Preferred citation: R.H.Atalla. Studies of polymorphy in native cellulose. In **Papermaking Raw Materials**, *Trans. of the VIIIth Fund. Res. Symp. Oxford, 1985*, (V.Punton, ed.), pp 59–75, FRC, Manchester, 2018. DOI: 10.15376/frc.1985.1.59.

STUDIES OF POLYMORPHY IN NATIVE CELLULOSE

Rajai H. Atalla
The Institute of Paper Chemistry, Appleton, Wisconsin 54912, USA

ABSTRACT

Our studies of cellulose structure based on X-ray diffractometry, Raman spectroscopy, and Solid State ^{13}C-NMR have led us to a model which addresses questions of structure at two levels. The first is that of the organization of individual chains. Two stable ordered states of cellulose chains are postulated, together with a disordered state in which there is less coherence between the orientations of adjacent anhydroglucose rings. The ordered states are identified as k_I and k_{II} based on their predominance in celluloses I and II, respectively; both conformations are based on the dimeric anhydrocellobiose as the basic repeat unit in the ordered chain. The disordered state is identified as k_o.

The second level of organization is that of aggregation of chains into three-dimensionally ordered crystalline domains. At this level our model recognizes two crystalline forms within the native state. These are identified as I_α and I_β, the first found to be dominant in bacterial and algal celluloses, the second dominant in celluloses from higher plants. These two forms are found to contain chains possessing the same molecular conformation k_I, but the patterns of hydrogen bonding are found to be different. Cellulose II, which is derived from the native state by mercerization or regeneration at low temperatures, is found to consist predominantly of chains in the k_{II} conformation in yet a third distinct crystalline lattice.

INTRODUCTION

Our earlier studies led us to propose a model of cellulose structure wherein the basic repeat unit of physical structure

is taken to be an anhydrocellobiose unit with nonequivalent adjacent anhydroglucose units. In order to reconcile information from Raman spectroscopy, conformational energy analyses, and ^{13}C-NMR studies, we also proposed that the nature of the nonequivalences between the adjacent anhydroglucose units differed between celluloses I and II. Thus, we view the two most common polymorphs as possessing different molecular conformations as well as different unit cells in their crystalline domains.

Later high resolution ^{13}C-NMR studies on crystalline native celluloses led us to the conclusions that, in their native states, celluloses are composites of two distinct crystalline modifications, and that the relative proportions of the two modifications are different in different native forms. This level of polymorphy within the native state was not previously recognized. Our most recent efforts have been directed at understanding the differences between these two crystalline modifications, and the degree to which their relative proportions in native celluloses influence their properties.

In this report we will review our model of structure, with particular attention to the native state. We will then discuss our most recent findings concerning native structures.

STRUCTURAL MODELS

Our proposals concerning the structure of cellulose have departed from earlier models at two levels. The first is that of individual cellulose molecules within an ordered lattice. The second is that of aggregates of chains in three dimensionally ordered crystal lattices. The following discussion will address the questions arising at each of these levels in sequence.

Molecular Conformations

The model that we have proposed was developed in order to reconcile the results of Raman spectroscopic observations with information from prior structural studies. From a detailed comparison of the Raman spectra of celluloses I and II, we concluded that the differences in the spectra could not be interpreted except in terms of different conformations of the molecular chains (1). Many of the structural models then

current had interpreted the two polymorphs as the result of different lattice packings of chains with the same skeletal conformation.

In order to establish the nature of the differences between the two conformations, we considered information from a number of sources. First, we examined the results of conformational energy calculations (2,3). They indicate that two different states of the glycosidic linkage are stable. These can be represented as relatively small left-handed and right-handed departures from the twofold helix conformation. They are also well approximated, respectively, by the experimentally observed conformations of the model disaccharides cellobiose (4) and methyl-β-cellobioside (5).

An analysis of the vibrational spectra in the OH stretching region for both the model disaccharides and for celluloses I and II led us to suggest that nonequivalent glycosidic linkages alternate along the molecular chains (6). The multiplicities of chemically equivalent carbons in the solid state ^{13}C-NMR spectra confirmed this conclusion, and the chemical shifts were consistent with a difference between the skeletal conformations in celluloses I and II (vide infra). Finally, the Raman spectra indicated that the C-6 carbons occur in nonequivalent sets in cellulose I but are merged into a single set in cellulose II (7).

