References and Notes

(1) (a) Central Research Laboratory, Mitsubishi Chemical Industries Ltd., Kawasaki, Kanagawa, Japan. (b) Katata Research Institute, Toyobo Co., Ltd., Otsu, Shiga, Japan.
(2) R. de P. Daubeny, C. W. Bunn, and C. J. Brown, *Proc. R. Soc. London, Ser. A*, **226**, 531 (1954).
(3) J. Bateman, R. E. Richards, G. Farrow, and I. M. Ward, *Polymer*, **1**, 63 (1960).
(4) I. Goodman, *Angew. Chem.*, **74**, 606 (1962).
(5) A. M. Joly, G. Nemoz, A. Douillard, and G. Vallet, *Makromol. Chem.*, **176**, 479 (1975).
(6) R. Jakeways, I. M. Ward, M. A. Wilding, I. H. Hall, I. J. Desborough, and M. G. Pass, *J. Polym. Sci., Polym. Phys. Ed.*, **13**, 799 (1975).
(7) I. M. Ward, private communication.
(8) M. Yokouchi, Y. Chatani, H. Tadokoro, and H. Tani, *Polym. J.*, **6**, 248 (1974).
(9) M. Yokouchi, H. Tadokoro, and Y. Chatani, *Macromolecules*, **7**, 769 (1974).
(10) A. Bondi, *J. Phys. Chem.*, **68**, 441 (1964).
(11) A. E. Tonelli, *J. Polym. Sci., Polym. Lett. Ed.*, **11**, 441 (1973).
(12) S. Arnott and A. J. Wonacott, *Polymer*, **7**, 157 (1966).
(13) Y. Takahashi, T. Sato, and H. Tadokoro, *J. Polym. Sci., Polym. Phys. Ed.*, **11**, 233 (1973).
(14) Y. Y. Tomashpol'skii and G. S. Markova, *Vysokomol. Soedin.*, **6**, 274 (1964).
(15) A. I. Kitaigorodskii, "Organic Chemical Crystallography", Academy of Science of USSR, Moscow, 1955 (in Russian).
(16) W. J. Dulmage and A. L. Geddes, *J. Polym. Sci.*, **31**, 499 (1958).
(17) H. Tadokoro, K. Tatsuka, and S. Murahashi, *J. Polym. Sci.*, **59**, 413 (1962).

Determination of the Structure of Cellulose II

F. J. Kolpak and J. Blackwell*

Department of Macromolecular Science, Case Western Reserve University, Cleveland, Ohio 44106.
Received October 2, 1975

Abstract: The structure of regenerated cellulose is shown by x-ray diffraction to be comprised of an array of antiparallel chain molecules. The determination was based on the intensity data from rayon fibers and utilized rigid-body least-squares refinement techniques. The unit cell is monoclinic with space group $P2_1$ and dimensions a = 8.01 Å, b = 9.04 Å, c = 10.36 Å (fiber axis), and γ = 117.1°. Models containing chains with the same sense (parallel) or alternating sense (antiparallel) were refined against the intensity data. The only acceptable model contains antiparallel chains. The $-CH_2OH$ groups of the corner chain are oriented near to the gt position while those of the center chain are near to the tg position. Both chains possess an O3–H···O5′ intramolecular hydrogen bond, and the center chain also has an O2′–H···O6 intramolecular bond. Intermolecular hydrogen bonding occurs along the 020 planes (O6–H···O2 bonds for the corner chains and O6–H···O3 bonds for the center chains) and also along the 110 planes with a hydrogen bond between the O2–H of the corner chain and the O2′ of the center chain. This center–corner chain hydrogen bonding is a major difference between the native and regenerated structures and may account for the stability of the latter form.

Cellulose is composed of 1 → 4 linked β-D-glucose residues, and is known to exist in at least four polymorphic crystalline forms, of which the structures and properties of cellulose I (native cellulose) and cellulose II (regenerated cellulose and Mercerized cellulose) have been most extensively studied. Cellulose II has a monoclinic unit cell with reported dimensions a = 7.93 Å, b = 9.18 Å, c = 10.34 Å, and γ = 117.3° (average values over a variety of cellulose II preparations).[1] The unit cell contains sections of two cellulose chains (four glucose residues) and the space group is generally thought to be $P2_1$. The cellulose chains are thought to possess twofold screw symmetry, and lie with their axes through the origin (0, 0, z) and the center (½, ½, z) of the unit cell. Thus, the space group requirements are satisfied whether the molecules have the same sense (parallel) or opposite sense (antiparallel).

Jones[2,3] compared the observed[4] and calculated x-ray intensities for models of the cellulose II structure consistent with the above requirements. The most favorable structure in terms of the infrared spectrum had the chains in an antiparallel arrangement with a relative shift between the glycosidic oxygens of the origin and center chain of approximately 0.29c. He was unable to determine the exact positions of the $-CH_2OH$ groups, but those of the center chain were thought to have an orientation different from those of the origin chain. More recently, Watanabe and Hayashi[5] proposed a "bent-twisted" model for the cellulose molecule. The proposed cellulose II structure with this modification of the twofold screw molecule had antiparallel chains with a relative shift of 0.185c. However, satisfactory match between the observed and calculated intensities was not achieved. In a potential energy packing analysis of the cellulose II structure, Sarko and Muggli[6] were unable to choose between a number of parallel and antiparallel chain models although they considered one antiparallel model to be most favorable because of extensive hydrogen bonding possibilities, where the $-CH_2OH$ side chains may have different rotational conformations.

