, Purdue University, West Lafayette, Indiana 47907-1160

Received March 26, 1998; Revised Manuscript Received July 20, 1998

ABSTRACT: The crystal structure of *Valonia* cellulose Iβ is determined by X-ray fiber diffraction analysis. A careful reanalysis of two existing X-ray diffraction data sets for *Valonia* cellulose I leads to a consistent, definitive structure of the β-phase. The results resolve ambiguities in existing X-ray analyses between "parallel up" and "parallel down" packing of the cellulose sheets in native *Valonia* cellulose. The molecular and sheet structures are essentially identical to those previously determined, but these results define precisely the packing of the sheets. The sheet packing is discussed in terms of the energy and density of the crystal structure. The structure of the β-phase has specific implications for the packing of the α-phase of *Valonia* cellulose I.

Introduction

The crystal structure of native cellulose, or cellulose I, has been the subject of investigation for 70 years, since the pioneering work of Meyer and Mark[1] and Meyer and Misch.[2] Although many of the features of the structure of native crystalline cellulose are now well understood, a number of ambiguities still remain. An important step in the study of the crystal structures of native cellulose was results obtained from a number of X-ray fiber diffraction analyses.[3−5] Although these studies produced similar results, there were some differences.

First, there were differences between celluloses from higher plants such as ramie[5] and those from algae such as *Valonia*[3,4] and *Chaetomorpha*.[6] Diffraction patterns from ramie cellulose showed a monoclinic unit cell through which two molecules pass, with 2-fold screw symmetry. On the other hand, celluloses from *Valonia* and *Chaetomorpha* give fiber diffraction patterns that were indexed on the basis of an eight-chain monoclinic unit cell, and there was some evidence that the chain may deviate from exact 2-fold screw symmetry. However, the lateral dimensions of the eight-chain unit cell are very close to twice those of the two-chain unit cell, the reflections that demand the larger unit cell are few in number and are weak relative to the other reflections, and the 001 and 003 reflections that are inconsistent with 2-fold screw symmetry of the molecules are also weak. In view of this, although there are clearly differences between the crystal structures of cellulose from higher plants and from algae, the latter can be considered to *approximate* a smaller two-chain unit cell, similar to that for ramie cellulose. The larger cell would then result from some subtle differences in the packing of second-nearest-neighbor chains that would double the lateral dimensions of the unit cell. Crystal structures of *Valonia* cellulose were determined on the basis of this approximate, smaller unit cell.[3,4]

Second, there were differences between the packings of the chains in the different structures determined on the basis of a two-chain unit cell. For the two-chain unit cell, there are three possible packings of the two (corner and center) chains, that are commonly referred to as "parallel up", "parallel down", and "antiparallel."[4,7,8] The (probably most reliable) crystal structure analyses, refs 3−5, indicate that antiparallel packing is inconsistent with the X-ray diffraction data. Although there have been some proposals for antiparallel packing of native cellulose (e.g. ref 9), further detailed analysis of the X-ray diffraction data for ramie cellulose[8] and analysis of *Valonia* microfibrils using silver staining[10−12] and cellobiohydrolase digestion[13] have all confirmed parallel packing for native cellulose. The crystal structures determined for native cellulose[3−5] are all consistent, on the face of it, since they all describe "parallel up" packing of the chains. However, as has been noted by a number of authors,[7,11,14] due to different definitions of the unit cell dimensions, the ramie structure[5] and one of the *Valonia* structures[3] have the same packing, whereas the other *Valonia* structure[4] has a different packing (opposite polarity).[7] The X-ray data for ramie cellulose has been reanalyzed,[8] and it seems clear that the polarity of the original structure determination[5] is correct. However, as a result of the above discrepancies, the precise packing of native *Valonia* cellulose, and whether it is different to the packing of ramie cellulose, is still unclear. This is the subject of this paper.

Recent molecular mechanics and dynamics studies have shown a preference for the parallel up structure of cellulose Iβ;[15−17] however, experimental corroboration of these results is important. A recent intricate experimental study[18] that used cellobiohydrolase digestion and silver-labeling to determine chain direction in a microfibril, followed by measuring electron diffraction patterns for different tilts of the same microfibril, gives strong evidence for parallel up packing. However, some doubt about the polarity of the packing still remains because of the current inconsistent X-ray results for *Valonia*.

Another important step in our understanding of native cellulose structure was the observation of multiple environments for some of the same carbon atoms from high-resolution CP/MAS ^{13}C NMR spectroscopy,[19,20] which implied that native cellulose consists of two allomorphs, which were labeled Iα and Iβ. This result was confirmed by electron diffraction analysis,[21] and it was concluded that the *Valonia* and bacterial celluloses are rich in Iα whereas ramie and cotton celluloses are dominated by Iβ. This work also showed that the Iα phase could be converted to the Iβ phase in the solid state. Further electron diffraction analysis[14,22] showed

* To whom correspondence should be addressed. Telephone: 765-494-9272. Fax: 765-494-7953. E-mail: rmillane@purdue.edu.
[†] Present address: USDA, Peoria, IL 61604-3999.

10.1021/ma9804895 CCC: $15.00
Published on Web 10/07/1998

that the Iβ phase corresponds to the two-chain monoclinic unit cell determined by X-ray diffraction, and the Iα phase to a one-chain triclinic unit cell. The electron diffraction results provide a very satisfactory interpretation of the X-ray diffraction data, since the triclinic unit cell for Iα indexes all the "extra" X-ray reflections that originally demanded the eight-chain monoclinic unit cell. The *Valonia* X-ray diffraction patterns are therefore due to a mixture of two phases, with monoclinic and triclinic unit cells, rather than to a single phase with an eight-chain monoclinic unit cell. Although the structure of the triclinic phase has not been determined in detail, the relationship between the triclinic and monoclinic unit cells indicates that the structure of the Iα phase is very similar to that of the Iβ phase, the primary difference being a unidirectional axial shift of each subsequent sheet of cellulose molecules in the Iα structure.[14] The primary difference between the Iα and Iβ structures is therefore that some chains in one structure are shifted by half the repeat distance (c) relative to those in the other. This explains the relatively good agreement obtained when structures for *Valonia* based on the monoclinic unit cell are refined against the X-ray data, since the $c/2$ shift of the chains produces only small changes in the diffraction.