With the above observations in mind, we formulated the new model of structure to meet three requirements. The first is that it allow for different ordered conformations of the molecular chains in celluloses I and II. The second is that adjacent anhydroglucose units in any cellulose chain be nonequivalent. That is, that the basic repeat unit in crystalline celluloses must be considered to be the dimeric anhydrocellobiose unit. The third, and perhaps most subtle requirement, is that the differences resulting in nonequivalence are centered at different points in the anhydroglucose rings in celluloses I and II.

The last of the requirements, concerning the sites of the nonequivalence is a key one. In cellulose I the C-6 hydroxyls on adjacent anhydroglucose rings are the centers of nonequivalence. The glycosidic linkages are also nonequivalent, but the ^{13}C-NMR spectra show them to be less distinct than in

cellulose II. In cellulose II, on the other hand, the nonequivalence is centered at the glycosidic linkages, whereas the C-6's are shown by the Raman spectrum to be essentially equivalent for adjacent anhydroglucose rings.

When, in addition to the requirements outlined above, we took into consideration the constraints imposed by packing of the chains in the lattice, the picture which emerged was one incorporating features of both experimental disaccharide structures. Thus, the model is of a cellulose chain with alternating glycosidic linkages corresponding alternately to small left-handed and right-handed departures from the twofold helix conformation. This feature is taken as common to both polymorphs.

The basic difference between molecular conformations in the two polymorphs is related to the nonequivalence of C-6's in cellulose I, and is illustrated in Fig. 1. On the basis of the vibrational spectra in the OH stretching region, we concluded that alternate C-6's in cellulose I have their primary hydroxyl groups participating in bifurcated intramolecular hydrogen bonds similar to the one known to occur in crystalline methyl-β-cellobioside. In this model disaccharide the dihedral angles at the glycosidic linkage have values which correspond to a relatively small right-handed departure from the twofold helix for cellulose. In this conformation the positions of the ring oxygen and the primary hydroxyl oxygen on C-6 are such that they can both function as electron donors in the hydrogen bond with the hydroxyl group on C-3 of the adjacent ring. The conformation of cellobiose, which is equivalent to a small left-handed departure from the twofold helix, does not allow the oxygen on C-6 to participate in the bifurcated bond, and the intramolecular hydrogen bond remains an isolated one. We believe that the participation of alternate primary hydroxyl groups in bifurcated intramolecular hydrogen bonds is the primary factor stabilizing the glycosidic linkages of cellulose I in positions where their nonequivalence is less pronounced than is the case in cellulose II.

In terms of the new model, the mechanism of conversion of cellulose from polymorph I to polymorph II is relatively simple. It corresponds to removal of the primary hydroxyl on every other C-6 from the bifurcated intramolecular hydrogen bonds. This then allows the glycosidic linkages to relax into