Thus the major details of the structure, including the polarity of the two chains in the unit cell and the hydrogen bonding network, remain undetermined. Rigid-body least-squares refinement techniques have recently been used by Gardner and Blackwell[7,8] to determine the structure of native cellulose (*Valonia*). The results of their refinement showed a significant preference for parallel rather than antiparallel chains. In the two-chain monoclinic unit cell, the chain at (½, ½, z) is staggered by 0.266c with respect to the chain at the origin and the orientation of the $-CH_2OH$ side groups are close to the tg position (see below) for both chains. The chains possess two intramolecular hydrogen bonds (O3–H···O5′ and O6···H–O2′) (see Figure 2 for numbering of the atoms) and interchain hydrogen bonds occur only along the 020 planes of the unit cell, resulting in an array of staggered sheets of hydrogen bonded cellulose chains.

The approach used in this refinement of cellulose II was similar to that employed for the native structure. The x-ray intensity data were obtained for rayon fibers and the least-squares refinement method of Arnott and Wonacott[9] was used to refine possible parallel and antiparallel chain mod-

TABLE I

Crystallographic Data for Regenerated Cellulose: Final Antiparallel Model

d Spacing (obs)	d Spacing (calc)	hkl	F(obs)[a]	F(calc)[b]
Zero Level				
	8.05	010	(4.0)	7.5
7.18	7.21	$1\bar{1}0$	44.6	39.7
	7.13	100		
	4.49	$1\bar{7}0$	(4.5)	8.0
4.44	4.43	110	94.8	96.8
4.05	4.02	020	100.0	92.2
	4.00	$2\bar{1}0$	(4.5)	3.8
3.56	3.61	$2\bar{2}0$	8.3	7.4
	3.56	200		
3.02	3.00	$1\bar{3}0$	5.8	16.0
	2.97	120		
	2.86	$2\bar{3}0$	(6.0)	0.6
	2.82	210	(6.0)	3.9
2.65	2.68	030	11.2	16.9
	2.64	$3\bar{2}0$		
	2.63	$3\bar{1}0$		
	2.40	$3\bar{3}0$	(7.5)	1.2
	2.38	300	(7.5)	2.6
	2.24	$2\bar{4}0$	(7.5)	4.9
2.21	2.22	$1\bar{4}0$	25.6	29.5
	2.21	220		
	2.20	130		
	2.07	$3\bar{4}0$	(7.5)	1.2
2.02	2.04	310	9.7	7.4
	2.01	040		
	2.00	$4\bar{2}0$		
	1.95	$4\bar{3}0$	(7.5)	2.0
	1.94	$4\bar{1}0$	(7.5)	1.3
1.80	1.81	$2\bar{5}0$	7.8	6.2
	1.80	$4\bar{4}0$		
	1.79	230		
	1.78	400		
First Level				
	10.36	001		0.0
6.36	6.20	011	4.4	7.6
	5.92	$1\bar{1}1$	(4.5)	0.8
	5.87	101	(4.5)	8.5
4.12	4.12	$1\bar{7}1$	8.0	7.6
	4.07	111	(6.0)	20.0
3.77	3.75	021	7.7	6.3
	3.74	$2\bar{1}1$		
	3.40	$2\bar{2}1$	(6.0)	4.0
3.37	3.36	201	5.9	10.1
2.86	2.88	$1\bar{3}1$	8.4	12.7
	2.86	121		
	2.76	$2\bar{3}1$	(6.0)	9.6
	2.72	211	(6.0)	5.2
	2.60	031	(6.0)	5.9
2.56	2.56	$3\bar{7}1$	15.7	19.4
	2.55	$3\bar{1}1$		
2.36	2.34	$3\bar{3}1$	17.0	19.3
	2.32	301	(6.0)	5.6
2.18	2.19	$2\bar{4}1$	11.0	11.9
	2.17	$1\bar{4}1$		
	2.16	221		
	2.15	131		
1.98	2.03	$3\bar{4}1$	9.7	10.4
	2.00	311		
	1.97	041		
	1.97	$4\bar{7}1$		
	1.92	$4\bar{5}1$	(5.0)	4.6
	1.90	$4\bar{1}1$	(5.0)	5.1
1.78	1.78	$2\bar{5}1$	5.3	3.3
	1.78	$4\bar{4}1$		
	1.76	231		
	1.76	401		
Second Level				
5.17	5.18	002	Moderate[c]	22.4
4.35	4.36	012	29.4	24.2
	4.21	$1\bar{2}2$		
	4.19	102		
3.39	3.39	$1\bar{7}2$	17.5	15.6
	3.37	112		
3.19	3.18	022	5.8	15.5
	3.17	$2\bar{1}2$		
2.97	2.96	$2\bar{2}2$	14.0	8.6
	2.94	202		
2.59	2.60	$1\bar{3}2$	12.1	15.0
	2.58	122		
	2.50	$2\bar{3}2$	(6.0)	7.1
	2.48	212	(6.0)	8.8
2.36	2.38	032	9.2	10.8
	2.35	$3\bar{2}2$		
	2.34	$3\bar{1}2$		
2.19	2.18	$3\bar{3}2$	6.9	5.8
	2.16	302		
2.04	2.06	$2\bar{4}2$	7.2	6.3
	2.04	$1\bar{4}2$		
	2.04	222		
	2.03	132		
	1.92	$3\bar{4}2$	(6.0)	6.2
	1.90	312	(6.0)	2.2
1.86	1.88	042	10.8	6.7
	1.87	$4\bar{2}2$		
	1.82	$4\bar{3}2$	(7.0)	4.5
	1.81	$4\bar{1}2$	(7.0)	5.9
1.68	1.71	$2\bar{5}2$	7.3	5.3
	1.70	$4\bar{4}2$		
	1.69	232		
	1.69	402		
Third Level				
	3.45	003		0.0
3.13	3.17	013	23.7	23.9
	3.13	$1\bar{1}3$		
	3.11	103		
2.73	2.74	$1\bar{7}3$	5.9	5.2
	2.72	113		
2.62	2.62	023	20.0	19.2
	2.62	$2\bar{1}3$		
2.48	2.49	$2\bar{7}3$	15.2	13.2
	2.48	203		
2.21	2.26	$1\bar{3}3$	20.5	18.2
	2.25	123		
	2.20	$2\bar{3}3$		
	2.18	213		
2.12	2.12	033	17.5	14.3
	2.10	$3\bar{7}3$		
	2.09	$3\bar{1}3$	(6.0)	4.1
1.98	1.97	$3\bar{3}3$	8.6	11.2
	1.96	303		
1.87	1.88	$2\bar{4}3$	6.8	8.9
	1.87	$1\bar{4}3$		
	1.86	223		
	1.86	133		
	1.78	$3\bar{4}3$	(4.0)	1.8
1.73	1.76	313	3.9	4.8
	1.74	043		
	1.73	$4\bar{7}3$		
	1.70	$4\bar{3}3$		
Fourth Level				
2.55	2.59	004	Strong[c]	72.9
2.46	2.46	014	8.1	10.8
	2.44	$1\bar{7}4$		
	2.43	104		
2.23	2.24	$1\bar{3}4$	19.4	16.7
	2.24	114		
	2.18	024		
	2.17	$2\bar{7}4$		
	2.10	$2\bar{2}4$	(7.5)	2.7
	2.10	204	(7.5)	2.3
	1.96	$1\bar{5}4$	(7.5)	2.0
	1.95	124	(7.5)	2.1
1.91	1.92	$2\bar{3}4$	11.9	7.1
	1.91	214		
	1.86	034	(9.0)	1.8
	1.85	$3\bar{7}4$	(9.0)	1.7
	1.84	$3\bar{1}4$	(9.0)	3.6
1.77	1.76	$3\bar{3}4$	9.4	5.9
	1.75	304		
Fifth Level				
	2.07	005		0.0
	2.01	015	(7.5)	2.6
	1.99	$1\bar{5}5$	(7.5)	6.7
	1.99	105	(7.5)	3.4
1.87	1.88	$1\bar{7}5$	7.5	7.0
	1.88	115		
	1.84	025		
	1.84	$2\bar{1}5$		
	1.80	$2\bar{7}5$	(7.5)	1.4
	1.79	205	(7.5)	2.6
1.70 / 1.65	1.71 / 1.70 / 1.68	$1\bar{3}5$ / $1\bar{7}5$ / $2\bar{3}5$	7.0	5.7
	1.67	215		
	1.64	035		