We describe here a reanalysis of the X-ray fiber diffraction data for native *Valonia* cellulose, based on a monoclinic unit cell, as reported by Sarko and Muggli[3] and Gardner and Blackwell[4] (henceforth referred to here as SM and GB, respectively), in an effort to resolve the question of the polarity of the packing and the relationship of the packing to that of ramie cellulose. As described above, the presence of the Iα phase is expected to have only a small effect on such an X-ray analysis. Furthermore, as will be shown, steric considerations are also important in distinguishing between the different packing models. Resolution of the packing of monoclinic *Valonia* cellulose is an important prelude to the more difficult problem of defining the structure of the mixed system. For example, the triclinic structure also has two possible, "up" and "down", packing arrangements, and in view of the solid-state I$\alpha \rightarrow$ Iβ transition, it is probable that the polarities of the two allomorphs are identical.

Experimental Section

X-ray Data. The X-ray fiber diffraction data used in this study were those published by SM[3] and GB.[4]

Unit Cell Determination. Unit cell dimensions were determined by minimizing the quantity $\Sigma_i (\rho_i^o - \rho_i^c)^2$, where ρ_i^o is the length of the reciprocal vector for the ith observed reflection and ρ_i^c is that calculated from the unit cell constants.

Crystal Structure Refinement. Molecular models were constructed and refined using the linked-atom least-squares (LALS) technique.[23] The D-glucosyl ring was initially set in the standard conformation,[24] and the glycosidic bridge angle was initially set to 116.5°. The space group is $P2_1$, each chain has 2_1 screw symmetry, and two chains pass through the unit cell at positions (0,0) and ($^1/_2$, $^1/_2$) in the $a-b$ plane. The two chains have independent conformations. Variable parameters were the glycosidic linkage conformation angles $\phi = \theta$(O5−C1−O4−C4) and $\psi = \theta$(C1−O4−C4−C5), the conformation angle $\chi = \theta$(C4−C5−C6−O6) for the primary hydroxyl group, the glycosidic bridge bond angle, the conformation angles and the endocyclic bond angles of the glucosyl ring (restrained to standard values), of each of the two chains, the rotations of the corner (μ_1) and center (μ_2) chains about their helix axes, the fractional translation of the center chain along its axis relative to the corner chain (w), and X-ray scale and overall isotropic temperature (B) factors.

In the LALS technique, the variable parameters are adjusted so as to minimize

$$\Omega = \Sigma_m w_m(F_m^o - F_m^c)^2 + \Sigma_m k_m(d_m - d_m^o)^2 + \Sigma_m e_m(\theta_m - \theta_m^o)^2 + \Sigma_m \lambda_m G_m \quad (1)$$

The first term represents the differences between the observed, F_m^o, and calculated, F_m^c, structure amplitudes. The second term involves close nonboned interatomic distances, d_m, that are driven beyond normally accepted contact limits. The third term is used to restrain certain parameters or quantities, such as conformation angles, bond angles, or hydrogen bond lengths, to their expected values. The last term is used to satisfy exact relationships (constraints) such as helix connectivity and ring-closure which are satisfied when $G_m = 0$, and the λ_m are Lagrange multipliers. The weights k_m are chosen to match nonbonded interatomic potentials and the weights e_m are inversely proportional to the variance of the values observed in crystal structures.[23,24] Constant X-ray weights (w_m) were used.

Consistency of a model with the observed X-ray data is quantified using the conventional

$$R = \frac{\Sigma_m |F_m^o - F_m^c|}{\Sigma_m F_m^o} \quad (2)$$

and quadratic

$$R'' = \frac{\Sigma_m (F_m^o - F_m^c)^2}{\Sigma_m (F_m^o)^2} \quad (3)$$

crystallographic R factors. Competing crystal structures are assessed by using the ratios of the respective values of R'' in Hamilton's test.[25] Taking the applied constraints and the effective data introduced by the nonbonded interatomic contacts into account, the data/parameter ratio for the refinements is typically about 9.

Results

The *Valonia* cellulose I crystal structures were determined using each of the published X-ray data sets from SM[3] and GB[4] and are described in turn below.

To avoid any confusion, it is necessary to clearly define the unit cell conventions and the definition of the polarity of the crystal packing. Most authors have defined the monoclinic unit cell dimensions of cellulose I such that a is the smaller dimension and b is the larger dimension. The cellulose sheets then lie along the b-axis. This is the convention adopted here. Following common convention also, a molecule is defined as being oriented "up" in the unit cell if the z-coordinate of O5 is greater than that of C5; otherwise, it is oriented "down". This means that the (1 → 4) linkage is directed along the $+ve$ z-axis for an up molecule, and along the $-ve$ z-axis for a down molecule.