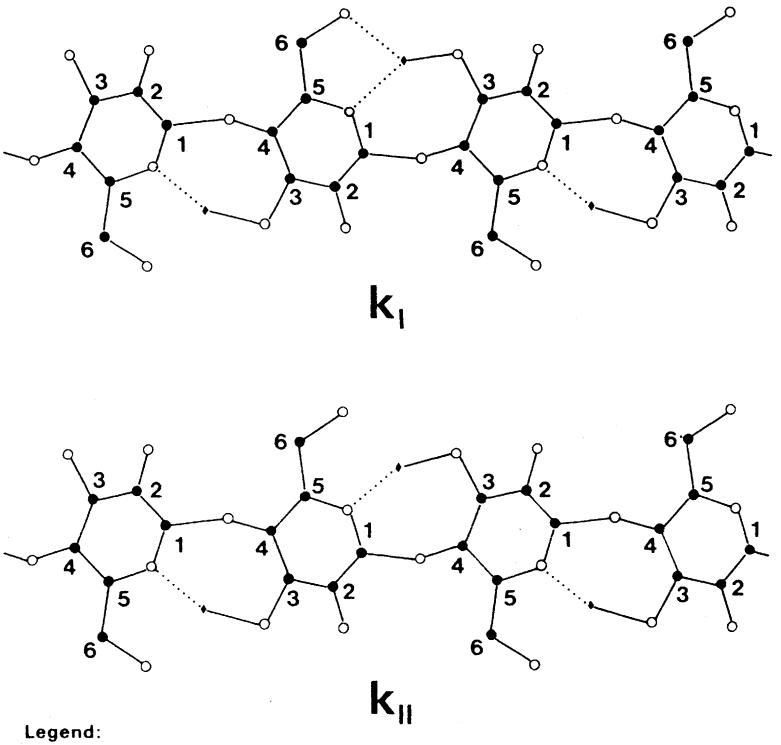

Fig 1—Schematic representation of conformations k_1 and k_{11}

the more distinct positions characteristic of cellulose II, and permits the C-6's to move into approximately equivalent positions. The relatively small contraction of the unit cell in the chain direction, when cellulose I is converted to cellulose II, is a logical consequence of the relaxation of the glycosidic linkages.

Conformational Distributions in Celluloses

Our Raman spectral studies of a variety of samples have led us to the conclusion that most celluloses of practical interest contain molecules in one or both of the stable ordered states described above, together with less ordered molecular chains. We have, therefore, used our model as a basis for deriving the distributions of molecular conformations in samples of practical interest from their Raman spectra. To accomplish this, three conformational states were defined. The two ordered states were identified as k_I and k_{II} on the basis of their predominance in celluloses I and II, respectively. In addition, a disordered state, identified as k_o has been defined as corresponding to the molecular organization dominant in the most amorphous celluloses.

In terms of the model outlined above, k_I corresponds to an ordered sequence of anhydrocellobiose units, with every other C-6 involved in a bifurcated intramolecular hydrogen bond, as shown in Fig. 1. Thus succesive C-6's in k_I are nonequivalent. Conformation k_{II}, also depicted in Fig. 1, has, in contrast, only isolated intramolecular hydrogen bonds, and the C-6's are approximately equivalent. The disordered state, k_o, differs from k_I and k_{II} in that no coherently ordered sequences of anhydrocellobiose units are expected. It is not expected to be completely random, however, since the hard sphere interactions between atoms on adjacent anhydroglucose units confine the dihedral angles on the glycosidic linkages to a much narrower range than would otherwise be expected.

The model has permitted quantitative resolution of the Raman spectra of celluloses to establish the distributions of conformational states in different samples (8). It is important to stress in this connection, that in many studies of model saccharides the Raman spectra could be well interpreted on the basis of the isolated molecule approximation. Thus, the

spectra are determined primarily by skeletal conformations; effects of lattice packing are secondary.

Our procedure permits quantitative characterization of the structure of native and processed celluloses, most of which are not pure polymorphic forms (9). The characterization reflects the distribution of cellulose among the conformational states, and does not always reflect the pattern of molecular aggregation into crystalline domains. It is true, however, that the crystalline polymorphs I and II are dominated by the conformations k_I and k_{II}, respectively.

Crystalline Forms

In our work with VanderHart (10,11) a number of crystalline celluloses were investigated using Solid State ^{13}C-NMR spectroscopy based on the method of Cross Polarization and Magic Angle Spinning (CP/MAS). The spectra which are central to development of the model were recorded from four native celluloses, and from regenerated celluloses I and II prepared by procedures developed in our laboratory to produce the pure polymorphic forms (12,13). The native celluloses studied included a bacterial cellulose from Acetobacter xylinum, an algal cellulose from Valonia ventricosa, and two fibrous celluloses, namely, cotton linters and ramie.

The spectra of the pure cellulose II could be rationalized in terms of nonequivalent sites within a unique unit cell. The spectra of the native celluloses, on the other hand, revealed multiplicities that suggested a more complex architecture.

The spectra of the native celluloses are shown together with the spectrum of the regenerated cellulose I in Fig. 2. The accepted assignments of resonances to the different carbons in the anhydroglucose units are as indicated at the top of the figure. The cluster between 70 and 81 ppm are assigned to C-2, C-3, and C-5, collectively, because at present there is no basis for individual assignment. The assignments of resonances due to C-1, C-4, and C-6, however, are firmly established on the basis of comparisons with oligosaccharides and model compounds.