[a] Values enclosed in parentheses indicate the estimated threshold for observation of that hkl plane. If F(calc) was greater than the estimated threshold, the plane was included in the refinement with F(obs) assigned a value of 2/3 times the estimated threshold.

[b] Calculated structure amplitudes for final antiparallel model.

[c] The observed meridional intensities are recorded here as moderate or strong due to the difficulties in applying correction factors for these reflections.

els for cellulose II. In this method, stereochemical restrictions in the form of allowed bond distances and bond angles are imposed upon a molecular model so as to form a completely rigid body, except for possible side chain rotations. For cellulose, this description reduces the number of refinable parameters from 66, i.e., the atomic coordinates of the nonhydrogen atoms, to five, thus increasing the ratio of observations to refinable parameters to the point where a least-squares refinement is meaningful. A preliminary report of our conclusions, described in detail below, has already been published in this journal.[10]

Experimental Section

Samples of regenerated cellulose (rayon fibers) were obtained from Celanese Fibers Co. Small individual fibers were drawn from the larger rayon fibers and arranged in a parallel bundle. The x-ray diffraction patterns of these specimens were recorded on Kodak No-Screen film using Ni-filtered Cu Kα radiation in an evacuated camera, with a collimator consisting of 400 μm diameter pinholes separated by 12 cm. X-ray photographs were recorded with the fiber axis perpendicular to the beam for the intensity data and tilted to observe the meridional reflections. The d spacings were calibrated by dusting the specimen with CaF_2 powder.

Intensities were obtained for 44 observed (nonmeridional) reflections. Quantitative values for 34 of these reflections were obtained using a Photometrics EDP Scanning Microdensitometer. The optical density data were recorded on magnetic tape and processed by computer as a two-dimensional numerical map of the x-ray pattern. Optical density contours were drawn, and the integrated intensities were obtained by summing the optical density values within a given reflection. The background correction was then subtracted, based on the optical density around the edge of the reflection. The intensities of the remaining ten (weak) reflections were estimated visually. Finally, the intensity data were corrected for Lorentz and polarization effects.