Refinement Against the SM X-ray Data. In the original SM analysis,[3] one parallel packing and the antiparallel packing were investigated. The parallel packing that they used actually corresponds to parallel up, so that the parallel down packing was not investigated. Although the structures were sterically refined using least-squares methods, optimization against the X-ray data was performed using a search procedure. The cellulose chain was fixed except for the rotation of the hydroxymethyl group. Hydrogen atoms were included. The antiparallel model was excluded in favor of the parallel model on the basis of better X-ray agreement. The crystallographic R factor for the final structure was

Table 1. Unit Cells Constant (and Standard Deviations) for *Valonia* Cellulose Iβ Calculated from the SM and GB X-ray Data

X-ray data	ref	a (Å)	b (Å)	c (Å)	γ (Å)
SM	3	7.88	8.21	10.34	96.8
SM	this work	7.85	8.27	10.38	96.3
GB	4	7.86	8.17	10.38	97.0
GB	this work	7.82	8.16	10.32	97.5

Table 2. Results from Refinement of Models of *Valonia* Cellulose Iβ against the SM X-ray Data

packing	R	R''	overshort interatomic distances
parallel up	0.238	0.213	none
parallel down	0.308	0.253	H6- - -H1 = 1.71 Å
antiparallel	0.359	0.311	none

$R = 0.32$. A more definitive analysis of the SM data is therefore expected by (1) considering the parallel down packing, (2) allowing variation of the molecular conformation, (3) allowing the two chains to be conformationally independent, and (4) conducting a full joint steric-X-ray least-squares refinement.

Spacings for 48 reflections (excluding those assigned to a triclinic phase) reported by SM[3] were used to refine the unit cell constants (although two of the reflections did not fit particularly well). The cell constants obtained are listed in Table 1 and are very similar to those obtained by SM.

The amplitudes of 48 above-threshold reflections were taken from SM[3] and multiplied by $\exp(26\rho^2/4)$ to remove an artificial isotropic temperature factor that had been applied.[3] The amplitudes of 11 below-threshold reflections were assigned threshold values equal to half that of the weakest observed reflection.

Crystal structure models for parallel up, parallel down, and antiparallel packing were constructed and refined against this X-ray data as described in the Experimental Section. A variety of relative translations, w, of the corner and center chains were explored, as were the three different domains for the hydroxymethyl conformation. In all cases, models with $w \approx 0.25$ and with the hydroxymethyl groups in the $\chi \approx -60°$ (tg) domain were substantially superior to models with other values of w and χ in the other domains.

The results of the refinements for the best model for each of the three packing modes are listed in Table 2. Inspection of the table shows that the parallel down model must be rejected since the contact between H6 and H1 of adjacent sheets is too short (the "outer limit" for H- - -H contacts is 1.90 Å[26]), and the R factor is also considerably higher than that for the parallel up model. This short contact can be relieved only by moving the chains away from the quarter-stagger ($w \approx 0.25$) position, which drives the R factor up even further and introduces other tight contacts. The X-ray agreement for the parallel up model is substantially better than that for the antiparallel model. Applying Hamilton's significance test to the ratio of R'' for the two models (for a one-dimensional hypothesis corresponding to changing the direction of one of the chains) shows that the former is preferred over the latter with better than 0.995 significance. The SM X-ray data therefore unequivocally support a parallel up structure. The R factors for this structure are low, indicating good data and an accurate structure. The varied bond angles and ring conformation angles differ from standard values by no more than 2 and 3°, respectively. The structure is qualitatively the same as that obtained by SM, although the fit with the X-ray data is much better. The conformation and packing parameters, atomic coordinates, and calculated and observed structure amplitudes, are listed in Tables 3−5, respectively. Views of the crystal structure are shown in Figure 1.

Table 3. Stereochemical Features of the Two Refined *Valonia* Cellulose Iβ Structures

	value			
	SM data		GB data	
parameter	corner	center	corner	center
ϕ (deg)	−95	−95	−95	−95
ψ (deg)	−145	−145	−143	−144
χ (deg)	−74	−75	−76	−77
O3- - -O5 (Å)	2.72	2.71	2.71	2.71
O2- - -O6 (Å)	2.71	2.80	2.68	2.69
O3- - -O6 (Å)	2.90	2.87	2.76	2.81
$\Delta\tau$ (deg)	0.0		−3.5	
w	0.249		0.258	
B (Å^2)	6.4		5.2	

Table 4. Atomic Coordinates (Å) for the Refined Cellulose Iβ Structure Based on the SM X-ray Data[a]

	corner chain			center chain		
Atom	x	y	z	x	y	z
C1	0.061	-0.362	0.474	4.029	3.349	3.071
C2	-0.324	-1.455	-0.516	3.747	2.226	2.081
C3	4.285	-1.145	-1.875	4.289	2.573	0.703
C4	-0.144	0.243	-2.333	3.782	3.941	0.266
C5	0.153	1.281	-1.256	4.072	4.988	1.335
C6	-0.412	2.646	-1.589	3.491	6.345	0.997
O2	0.121	-2.713	-0.023	4.329	1.019	2.561
O3	-0.142	-2.123	-2.825	3.875	1.581	-0.239
O4	0.572	0.621	-3.507	4.434	4.352	-0.933
O5	-0.425	0.895	0.000	3.496	4.583	2.586
O6	0.346	3.300	-2.606	4.251	7.011	-0.010
H1	1.156	-0.320	0.568	5.116	3.446	3.203
H2	-1.420	-1.506	-0.597	2.662	2.059	2.015
H3	1.382	-1.189	-1.804	5.389	2.581	0.733
H4	-1.222	0.225	-2.551	2.698	3.875	0.090
H5	1.242	1.383	-1.135	5.160	5.103	1.450
H61	-1.461	2.542	-1.904	2.447	6.225	0.670
H62	-0.439	3.263	-0.679	3.443	6.961	1.908

[a] Coordinates for the first residue only are shown; the second residues being generated by a 2-fold screw rotation about axes at (0,0) for the corner chain, and (3.901, 3.702) for the center chain.