The sharper components of the C-4 and C-6 resonances, as well as the C-1 resonance possess a multiplicity indicating

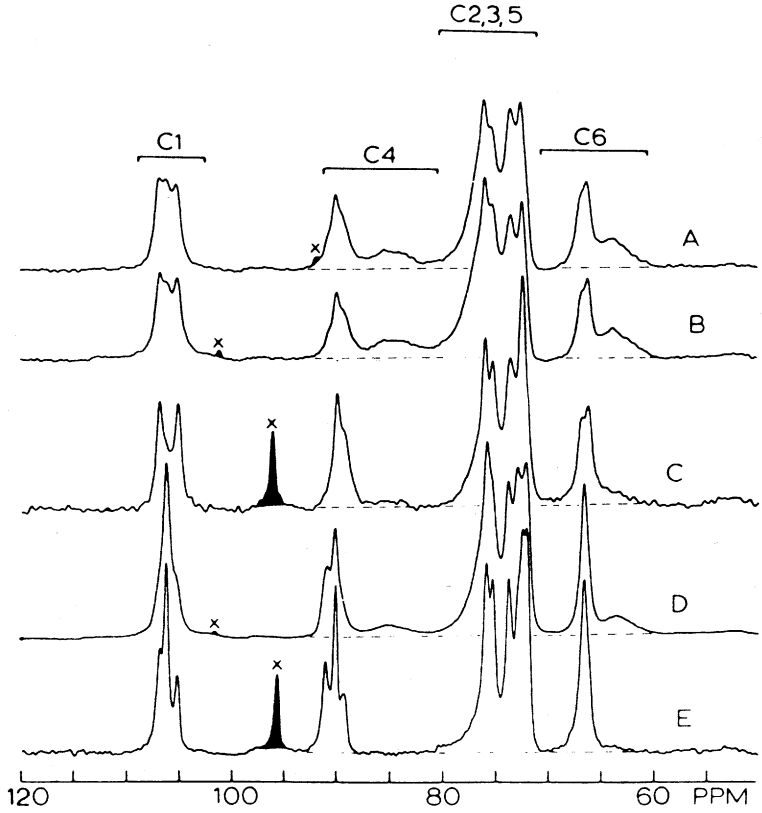

Fig 2— ^{13}C CP-MAS spectra of various celluloses: A - Ramie; B - cotton linters; C - regenerated cellulose I; D - **Acetobacter xylinum** cellulose; E - **Valonia ventricosa** cellulose. The "X" marks the small first spinning side band of lindear polyethylene added as an internal standard; its centerband at 33.6 ppm is not included in this display

that they arise from magnetically nonequivalent sites within crystalline domains. The most important feature in Fig. 2 is the pattern of variation of the multiplets, primarily at C-4, but also at C-1 and C-6. These narrow lines, which arise from molecular chains within crystalline domains, differ among the samples; the relative intensities are not constant, and they are not in the ratios of small whole numbers as would be expected if they arise from nonequivalent sites within a single unit cell. We concluded, therefore, that the native celluloses must be composites of two or more crystalline forms. But the spectral intensities are not consistent with the possibility of three independent crystal forms, each giving rise to a single resonance in the C-1, C-4, and C-6 regions. A model based on two independent crystalline forms remained the most plausible alternate proposal. And indeed a decomposition of the spectra was accomplished on the basis of such a model.

The approximate component spectra of the two forms are illustrated in Fig. 3, and designated I_α and I_β. The spectra were obtained by taking appropriate linear combinations of the spectrum of the regenerated cellulose I and that of the cellulose from Acetobacter xylinum. They were judged to be the closest to the two extremes on the basis of a two component model. Comparison of the I_α and I_β spectra with the spectrum of cellulose II, which is included in Fig. 3, clearly shows that the multiplicity of I forms is quite distinct from the well established polymorphic variation.

The spectra in Fig. 2 indicate the following ranking with respect to the relative amounts of I_α as a fraction of the total: Acetobacter xylinum $\cong$ Valonia ventricosa $>$ cotton $\cong$ ramie $>$ regenerated cellulose I. The spectra in Fig. 2 and those recorded for many other native celluloses have led us to propose that all native celluloses are composites of the two distinct forms. We believe that our finding confirms earlier reports (14) that Acetobacter and Valonia celluloses are structurally different from other celluloses such as cotton and ramie. A very crude estimate based on our model suggests that Acetobacter cellulose is 60 to 70% I_α, whereas cotton is 60 to 70% I_β.