In addition to the 44 observed reflections, there were 41 unobserved reflections that were predicted to fall within the scattering angle recorded by the x-ray photograph. For each least-squares cycle all these unobserved reflections were considered and those which have calculated structure amplitudes higher than estimated threshold values (F_{thres}) were assigned an F(obsd) of two-thirds F_{thres} and included in the data. The corrected intensity data, in the

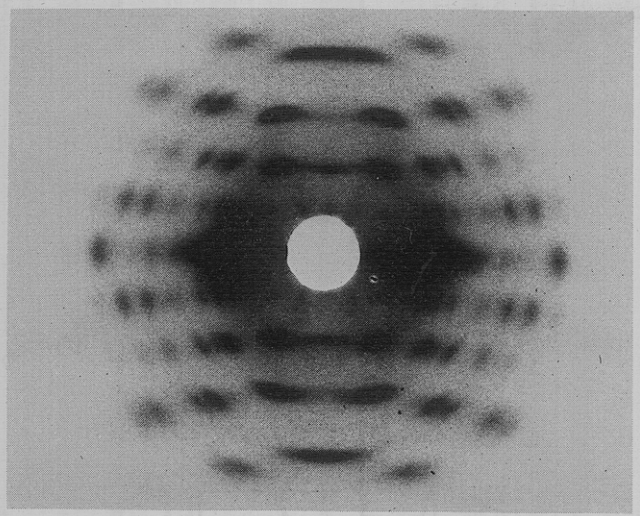

Figure 1. X-ray diffraction pattern of a bundle of oriented rayon fibers. The bundle has been tilted slightly to observe the 002 reflection.

form of structure amplitudes F(obsd) for the observed reflections and estimated threshold values (F_{thres}) are given in Table I.

Structure Determination. Unit Cell Parameters. The x-ray diffraction pattern of the rayon fiber specimen is shown in Figure 1. Least-squares refinement of the monoclinic unit cell parameters using the 44 observed d spacings resulted in values of $a = 8.01 \pm 0.05$ Å, $b = 9.04 \pm 0.05$ Å, $c = 10.36 \pm 0.05$ Å, and $\gamma = 117.1 \pm 0.1°$. The observed and calculated d spacings are given in Table I. Absence of odd-ordered $00l$ reflections indicates a $P2_1$ space group for this structure. (Legrand[11] also investigated the possibility of true odd-ordered meridionals in regenerated cellulose as distinct from streaking accompanying $(10l)$ or $(01l)$ reflections and found no evidence thereof.) The refined value of the helix repeat c is within experimental error of the 10.38 Å determined[8] for cellulose I and the backbone conformation of the cellulose molecule can be expected to be the same for cellulose I and II.

Molecular Model for the Isolated Chain. Consistent with the $P2_1$ symmetry, a model for the cellulose chain was constructed as a twofold helix repeating in 10.36 Å. Standard bond distances and bond angles[12] were used, and the molecular model incorporated an O3–H···O5′ intramolecular hydrogen bond of length 2.69 Å. The bond lengths and angles for the model are shown in Figure 2. The disaccharide residue has a glycosidic bond angle of 114.8° and glycosidic torsion angles of $\Phi = 24.7°$, $\psi = -26.2°$ (using the convention followed by Sundararajan and Rao[13]).

This molecular model is completely rigid except for the allowed rotation of the $-CH_2OH$ side group about the C5–C6 bond. This rotation is described by the dihedral angle χ, where χ has a value of zero when the C6–O6 bond is cis to the C4–C5 bond. Counterclockwise rotation of the group when looking down the C5–C6 bond represents positive rotation. The $-CH_2OH$ orientation is also described in terms of its orientation relative to the C4–C5 and C5–O5 bonds: gg, gauche to C5–O5 and gauche to C4–C5 ($\chi = -60°$); gt, gauche to C5–O5 and trans to C4–C5 ($\chi = 180°$); and tg, trans to C5–O5 and gauche to C4–C5 ($\chi = 60°$).[14]

Chain Packing. The positions of the rigid cellulose chains are completely defined by three packing parameters, whether they have the same or opposite sense. These parameters are the shift of one chain (along c) with respect to the other chain, and two parameters defining the orientation of the two chains (about their helix axes). A survey of the calculated intensities for potential packing arrangements of the two cellulose chains indicates that four basic models need to be considered, as was the case for the native structure.[8] The relative intensities of the 002 (moderately weak) and 004 (strong) indicate an approximate $c/4$ stagger of the chains. In each case the first chain through (0, 0, z) has the glycosidic oxygen O1′ at $z = 0$, and "shift" describes the c axis displacement of O1′ in the second chain through ($\frac{1}{2}$, $\frac{1}{2}$, z); the chain sense is defined as "up" when $Z_{O5} > Z_{C5}$. The four models are the following: p_1, parallel chains oriented up with a shift of $+c/4$ for the second chain; p_2, parallel chains oriented down with a shift of $+c/4$ for the second chain; a_1, antiparallel chains with an up chain at the origin and a down chain at ($\frac{1}{2}$, $\frac{1}{2}$, z), with a shift of $-c/4$; and a_2, antiparallel chains as in a_1, except with a shift of $+c/4$ for the second chain.

Thus, for each of the four possible models there are seven refinable parameters. Three packing parameters: (1) SHIFT, the stagger of the center chain along its helix axis with respect to that at the origin; (2) PHI, the rotation of the origin chain about its helix axis; (3) PHI′, the rotation of the center chain about its helix axis. Two molecular parameters: (4) the dihedral angle χ, which determines the orientation of the $-CH_2OH$ group in the origin chain; (5) the dihedral angle χ', which determines the orientation of the $-CH_2OH$ group in the center chain. Two crystallographic parameters: (6) K, the scale factor; and (7) B, the average isotropic temperature factor.