Refinement Against the GB X-ray Data. In the original GB analysis, all three possible packings (parallel up, parallel down, and antiparallel) were investigated.[4] Molecular models were constructed with standard stereochemistry and consistent with the observed repeat distance and symmetry. Crystal structures were then constructed and optimized, using least-squares and rigid cellulose chains except for χ, against the X-ray data. Hydrogen atoms were not included however. On the basis of the X-ray agreement for certain sets of structure amplitudes (either observed or all reflections) and the resulting hydrogen bonding pattern, the parallel down and antiparallel structures were rejected in favor of the parallel up structure. The final structure has a quadratic crystallographic R factor $R'' = 0.22$. However, as has been noted previously,[7] because the GB unit is defined with $a > b$, their parallel up structure is in fact *parallel down* in our definition. The GB structure therefore has a polarity opposite to that of the SM

Table 5. Calculated and Observed Structure Amplitudes for the Refined Cellulose Iβ Structure Based on the SM X-ray Data[a]

l	h	k	F_c	F_o
0	0	1	1.1	(3.0)
	1	0	0.6	(3.0)
	-1	1	100.3	84.3
	1	1	109.4	91.8
	0	2	18.4	15.3
	2	0		
	-1	2	262.1	271.8
	-2	1	9.7	12.1
	1	2	1.5	17.2
	2	1	14.9	16.4
	-2	2	8.3	17.5
	0	3		
	2	2	18.1	38.7
	-1	3		
	3	0		
	-3	1	35.0	38.9
	1	3	19.3	13.6
	3	1	35.4	41.5
	-2	3	0.7	(3.0)
	-3	2	4.4	(3.0)
	2	3	5.7	(3.0)
	-1	4		
	-3	3	20.5	(3.0)
	3	2		
	0	4	31.5	44.8
	-4	1		
	4	0	74.4	72.9
	1	4		
	-2	4	46.0	39.7
1	1	0		
	0	1	31.3	13.5
	-1	1	10.2	6.0
	1	1	26.5	12.4
	0	2	15.9	25.1
	2	0		
	-1	2	26.7	19.5

l	h	k	F_c	F_o
1	-2	1	4.8	12.3
	1	2	10.2	10.7
	2	1	26.5	18.4
	-2	2	25.9	23.0
	0	3	52.7	73.4
	2	2		
	-1	3	60.3	66.5
	-3	1		
	3	0	37.8	64.8
	1	3	34.3	18.6
	3	1		
	-2	3	56.9	48.5
	-3	2	21.8	49.7
2	0	0	0.3	W
	0	1		
	1	0	54.5	58.6
	-1	1		
	1	1	4.4	(3.0)
	0	2	1.3	(3.0)
	2	0		
	-1	2		
	-2	1	47.0	58.0
	1	2		
	2	1	55.9	65.4
	-2	2	4.1	(3.0)
	0	3	25.4	41.7
	-1	3		
	2	2	3.2	(3.0)
	3	0		
	-3	1		
	1	3		
	3	1		
	-2	3		
	-3	2	48.7	54.6
	3	2		
	2	3	20.5	26.9

l	h	k	F_c	F_o
3	1	0		
	0	1	33.2	33.5
	-1	1		
	1	1	41.2	38.8
	0	2	53.7	73.5
	2	0		
	-2	1		
	-1	2	51.5	58.1
	2	1		
	1	2	59.1	47.0
	-2	2	32.8	30.6
	0	3	47.0	47.1
	2	2		
	-1	3		
	3	0	42.7	36.0
	-3	1		
	1	3	50.6	23.9
	3	1		
	-2	3	32.3	15.8
4	0	0	76.3	M
	1	0		
	0	1	1.4	(3.0)
	-1	1		
	1	1	32.9	38.1
	0	2		
	2	0		
	-1	2	62.8	50.8

[a] Values for overlapping reflections are given next to the last (hk) value in the group. Parentheses indicate threshold values. W and M denote weak and medium amplitude meridional reflections (that were not included in the refinement).

structure. A more definitive analysis of the GB data is therefore expected by (1) including the hydrogen atoms, (2) allowing variation of the molecular conformation, (3) allowing the two chains to be conformationally independent, and (4) conducting a joint steric-X-ray refinement.

Spacings for the 36 reflections (excluding three reflections assigned to the larger eight-chain unit cell) reported by GB[4] were used to refine the cell constants. The cell constants obtained are listed in Table 1 and are very similar to those obtained by GB.

The amplitudes of 36 above-threshold reflections were taken from GB.[4] The 43 planes for which the observations were below the threshold of observation were combined, based on their spacings, into 33 below-threshold reflections and assigned threshold values based on the values given by GB.[4]

Crystal structure models for parallel up, parallel down, and antiparallel packing were constructed and refined against this X-ray data as described in the Experimental Section. A variety of relative translations, w, of the corner and center chains were explored, as were the three different domains for the hydroxymethyl conformation. In all cases, models with $w \approx 0.25$ and with the hydroxymethyl groups in the $\chi \approx -60°$ (tg) domain were substantially superior to models with other values of w, and χ in the other domains.

The results of the refinements for the best model for each of the three packing models are listed in Table 6. Inspection of the table shows that although the X-ray agreement for the parallel down model is comparable to that for the parallel up model, the parallel down model must be rejected since, as with the SM model, the contact between H6 and H1 of adjacent sheets is much too short. This short contact can be relieved only by moving the chains away from the quarter-stagger position, which drives the R factor up to $R > 0.30$ and introduces other tight contacts. The X-ray agreement for the parallel up model is significantly better than for the antiparallel model, and applying Hamilton's significance test to the ratio of R'' for the two models shows that the former is preferred over the latter at better than 0.995 significance. The GB X-ray data therefore unequivocally support a parallel up structure. The R