CONFORMATIONAL VARIATION AND POLYMORPHY

Analyses of the SS ^{13}C-NMR spectra of the different native forms led us to the conclusion that two distinct crystalline

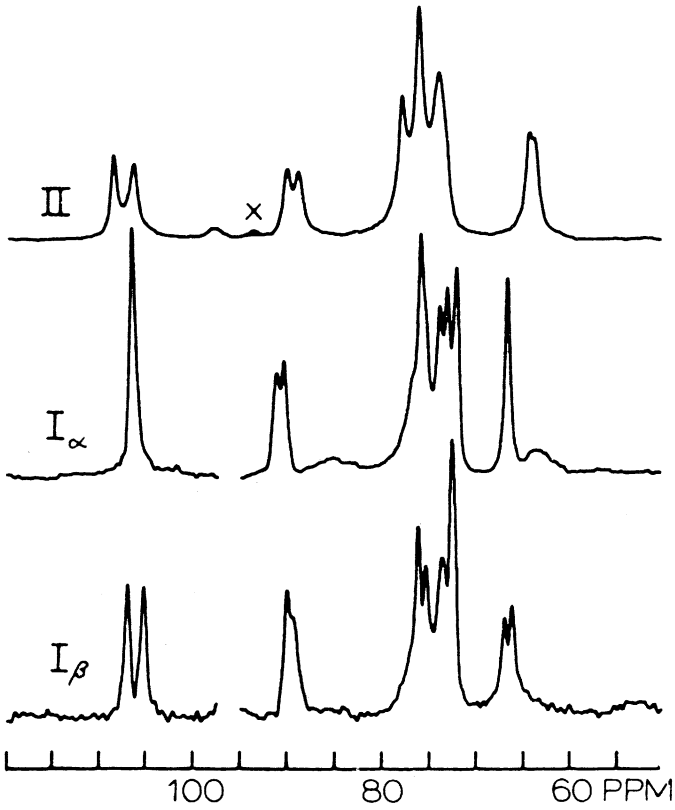

Fig 3—Comparison of the ^{13}C CP-MAS spectra of cellulose II and the spectra of the two proposed crystalline forms of cellulose I, namely, 1_a, and I_β. An "X" or a gap mark locations of the first spinning sideband of the linear polyethylene chemical shift standard.

forms coexist in native states of cellulose. In essence, the spectra indicate that polymorphy occurs within the cellulose I subset of structures, in addition to, and clearly apart from the well recognized cellulose II form. The Raman spectral studies, on the other hand, led us to propose that two distinct ordered states of the cellulose chains are stable, and that these can coexist together with a less ordered state in most celluloses of practical interest. The question thus arose whether the two forms I_α and I_β represent true solid state polymorphs with molecules possessing identical conformations, or whether like celluloses I and II they contain molecules with different conformations.

In light of our earlier extensive analysis of the effect of conformational changes on vibrational spectra, it seemed to us that a comparison of the Raman spectra of the different native forms would resolve this question. Most of our earlier studies had focused on celluloses from higher plants where I_β is the dominant component. We, therefore, extended our investigations to include the Raman spectra of bacterial and Valonia celluloses.

The bacterial cellulose was a sample from Acetobacter xylinum previously grown at the Institute in the course of other studies. The Valonia cellulose was isolated from a culture of Valonia ventricosa currently maintained in our laboratory.

The Raman spectra of the native celluloses are shown, together with a spectrum of cellulose II, in Fig. 4. The region of the spectra shown is the one between 250 and 1500 cm^{-1} which is generally dominated by the skeletal motions most sensitive to molecular conformation. Comparison of the spectra reveals great similarity between most of the major features in the spectra of the three native celluloses. The spectrum of cellulose II is in sharp contrast, however; most of its features, particularly in the region below 600 cm^{-1}, are quite different from those of the native celluloses.