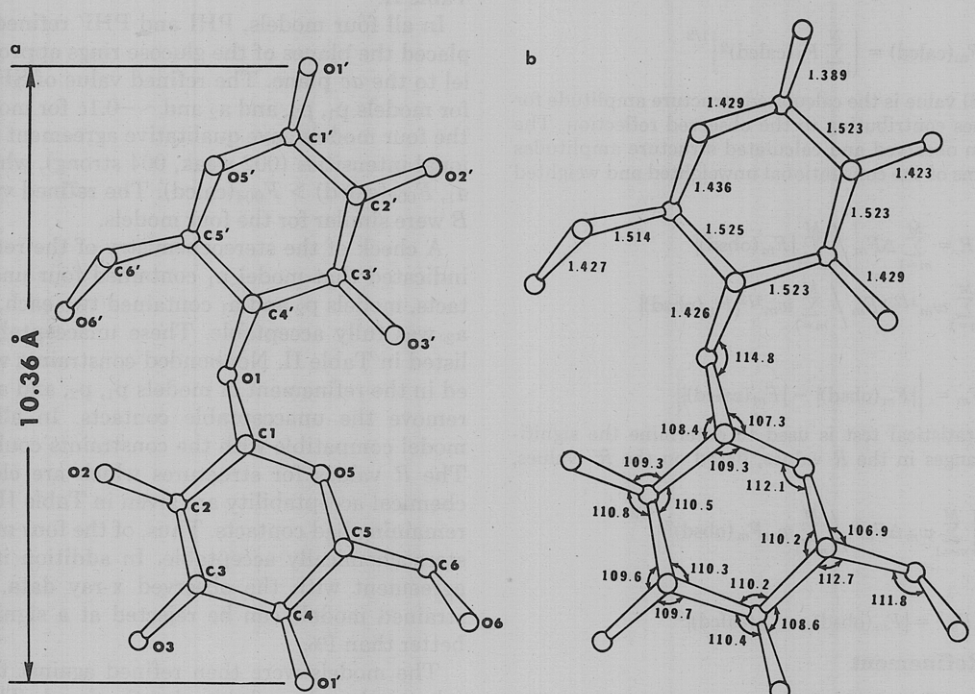

Figure 2. Molecular model for the cellulose chain. (a) Atom numbering in cellobiose repeating unit. (b) Bond lengths and angles in the model.

Table II
Refinement Results for the Four Possible Cellulose II Models

	p_1	p_2	a_1	a_2
	Observed Data Only			
R	0.254	0.188	0.195	0.171
R'	0.254	0.188	0.195	0.171
R''	0.201	0.154	0.170	0.134
Bad contacts[a]				
C···O,Å	2.63	2.49	1.98	None
O···O,Å	2.17	2.05	2.11	None
	2.47			
	2.43			
	Observed Data with Applied Constraints			
R	0.272	0.219	0.230	
R'	0.272	0.219	0.230	
R''	0.217	0.170	0.191	
Bad contacts,				
C···O,Å	2.52	2.50	2.97[b]	
O···O,Å	2.21	2.12	1.87	
	2.37			
	2.42			

[a] Unacceptable stereochemical contacts were taken as oxygen–oxygen distances less than 2.60 Å and oxygen–carbon distances less than 2.80 Å. [b] This nonbonded contact distance is acceptable; its value was included in the table to demonstrate the effect of the constraint.

Refinement. The possible structural models for cellulose were refined by adjusting the above parameters using a least-squares process[9] to provide the best fit between observed and calculated structure factors, minimizing the function

$$\Phi = \sum^{M} w_m |F_m(\text{obsd}) - F_m\ (\text{calcd})|^2$$

where F_m(calcd), defined by

$$F_m(\text{calcd}) = (1/K)F_m \exp(-B\rho_m^2/4)$$

is the calculated value to be compared with F_m(obsd), the mth of the M observed structure factor amplitudes. w_m is the weight to be applied to the observation; K is the scale factor; B is the average isotropic temperature factor; and ρ_m is the reciprocal d spacing. In cases where more than one hkl plane contributes to an observed reflection, the calculated structure amplitude used for comparison is given by

$$F_m(\text{calcd}) = \left\{ \sum_{n=1}^{N} F_n(\text{calcd})^2 \right\}^{1/2}$$

where the F_n(calcd) value is the calculated structure amplitude for the nth of N planes contributing to the observed reflection. The agreement between observed and calculated structure amplitudes is expressed in terms of the conventional unweighted and weighted R factors:

$$R = \sum_{m=1}^{M} \Delta F_m \Big/ \sum_{m=1}^{M} |F_m(\text{obsd})|$$

$$R' = \sum_{m=1}^{M} w_m{}^{1/2} \Delta F_m \Big/ \sum_{m=1}^{M} w_m{}^{1/2} |F_m(\text{obsd})|$$

where

$$\Delta F_m = \Big| |F_m(\text{obsd})| - |F_m(\text{calcd})| \Big|$$

The Hamilton[15] statistical test is used to determine the significance of small changes in the R values, based on the R'' values, where

$$R'' = \left\{ \sum_{m=1}^{M} w_m \Delta F_m{}^2 \Big/ \sum_{m=1}^{M} w_m F_m(\text{obsd})^2 \right\}^{1/2}$$

and

$$\Delta F_m{}^2 = |F_m(\text{obsd}) - F_m(\text{calcd})|^2$$

Course of the Refinement

Least-squares refinement of the seven parameters was performed for the two parallel and two antiparallel models. The models were first refined against the observed reflections only and the results of these refinements are shown in Table II.