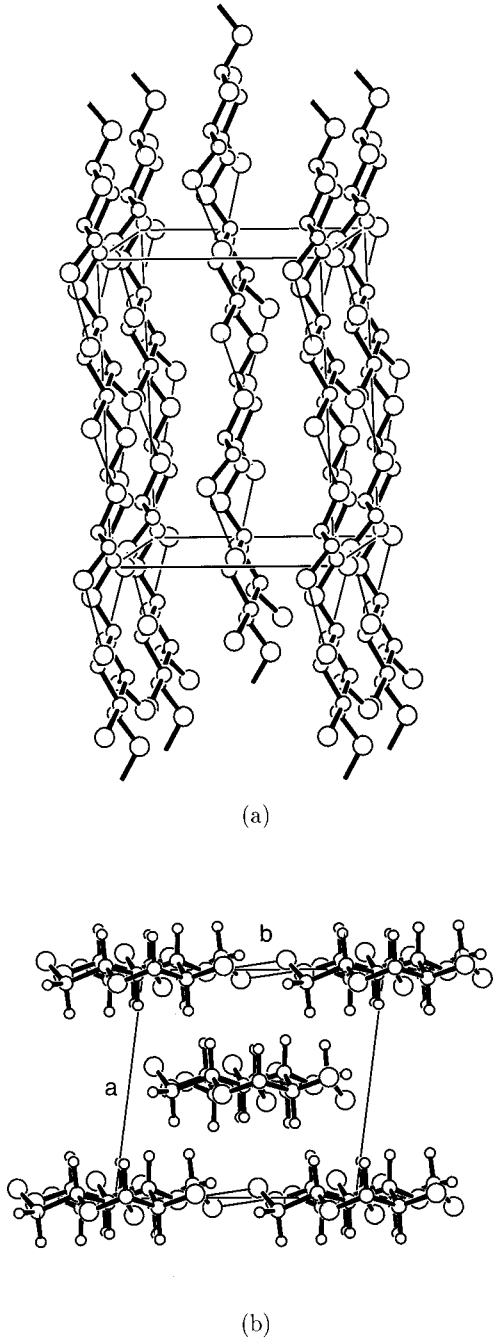


Figure 1. Views of the refined cellulose Iβ crystal structure (a) obliquely to and (b) along the *c*-axis. Thin lines show hydrogen bonds. The hydrogen atoms are excluded from part a for clarity.

Table 6. Results from Refinement of Models of *Valonia* Cellulose Iβ against the GB X-ray Data

packing	R	R''	overshort interatomic distances (Å)
parallel up	0.213	0.199	none
parallel down	0.221	0.218	H6- - -H1 =1.60 Å
antiparallel	0.251	0.240	none

factors for the structure are low, indicating good data and an accurate structure. The varied bond angles and ring conformation angles differ from standard values by no more than 2° and 3°, respectively. We therefore conclude that the parallel down model preferred by GB is incorrect, probably due to exclusion of the hydrogen atoms so that the short H- - - H contact did not show up. The conformation and packing parameters, atomic coordinates, and calculated and observed structure amplitudes, are listed in Tables 3, 7, and 8, respectively. The crystal structure appears insignificantly different from that shown in Figure 1.

Table 7. Atomic Coordinates (Å) for the Refined Cellulose Iβ Structure Based on the GB X-ray Data[a]

	corner chain			center chain		
Atom	x	y	z	x	y	z
C1	0.042	-0.359	0.480	3.977	3.221	3.149
C2	-0.321	-1.457	-0.511	3.632	2.106	2.170
C3	0.328	-1.149	-1.852	4.231	2.402	0.804
C4	-0.119	0.231	-2.317	3.785	3.790	0.362
C5	0.154	1.282	-1.247	4.134	4.847	1.403
C6	-0.416	2.639	-1.603	3.621	6.223	1.034
O2	0.107	-2.714	0.001	4.113	0.867	2.680
O3	-0.052	-2.142	-2.808	3.796	1.416	-0.134
O4	0.607	0.622	-3.481	4.467	4.174	-0.831
O5	-0.445	0.896	0.000	3.537	4.488	2.659
O6	0.370	3.300	-2.593	4.431	6.840	0.034
H1	1.135	-0.311	0.588	5.068	3.240	3.290
H2	-1.414	-1.506	-0.621	2.539	2.020	2.084
H3	1.423	-1.176	-1.745	5.328	2.357	0.867
H4	-1.193	0.193	-2.550	2.702	3.769	0.175
H5	1.240	1.394	-1.110	5.227	4.909	1.512
H61	-1.451	2.521	-1.955	2.581	6.145	0.685
H62	-0.484	3.258	-0.696	3.581	6.854	1.934

[a] Coordinates for the first residue only are shown; the second residues being generated by a 2-fold screw rotation about axes at (0,0) for the corner chain, and (3.877, 3.573) for the center chain.

Discussion

Full refinement of the cellulose Iβ crystal structure against two independent X-ray data sets shows, unequivocally, that *Valonia* cellulose I packs in a parallel up arrangement. Doubts about the packing as derived from the two X-ray studies conducted in the 1970s resulted from an inability to conduct full refinements, primarily as a result of the limited computing power available at that time. The two structures obtained are practically identical and have very good stereochemistry and X-ray agreements, indicating precise structures. An independent assessment of the reliability of a structure determined by X-ray fiber diffraction can be obtained by comparing the R factor for the structure with the "largest likely R factor" for the particular X-ray data set used.[27−29] For a well-determined structure, the R factor obtained should be less than half the largest likely R factor. The largest likely R factors for the two X-ray data sets used here are calculated[28] to be 0.45 for the SM data set and 0.43 for the GB data set. The R factors obtained for the structures reported here therefore indicate well-determined structures. Our results are consistent with recent molecular modeling and electron diffraction results.[15−18]