On the basis of our prior analysis of the dependence of vibrational spectra on molecular conformation, we concluded that the spectra in Fig. 4 indicate the same conformation of the cellulose chain in the three native celluloses, and a distinctly different conformation in cellulose II. In terms of

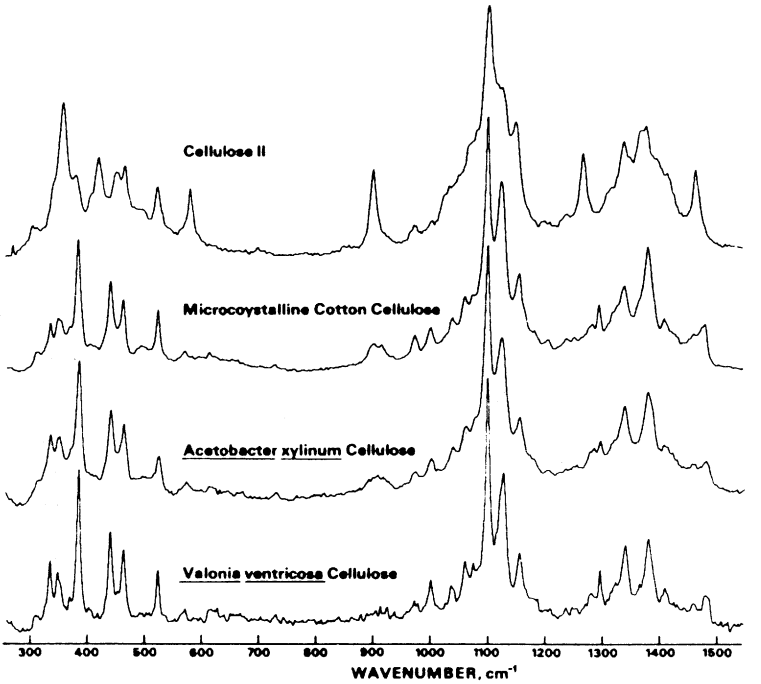

Fig 4—Raman spectra of **Valonia ventricosa** cellulose, **Acetobacter xylinum** cellulose, Microcystalline cotton cellulose, and high-crystallinity cellulose I.

the conformational model, k_I is dominant in native celluloses, and k_{II} in cellulose II.

When these results were considered together with the ^{13}C-NMR results shown in Fig. 3, we concluded further that the molecular conformation k_I is capable of aggregating into more than one stable crystalline lattice. Thus, I_α and I_β are two different crystalline lattices containing cellulose in the same k_I conformation. Cellulose II, in contrast, is an aggregate of molecules in the conformation k_{II} in yet a third distinct lattice.

A pattern similar to that noted above has been observed among the polymorphs of amylose (15). We have previously commented (1) on the Raman spectra of forms V_a and V_h which differ in their hydration but have the same molecular conformation and nearly identical spectra. In contrast, form B, which is known to have a distinctly different helix period, has a spectrum which differs from those of forms V_a and V_h.

Another aspect of the ^{13}C-NMR spectra, the chemical shift variation, is also consistent with the proposal of the same conformation, k_I, for celluloses I_α and I_β, and a different one, k_{II}, for cellulose II. Comparison of the spectra in Fig. 3 shows that, on average, the chemical shifts in cellulose II differ from those of celluloses I_α and I_β, more than the shifts of the latter two differ among themselves.

In search of further information concerning the difference between I_α and I_β forms of native cellulose, we explored the Raman spectra in the OH stretching region. The possibility that this particular region of the spectra may be informative was suggested by the work of Marrinan and Mann (14) wherein the differences between the different classes of native celluloses had been established on the basis of their infrared spectra in the OH region. In our experiment, Raman spectra were recorded for parallel aggregates of elementary fibrils from the cell walls of Valonia and for individual ramie fibers. The spectra were acquired with a Raman microprobe, and were recorded with the electric vectors of the exciting radiation parallel and perpendicular to the direction of the fibrils.

The spectra in the CH and OH stretching regions are shown in Fig. 5 for both ramie and Valonia ventricosa fibrils. The