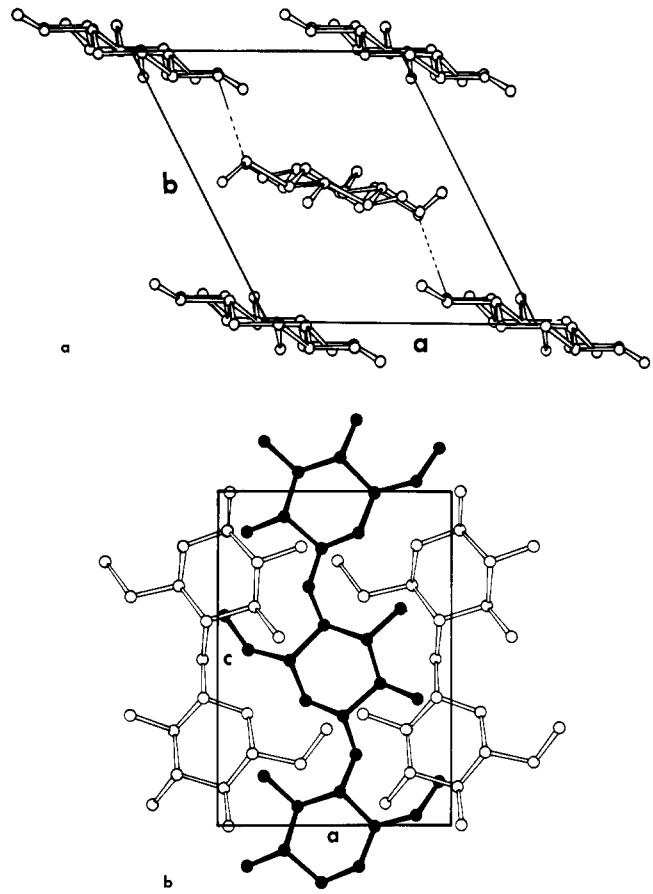

Figure 3. (a) Projection of the cellulose chains down the c axis of the unit cell; hydrogen bonding along the 110 plane is shown (only carbon and oxygen atoms are drawn in, hydrogen bonds are indicated by the paritially dashed lines; the hydrogen position given by the solid portion). (b) Projection of the chains perpendicular to the ac face of the unit cell.

In all four models, PHI and PHI′ refined to values that placed the planes of the glucose rings approximately parallel to the ac plane. The refined value of SHIFT was $\sim 0.2c$ for models p_1, p_2, and a_2 and $\sim -0.1c$ for model a_1. Three of the four models gave qualitative agreement with the meridional intensities (002 weak, 004 strong), whereas for model a_1, F_{002}(calcd) > F_{004}(calcd). The refined values for K and B were similar for the four models.

A check of the stereochemistry of the refined structures indicated that model p_1 contained four unacceptable contacts, models p_2 and a_1 contained two each, whereas model a_2 was fully acceptable. These unacceptable contacts are listed in Table II. Nonbonded constraints were incorporated in the refinement of models p_1, p_2, and a_1 in an effort to remove the unacceptable contacts. In all three cases a model compatible with the constraints could not be found. The R values for structures which are closest to stereochemical acceptability are given in Table II, along with the remaining bad contacts. Thus, of the four models, only a_2 is stereochemically acceptable. In addition it gives the best agreement with the observed x-ray data. All three constrained models can be rejected at a significance level of better than 1%.

The models were then refined against the full data set (observed and unobserved reflections). The bad contacts for models p_1, p_2, and a_1 were not removed and these models remain unacceptable. For model a_2, the resulting R

Table III
Fractional Atomic Coordinates for One Glucose Residue of Each Chain (Final Antiparallel Model[a])

Atom	x/a	y/b	z/c
		Origin Chain	
C1	−0.042	0.006	0.386
C2	−0.112	0.093	0.290
C3	−0.142	0.010	0.158
C4	0.034	0.000	0.115
C5	0.099	−0.080	0.220
C6	0.283	−0.081	0.188
O1′	−0.001	−0.097	0.000
O2	−0.281	0.088	0.336
O3	−0.188	0.103	0.066
O5	0.127	0.009	0.339
O6	0.351	−0.140	0.293
		Center Chain	
C1	0.467	0.518	−0.170
C2	0.316	0.500	−0.074
C3	0.402	0.566	0.058
C4	0.521	0.483	0.101
C5	0.661	0.498	−0.004
C6	0.773	0.406	0.028
O1′	0.621	0.559	0.216
O2	0.219	0.588	−0.120
O3	0.256	0.533	0.150
O5	0.565	0.430	−0.123
O6	0.902	0.484	0.131

[a] The complementary half of the unit cell can be generated by the symmetry operation $-x, -y, z + 0.5$.

values were $R = 0.213$, $R' = 0.200$, and $R'' = 0.155$. For R' and R'', weighting factors of $w = 1$ for the observed reflections and $w = \frac{1}{2}$ for unobserved reflections were used (the lack of systematic variations of ΔF_m with F(obsd) precluded the use of any other weighting schemes[16,17]). The new a_2 model did contain a bad O···O contact (2.49 Å) which was removed with an appropriate constraint. The R values of the stereochemically acceptable constrained model were: $R = 0.235$, $R' = 0.219$, and $R'' = 0.167$. For this model, only 9 of the 41 unobserved reflections had an F(calcd) > F_{thres}.