The molecular conformation, sheet structure, and hydrogen-bonding pattern are similar to those reported previously[3,4] (Table 3). What this study defines more precisely is the packing of the sheets relative to each other. Inspection of Table 3 shows that the conformations of the corner and center chains are essentially identical. A difference in the orientations of the corner and center chains about their axes is an important parameter related to any differences between the sheets containing the corner and center chains. Since the

Table 8. Calculated and Observed Structure Amplitudes for the Refined Cellulose Iβ Structure Based on the GB X-ray Data[a]

l	h	k	F_c	F_o
0	0	1		
	1	0	0.8	(11.5)
	-1	1	103.5	104.9
	1	1	94.4	73.7
	0	2	16.2	(13.6)
	2	0		
	-1	2	264.8	270.9
	-2	1	6.4	(8.1)
	1	2		
	2	1	0.2	(11.5)
	-2	2	13.6	(8.1)
	0	3	0.9	(8.1)
	-1	3		
	2	2	35.2	39.3
	3	0		
	-3	1	31.2	42.0
	1	3	18.3	33.1
	3	1		
	2	3	23.9	28.5
	-3	2	2.0	(10.8)
	2	3	2.9	(10.8)
	3	2	2.1	(10.8)
	0	4		
	-1	4	35.3	38.2
	-3	3	31.3	(24.4)
	-4	1		
	4	0	78.3	95.1
	1	4	0.8	(24.4)
	-2	4	46.6	(24.4)
1	0	1		
	1	0	36.3	20.9
	-1	1	10.2	(13.6)
	1	1	26.4	14.9
	0	2	15.4	19.2
	2	0	23.0	(13.6)
	-1	2		
	-2	1	22.0	28.5
	1	2		
	2	1	25.6	(19.2)

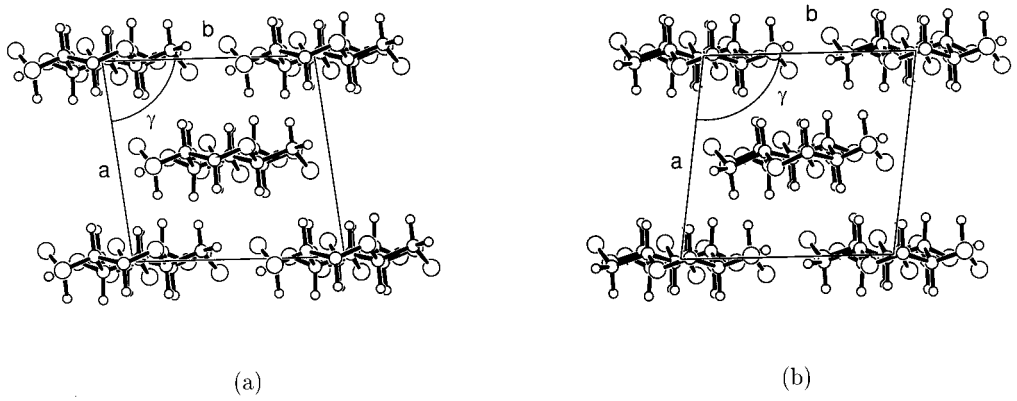
l	h	k	F_c	F_o
1	-2	2	25.4	(13.6)
	0	3	55.8	66.4
	-1	3		
	2	2	61.4	62.3
	-3	0		
	3	-1	43.8	65.3
	1	3	35.2	41.2
	3	1		
	-2	3		
	-3	2	65.3	48.5
	2	3	25.9	(24.4)
	3	2	8.3	(24.4)
	0	4		
	-1	4	29.4	(24.4)
2	0	0	1.4	W
	0	1		
	1	0	54.1	51.5
	-1	1	4.0	(13.6)
	1	1	5.1	(13.6)
	0	2	3.4	(13.6)
	2	0	4.8	(13.6)
	-1	2		
	-2	1	47.7	58.8
	1	2		
	2	1	51.9	62.9
	-2	2	6.1	(16.3)
	0	3		
	-1	3		
	2	2	29.9	25.7
	3	0		
	-3	1	19.5	(23.0)
	1	3	3.1	(16.3)
	3	1		
	-2	3		
	-3	2	54.9	68.0
	2	3	26.4	(24.4)
	3	2	4.2	(24.4)
	0	4		
	-1	4	39.8	28.5

l	h	k	F_c	F_o
3	1	0		
	0	1	29.6	35.2
	-1	1	24.0	34.1
	1	1	40.8	35.2
	2	0		
	0	2	63.2	71.8
	-1	2		
	-2	1	49.2	61.2
	1	2		
	2	1	56.0	36.3
	-2	2	33.8	(16.3)
	0	3		
	-1	3		
	2	2	67.3	65.3
	3	0		
	-3	1	21.7	59.9
	1	3	52.9	38.2
4	0	0		
	1	0		
	0	1	55.9	M
	-1	1	17.3	(16.3)
	1	1	32.5	20.9
	0	2		
	2	0		
	-1	2		
	-2	1	65.9	74.8
	1	2		
	2	1	5.8	(34.5)
	-2	2	33.1	(24.4)
	0	3	4.5	(24.4)
	-1	3		
	2	2	54.9	48.5

[a] Values for overlapping reflections are given next to the last (*hk*) values in the group. Parentheses indicate threshold values. W and M denote weak and medium amplitude meridional reflections (that were not included in the refinement).