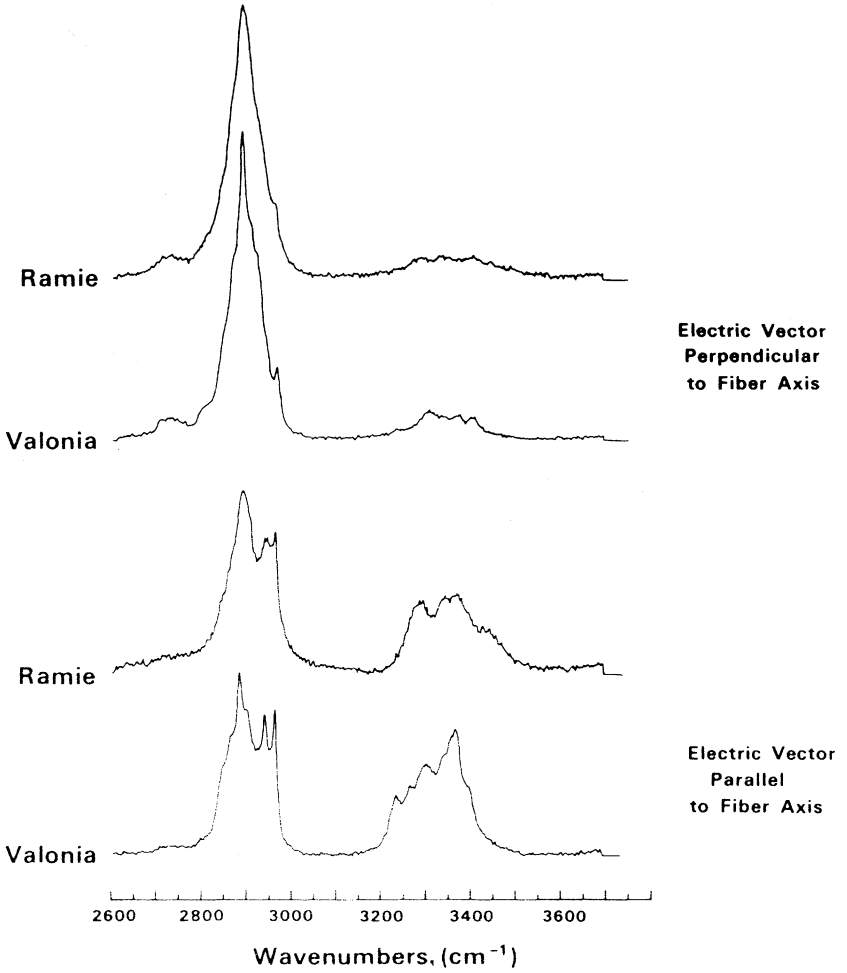

Fig 5—Raman microprobe spectra in the CH and OH stretching regions for ramie and **Valonia Vcentricosa** celluloses. The spectra were recorded for fibrillar aggregates of **Valonia** cellulose and for on individual fiber of ramie.

spectra are best resolved in the parallel mode for both types of native cellulose. The OH bands are also more intense in this mode, suggesting that a larger fraction of the hydroxyl groups are aligned parallel to the chain direction rather than perpendicular to it.

The outstanding difference between the ramie and Valonia celluloses is that the latter seem to have one additional OH band at lower frequency than any in the ramie spectrum. The ramie spectrum, on the other hand, shows one somewhat broader band at higher frequencies than any in the Valonia spectrum. Thus each of the two forms of cellulose seems to have correlated with it at least one hydrogen bond that does not occur in the other form. These observations, when taken together with the considerable similarity of the Raman spectra in the skeletal region, lead us to the conclusion that the key difference between the I_α and I_β forms of cellulose must be related to differences in the intermolecular hydrogen bonding patterns in the respective lattices. Such differences could well account for the differences in the SS ^{13}C-NMR spectra of the two forms.

The Raman spectra in the OH stretching region also provide important evidence concerning the much lower susceptibility of algal celluloses to mercerization. The OH band at the lowest frequency suggests the presence of a hydrogen bond that is stronger than any in ramie cellulose. The band at the high end of the frequency range in ramie indicates the presence of some intermolecular bonds which are weaker than any in the Valonia cellulose. Thus, on balance it is not surprising that the algal celluloses are less susceptible to mercerization than the ramie and related celluloses.

CONCLUSIONS

The central features of our model of cellulose are that anhydrocellobiose is the basic repeat unit of structure in cellulosic chains, and that it can occur in ordered sequences defining two distinct conformations, k_I and k_{II}, as well as incoherent sequences corresponding to a disordered state, k_o. Furthermore, the ordered sequences can aggregate into different crystalline forms, some incorporating the same molecular conformation, others incorporating different conformations. The key departure from prior models developed in our work, is

recognition of the native states of cellulose as composites of two distinct crystalline forms, which nevertheless are aggregates of chains in the same molecular conformation. The primary difference between these forms, revealed in the SS ^{13}C-NMR spectra and in the OH stretching regions of the Raman spectra, appears to arise from differences in the patterns of hydrogen bonding. We believe that a deeper understanding of the nature of these two components of native cellulose will lead us to clarification of the differences between the properties of celluloses from different sources, as well as the patterns of their response to process conditions.