Thus an antiparallel model with an approximate quarter stagger of the chains is proposed for regenerated cellulose. The observed and calculated structure amplitudes for model a_2 are given in Table I. The *ab* and *ac* projections for the proposed antiparallel model are shown in Figure 3. The refined values of PHI = 22.8° (σ_{PHI} = 0.3°) and PHI′ = 61.8° ($\sigma_{PHI'}$ = 0.3°) orient the chains such that the "planes" of the glucose rings are approximately parallel to the *ac* plane. The refined value for SHIFT, the relative stagger of the chains, is 0.216*c* (σ_{SHIFT} = 0.018*c*). For the side groups, χ refines to 186.3° (σ_χ = 12.0°), placing the $-CH_2OH$ group of the origin chain within ~6° of the gt position. The value of χ' is 70.2° ($\sigma_{\chi'}$ = 15.2°), which positions the $-CH_2OH$ group of the center chain within ~10° of the tg position. The refined isotropic temperature factor is 19.96 (σ_B = 2.09). The fractional coordinates for the final model are given in Table III.

Hydrogen Bonding Network

The proposed hydrogen bonding network for regenerated cellulose is shown in Figure 4. With the refined values of χ and χ' all of the hydroxyl groups can form hydrogen bonds with acceptable O–H···O bond distances and C–O···O bond angles. Infrared spectroscopy[18,19] confirms that all the hydroxyl groups are hydrogen bonded.

Each chain possesses the O3–H···O5′ hydrogen bond of length 2.69 Å as defined in the model. The $-CH_2OH$ groups of the "down" (center) chains are close to the tg position such that these chains possess a second intramolecular hydrogen bond: O2′–H···O6. The –O6–H groups then form an intermolecular O6–H···O3 hydrogen bond to the next chain along the *a* axis, i.e., in the 020 plane. Thus the down chains form hydrogen bonded sheets, as shown in Figure 4a, which are essentially the same as those in cellulose I.

For the "up" (corner) chains the $-CH_2OH$ groups are close to the gt position,[21] and form intermolecular O6–H··· O2 hydrogen bonds to the next chain along the *a* axis. The sheet of up chains is shown in Figure 4b. The O2–H group cannot form an intramolecular hydrogen bond but is involved in an intermolecular O2–H···O2′ hydrogen bond to the neighboring down chain in the 110 plane, as shown in Figure 4c. The bond lengths and angles for the hydrogen bonds are given in Table IV.

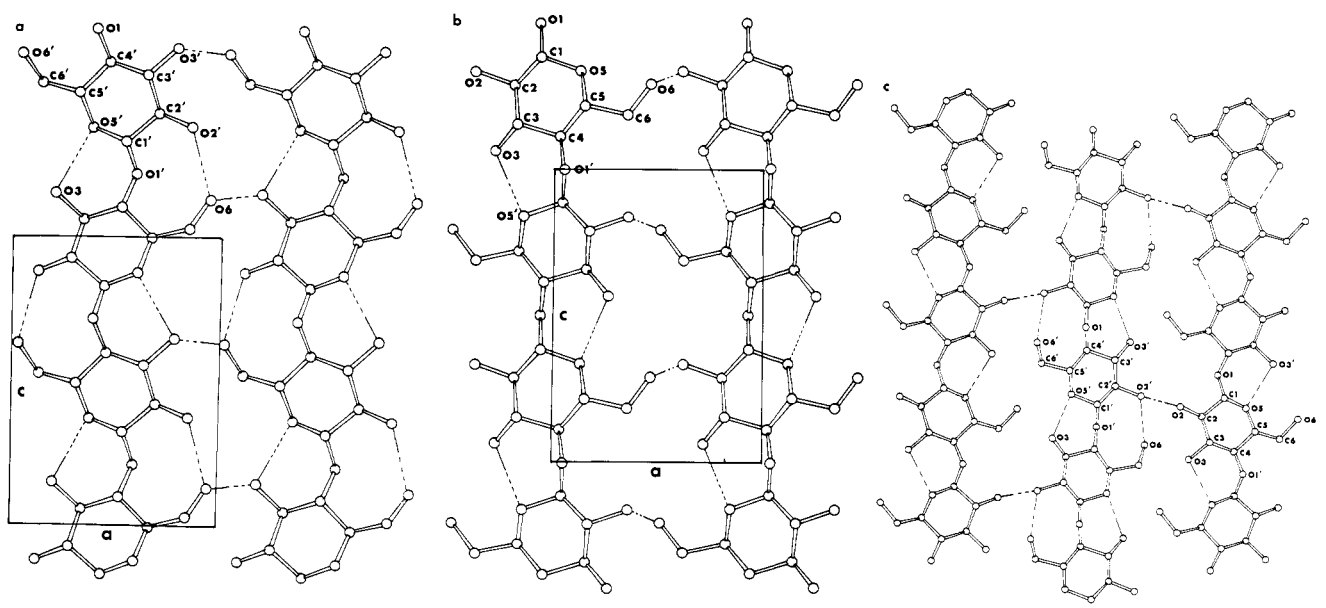

Figure 4. Intrachain and interchain hydrogen bonding in regenerated cellulose: (a) along the 020 plane for the "down" chains, (b) along the 020 plane for the "up" chains, (c) along the 110 plane (view from the origin of the unit cell).

Table IV
A Summary of Hydrogen Bonding Information for the Final Antiparallel Model[a]

Bond lengths, Å		Bond angles, deg	
O3–O5′	2.69	C3–O3···O5′	102.6
O2*′–O6*	2.73	C2*′–O2*′···O6*	114.3
O6–O2$_a$	2.76	C6–O6···O2$_a$	107.7
O6*–O3*$_a$	2.67	C6*–O6*···O3*$_a$	127.9
O2$_a$–O2*′	2.77	C2$_a$–O2$_a$···O2*′	108.5

[a] The symbols used are the following: (*) denotes atom on the center "down" chain; (′) denotes an atom on the next residue up from the asymmetric residue; and (a) denotes an atom on the next chain along the positive *a* axis.