corner and center chains in this analysis are not conformationally identical, one cannot define an exact rotational relationship between the two chains. However, since C6 is the glucosyl atom most distant from the helix axis, we describe the relative orientation of the two chains by $\Delta\tau = \tau$(center chain) $- \tau$(corner chain) where τ is the cylindrical polar angle of C6. The values of $\Delta\tau$ are 0 and $-3.5°$ for the structures based on the SM and GB data, respectively (Table 3). Larger values of this angle ($-11.5°$) have been calculated from molecular dynamics studies and have been related to the stability of the Iβ phase relative to the Iα phase.[16] The effect of increasing the angle between the chains was investigated by forcing $\Delta\tau$ to be $\pm 10°$, while refining μ_1 and μ_2, for both structures. For $\Delta\tau = -10°$, the *R* factors increased by 0.03 and 0.02 for the SM and GB structures, respectively. For $\Delta\tau = 10°$, the *R* factors increased by 0.07 and 0.05, for the SM and GB structures, respectively. Acceptable stereochemistry was retained in the modified structures. The rather small increases in the *R* factor for $\Delta\tau$ negative indicate that larger rotations between the chains are not inconsistent with the X-ray data. The opposite direction of rotation ($\Delta\tau$ positive) leads to larger increases in the *R* factor and is not supported by the X-ray data. The preferred direction of the rotation is the same as that calculated by molecular dynamics.[16] Releasing the center chain from the ($^1/_2$, $^1/_2$) position led to shifts of the chain axis of less than 0.05 Å, and neither the X-ray agreement nor the stereochemistry was improved. The intermolecular (O3- - -O6) hydrogen bonds are longer for the structure based on the SM data because of the slightly larger value of *b* in this case. Other small variations in the lengths of these hydrogen bonds are related to the different values of μ_1 and μ_2.

The packing of *Valonia* cellulose is therefore the same as that of ramie cellulose, suggesting the same packing throughout the algae and higher plants and probably



(a) (b)

Figure 2. Views of the cellulose Iβ crystal packing with the (1 → 4) linkage directed out of the page for (a) parallel down and (b) parallel up packing.

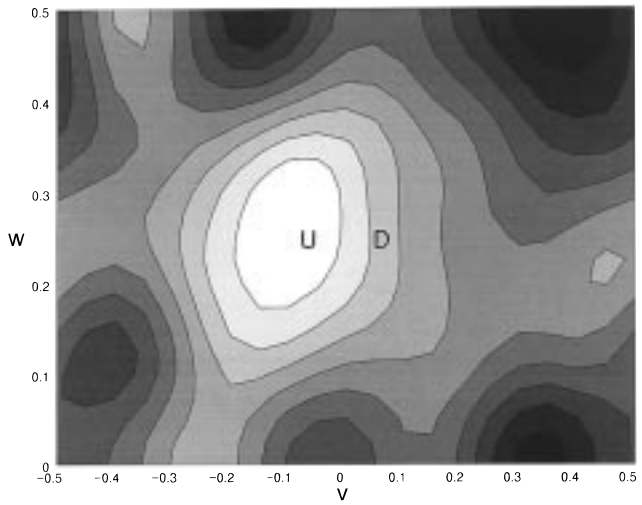


Figure 3. Contour plot of the steric compression (arbitrary units) between two sheets of cellulose molecules separated by 3.88 Å as a function of their relative positions described by v and w. The darker regions denote larger steric compression. The parallel up and parallel down positions are marked by U and D, respectively.

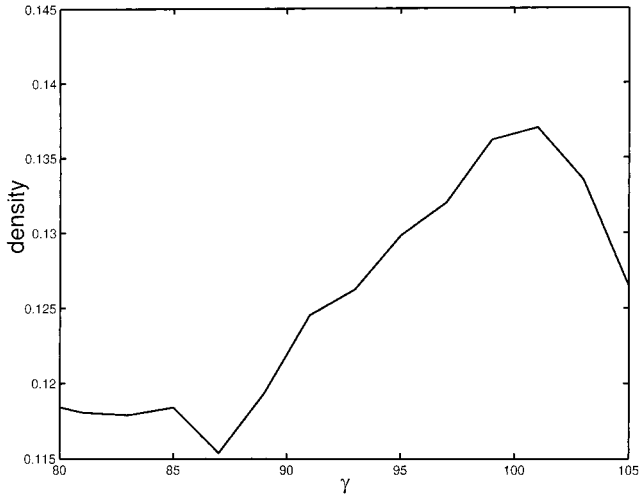


Figure 4. Packing density (arbitrary units) for sheets of cellulose molecules packed in a crystalline array as a function of γ.

for bacteria as well. This indicates considerable commonalties in the mechanism of biosynthesis of cellulose between different species, although there are clearly some differences that result in different proportions of the Iα and Iβ phases. The packing of the sheets in native cellulose has a direct effect on the structure of the surface of cellulose microfibrils and is expected to be relevant to the manner in which molecules such as xyloglucans interact with the microfibril surface.[27]

The results presented here have important implications for the structure of cellulose Iα. Cellulose Iα, since it contains one molecule per unit cell, is necessarily a parallel packed structure. However, both "up" and "down" packings exist as possibilities for the triclinic structure. Given the relationship between the monoclinic and triclinic unit cells,[14] as well as the Iα → Iβ transformation by annealing in the solid state, it is likely that cellulose Iα also packs in an "up" fashion.

An interesting question is "why does cellulose pack in the parallel up, rather than the parallel down, arrangement"? Parts a and b of Figure 2 show the parallel up and parallel down packing, respectively, displayed such that the polarities of the molecular chains are the same. Defining the unit cells this way means that $\gamma \approx 83°$ and $\gamma \approx 97°$ in parts a and b, respectively (Figure 2). Hence, relative to the molecules, parallel down and parallel up packings correspond to $\gamma \approx 83°$ and $\gamma \approx 97°$, respectively. The question posed above can then be recast as "why does cellulose pack with $\gamma \approx 97°$ rather than $\gamma \approx 83°$"? Or more generally as "why does cellulose pack with $\gamma \approx 97°$, rather than some other value of γ"? Referring to Figure 2, what γ determines, together with w, is the relative position with which the two sheets of cellulose molecules pack. Two calculations were made to shed some light on the answer to this question.