ACKNOWLEDGEMENTS

Support of this work through the Institutional Research Fund of The Institute of Paper Chemistry is gratefully acknowledged. The collaboration of Dr. David L. VanderHart in the ^{13}C-NMR studies has been invaluable. Ms. Rebecca E. Whitmore and Mr. Clark P. Woitkovich, the author's staff associates at the Institute, have contributed much to the experimental programs undertaken in support of the theses presented in this report. The spectra in Fig. 5 were acquired by Mr. James H. Wiley in the course of his doctoral research program.

REFERENCES

1. Atalla, R. H. Proceedings of the 8th Cellulose Conference, Appl. Polymer Symp. 28:659(1976).

2. Reese, D. A., and Skerrett, R. J., Carbohyd. Res. 7:334 (1968).

3. Melberg, S., and Rasmussen, K., Carbohyd. Res. 71:25(1979).

4. Chu, S. S. C., and Jeffrey, G. A., Acta Crystallogr. B24: 830(1968).

5. Ham, J. T., and Williams, D. G., Acta Crystallogr. B29: 1373(1970).

6. Atalla, R. H., Advances in Chemistry Series, No. 181, 1979, p. 55.

7. Atalla, R. H. Proceedings of the International Symposium on Wood and Pulping Chemistry, Stockholm, June, 1981. SPCI Report No. 38, Vol.1, 1981, p.57.

8. Atalla, R. H. Proceedings of the 9th Cellulose Conference, Appl. Polymer Symp., No. 37, 1983, p. 295.

9. Atalla, R. H., Ranua, J., and Malcolm, E. W., Tappi J. 67:96(1984).

10. Atalla, R. H., and VanderHart, D. L., Science 223:283 (1984).

11. VanderHart, D. L., and Atalla, R. H., Macromol. 17:1465 (1984).

12. Atalla, R. H., and Nagel, S. C., Science 185:522(1974).

13. Whitmore, R. E., and R. H. Atalla, R. H., Int. J. of Biol. Macromol., in press.

14. Marrinan, H. J., and Mann, J., J. Polym. Sci. 21:301(1956).

15. Sarko, A., and Zugenmaier, P., ACS Symposium Series, No. 141, 1980, p. 459.

Transcription of Discussion

Studies on Polymorphy in Native Cellulose
by R.H. Atalla

Marchessault A recent paper by Cael Choi and Battarchargee (Macro molecules - 1985) interprets the multiplicity of the C-1 resonance of native cellulose in CP/MAS NMR as relating to varying contributions of the "eight-chain" and "two-chain" unit cells. Why is this not a more straight forward interpretation of your C-1 data rather than I_α and I_β phases?

Prof. R.H. Atalla The same question was raised in a letter to "Science" in January, 1985. My response to that was that the present data tell us that we are dealing with a composite of two crystalline forms which we have labelled I and I and the data do not allow us to indicate what the structure is. I am proposing that it includes the same molecular conformation but with different hydrogen bonding patterns. Cael et al's paper is purely speculative and therefore does not lead me to take a different view from what I have just presented.

Mason I there any place in your structural pictures in wood-cellulose for lignin-carbohydrate bonds of which one used to hear a great deal?

Atalla Based on the spectroscopic studies I do not think it is possible to comment on that. NMR or Raman spectroscopy at this stage in our understanding cannot provide evidence for or against this view.

Roberts I am not quite sure why the I_α spectrum that you propose has C-4 in two different environments with C-1 in one environment. Is that a natural consequence of the structures which you are proposing?

Atalla We have not proposed a particular structure. We are proposing that native cellulose is a composite and at this stage we do not have a good answer to that question.

Marchessault The C-6 resonance of native cellulose is interpreted by Horii *et al* in terms of the *tg* conformer for O-6. None of your models propose this conformer. Why not?

Atalla The model I have shown is essentially an adaptation of the structure of cellobiose.