Discussion

An antiparallel chain structure is the only acceptable model for regenerated cellulose II, in contrast to the parallel chain system for cellulose I. The regenerated and native structures are similar in many ways, each consisting of an array of hydrogen bonded sheets of chains parallel to the 020 planes. However, it is not possible stereochemically to pack antiparallel sheets in the cellulose II lattice when all the $-CH_2OH$ side chains have the same (tg) conformation as in cellulose I. Rather the tg conformation is retained by alternate (down) sheets while the other (up) sheets have the gt conformation. This still allows for intermolecular intrasheet hydrogen bonding of the O6–H groups and introduces the attractive feature of an O2–H···O2′ intersheet hydrogen bond. Cellulose II is the stable polymorphic form; cellulose I is produced only in nature, and is converted into form II by solution and reprecipitation or by swelling, after which it cannot be reconverted to form I. The hydrogen bonding between center and corner chains may contribute significantly to the stability of form II.

Of the various possibilities for the chain polarities in the two polymorphic forms, the parallel I–antiparallel II solution seems the most reasonable. If cellulose I were to consist of antiparallel chains it is difficult to see why these would not adopt the cellulose II lattice. Parallel chains however lead to production of extended chain polymer single crystals, which are particularly suitable as structural fibers. The chains in form I are bound together sufficiently tightly that the crystallites are inpenetrable to water. However, cellulose solvents break these bonds whereupon the chains would be expected to act like other polymers and fold back on themselves when they are precipitated.

In the Mercerization process, cellulose I is converted to cellulose II by swelling in caustic soda solution. This change is accompanied by only a small decrease in length, which has been used as an argument against chain folding. We are currently examining the x-ray patterns of Mercerized cellulose to see if the same antiparallel structure is present. Nevertheless, Chanzy et al.[20] have recently examined shish kebab structures given by low molecular weight cellulose (DP ~ 30) epitaxially crystallized on cellulose I fibrils; the low molecular weight material adopts the cellulose II lattice. Such epitaxial crystallization of cellulose II on cellulose I is to be expected since half of the sheets in form II are common to form I. Mercerization proceeds slowly and never goes to completion. The unconverted cellulose I could maintain the fiber dimensions and serve as a template for crystallization of the cellulose II chain folded domains. It should be emphasized, however, that although our results indicating alternating chains are compatible with regular chain folding, they do not prove that such folding does in fact occur.

Acknowledgment. This research was supported by N.S.F. Grant No. DMS 75-01028 and N.I.H. Research Career Development Award No. AM 70642 (to J.B.). A preliminary study of the cellulose II structure by K. H. Gardner provided valuable assistance in this work. (Ph.D. dissertation, Case Western Reserve University, 1974.)

Miniprint Material Available: Full-sized photocopies of Table I (8 pages). Ordering information is given on any current masthead page.

References and Notes

(1) N. J. Wellard, *J. Polym. Sci.*, **13**, 471–476 (1954).
(2) D. W. Jones, *J. Polym. Sci.*, **32**, 371–394 (1958).
(3) D. W. Jones, *J. Polym. Sci.*, **42**, 173–188 (1960).
(4) J. Mann, L. Roldan-Gonzales, and J. J. Wellard, *J. Polym. Sci.*, **42**, 165–171 (1960).
(5) S. Watanabe and J. Hayashi, *Kogyo Kagaku Zasshi*, **73**, 1890–1891 (1970).
(6) A. Sarko and R. Muggli, *Macromolecules*, **7**, 486–494 (1974).
(7) K. H. Gardner and J. Blackwell, *Biochim. Biophys. Acta*, **343**, 232–237 (1974).
(8) K. H. Gardner and J. Blackwell, *Biopolymers*, **13**, 1975–2001 (1974).
(9) S. Arnott and A. J. Wonacott, *Polymer*, **7**, 157–166 (1966).
(10) F. J. Kolpak and J. Blackwell, *Macromolecules*, **8**, 563–564 (1975).
(11) Ch. Legrand, *Acta Crystallogr.*, **5**, 800–802 (1952).
(12) S. Arnott and W. E. Scott, *J. Chem. Soc., Perkin Trans.* **2**, 324–335 (1972).
(13) P. R. Sundararajan and V. S. R. Rao, *Biopolymers*, **8**, 305–312 (1969).
(14) M. Sundaraligam, *Biopolymers*, **6**, 189–213 (1966).
(15) W. C. Hamilton, *Acta Crystallogr.*, **18**, 502–510 (1951).
(16) E. W. Hughes, *J. Am. Chem. Soc.*, **63**, 1737–1752 (1941).
(17) D. W. J. Cruickshank and D. E. Pilling, "Computing Methods and Phase Problem in X-ray Crystal Analysis", Pergamon, Oxford, England, 1961, p 46.
(18) H. J. Marrinan and J. Mann, *J. Polym. Sci.*, **21**, 301–311 (1956).
(19) J. Mann and H. J. Marrinan, *J. Polym. Sci.*, **32**, 357–370 (1958).
(20) H. D. Chanzy and E. J. Rochs, *J. Polym. Sci., Polym. Symp.*, in press.
(21) In our preliminary communication, the side-chain conformations were incorrectly given as tg for the "up" chain and gt for the "down" chain. These conformations should be gt for "up" and tg for "down" as described here.