First we considered two identical sheets of cellulose molecules (the corner chains in our refined GB structure) that are separated by 3.88 Å, the sheet separation in the crystal structure. The steric compression (the second term in eq 1) was calculated as a function of the relative position of the two sheets. The relative position of the sheets is defined by v and w, which are the fractional shifts perpendicular and parallel to the molecular axes, respectively. The molecules in the two sheets exactly interleave when $v = 0$. The results are shown in Figure 3. The parallel up position corresponds to $(v, w) = (-0.063, 0.25)$ and the parallel down position to (0.063, 0.25), shown by U and D, respectively, in Figure 3. Inspection of Figure 3 shows that there is a single minimum near the parallel up position. The parallel up position therefore allows the sheets to pack closer together. The packing energy is considerably higher at the parallel down position.

The analysis described above is for two identical sheets of cellulose molecules. In the second calculation we considered the whole crystal structure and allowed the structure of the sheets, as well as their packing, to vary. This was done by varying γ for the crystal structure between 80 and 105° and, while allowing μ_1, μ_2, and w to vary, finding the minimum value of a for which the structure has no serious overshort contacts. Serious overshort contacts were considered to be those that are shorter than the "outer limits".[26] From this calculation one can calculate an approximate packing density of a cellulose crystal structure as a function of γ, as shown in Figure 4. Inspection of Figure 4 shows that the parallel up packing ($\gamma \approx 97°$) is close to the predicted maximum of the packing density at $\gamma \approx 101°$. This calculation, based only on the single shortest contact, is only approximate of course, but one can see that the parallel down packing ($\gamma \approx 83°$) is well away from the maximum density. The cellulose sheets therefore pack parallel up because this arrangement minimizes the packing energy and maximizes the crystal density.

Acknowledgment. This work was supported by the National Science Foundations (DBI-9722862) and the Industrial Consortium of the Whistler Center for Carbohydrate Research.

References and Notes

(1) Mark, H,; Meyer, K. *H. Ber. Dtsch. Chem. Ges.* **1928**, *61B*, 593.
(2) Meyer, K. H.; Misch, L. *Helv. Chim. Acta* **1937**, *20*, 232−244.
(3) Sarko, A.; Muggli, R. *Macromolecules* **1974**, *7*, 486-494.
(4) Gardner, K. H.; Blackwell, *Biopolymers* **1974**, *13*, 1975−2001.
(5) Woodcock, C.; Sarko, A. *Macromolecules* **1980**, *13*, 1183−1187.
(6) Nieduszynski, I.; Atkins, E. *Biochim. Biophys. Acta* **1970**, *222*, 109−118.
(7) French, A. D.; Howley, P. S. In *Cellulose and Wood*; Schuerch, C., Ed.; Wiley: New York, 1989; p 159.
(8) Millane, R. P.; Narasaiah, T. V. In *Cellulose and Wood*; Schuerch, C., Ed.; Wiley: New York, 1989; p 39.
(9) French, A. D. *Carbohydr. Res.* **1978**, *61*, 67−80.
(10) Heita, K.; Kuga, S.; Usuda, M. *Biopolymers* **1984**, *23*, 1807.
(11) Kuga, S.; Brown Jr, R. M. In *Cellulose and Wood*; Schuerch, C., Ed.; Wiley: New York, 1989; p 677.
(12) Kuga, S.; Brown Jr, R. M. *Carbohydr. Res.* **1988,** *180*, 345−350.
(13) Chanzy, H.; Henrissat, B. *FEBS Lett.* **1985**, *184*, 285−288.
(14) Sugiyama, J.; Vuong, R.; Chanzy, H. *Macromolecules* **1991**, *24*, 4168−4175.
(15) Aabloo, A.; French, A. D.; Mikelsaar, R. H.; Pertsin, A. J. *Cellulose* **1994**, *1*, 161−168.
(16) Heiner, A. P.; Sugiyama, J.; Teleman, O. *Carbohydr. Res.* **1995**, *275*, 207−223.
(17) Kroon-Batenburg, L. M. J.; Bouma, B.; Kroon, J. *Macromolecules* **1996**, *29*, 5695−5699.
(18) Koyama, M.; Helbert, W.; Imai, T.; Sugiyama, J.; Henrissat, B. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9091−9095.
(19) Atalla, R. H.; van der Hart, D. L. *Science* **1984**, *223*, 283−285.
(20) van der Hart, D. L.; Atalla, R. *Macromolecules* **1984**, *17*, 1465−1472.
(21) Sugiyama, J.; Okano, T.; Yamamoto, H.; Horii, F. *Macromolecules* **1990**, *23*, 3196−3198.
(22) Sugiyama, J.; Persson, J.; Chanzy, H. *Macromolecules* **1991**, *24*, 2461−2466.
(23) Smith, P.; Arnott, S. *Acta Crystallogr.* **1978**, *34*, 3−11.
(24) Arnott, S.; Scott., W. *J. Chem. Soc., Perkin Trans. 2* **1972**, *2*, 324−335.
(25) Hamilton, W. *Acta Crystallogr.* **1965**, *18*, 502−510.
(26) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. In *Aspects of Protein Structure*; Ramachandran, G. N., Ed.; Academic Press: New York, 1963; p 121.
(27) Stubbs, G. *Acta Crystallogr.* **1989**, *A45*, 254−258.
(28) Millane, R. P. *Acta Crystallogr.* **1989**, *A45*, 573−576.
(29) Millane, R. P., Stubbs, G. *Polym. Prepr.* **1992**, *33* (1), 249−250.
(30) Finkenstadt, V. L.; Hendrixson, T. L.; Millane, R. P. *Carbohydr. Chem.* **1995**, *14*, 601−611.

MA9804895