特別講演 3

The Individual Structures of Native Celluloses: Species Specificity of Secondary and Tertiary Structures and its Implications with respect to Processes of Biogenesis

R H. Atalla

USDA Forest Service, Forest Products Laboratory
Madison, Wisconsin 53705, USA

INTRODUCTION

The diversity of the structures of native celluloses is important because, more than any other natural substrate, the susceptibility of cellulose to chemical reagents or to enzyme action is as much a function of its state of aggregation as it is of its primary structure. And the state of aggregation of each native cellulose is specific to the species within which it is formed. This has not been generally recognized, in part, because studies of cellulose structure view cellulose within the conceptual framework of polymer science without adequate attention to its biological character. Studies of structure usually seek to establish an ideal crystal lattice for native cellulose and regard all real celluloses as departing from this ideal form with respect to the degree of disorder. The disorder is viewed as imparting some amorphous character and methods to quantify the "amorphous" fraction are sought.

In contrast, recent studies have shown that all celluloses that are relatively pure in their native states are composites of two distinct forms that are intimately blended in a manner and in proportions that are unique to the species producing them. The far wider category of celluloses that, in their native states, are intimately blended with other polysaccharides and with lignin, are even more diverse with respect to their patterns of aggregation. This report will discuss efforts to characterize the nature of native celluloses and their states of aggregation in a manner that may allow more comprehensive investigations of the nature of the interaction between the cellulosic substrates and the chemical and biological systems that act upon them. One intent of the discussion is to make the case that, in the context of biological studies, it is no longer adequate to regard cellulose as the β-1,4-linked polymer of anhydroglucose and to apply to it the usual principles and methods of polymer science; these principles and methods do not begin to capture the distinctive features of the molecule which are derived from its biological origin.

In this report we first address the issues that complicate characterization of the structure of cellulose. We then review recent developments concerning our understanding of structure in native celluloses. Information developed from spectroscopic investigations that have been undertaken in recent decades is presented. The patterns of self-assembly are then considered on the basis of observations of its biogenesis in model systems incorporating cultures of *Acetobacter xylinum.*

BACKGROUND

In recognition of the important biological functions of cellulose, there is a need to reclaim it from the field of polymer science and to recognize that it is a unique biological molecule with many distinctive properties that are key to its role in biological processes. Over the past 100 years much of the research on cellulose has been undertaken in relation to its utilization in industrial processes. It is in this context that many of the principles of polymer science were first developed, though they were also extended considerably in relation to synthetic polymers. These methods, however, are not adequate for the native celluloses. The majority of crystallizable synthetic polymers, as well as cellulose regenerated from solution, can be viewed in terms of heterogeneous combinations of separate domains that are amorphous or crystalline, depending on the past history of the sample. For native celluloses, in contrast, it is increasingly obvious that they must be regarded as self-assembling naturally occurring substances, and that the self-assembly occurs at a hierarchy of levels that cannot be understood in terms of the simplistic notions of crystallization and separate phases, both of which are based on the classical thermodynamics of macroscopic systems.

When considering the structures of native celluloses we distinguish two categories of native forms. The first category, the Pure Celluloses, includes those that occur in relatively pure form in their native state and that can be isolated using mild procedures that do not alter the state of aggregation to any significant degree. Such celluloses have been the subject of most structural studies. Examples are cotton, bacterial cellulose and some of the algal celluloses. The second category, the *Complex Celluloses*, which includes the vast majority of naturally occurring forms, consists of the celluloses that are an integral part of the complex architecture higher plant cell walls. In the vast majority of such plants, the celluloses are intimately blended with the other cell wall constituents. The celluloses in this second category undergo significant changes in their state of aggregation during traditional isolation procedures (1). However, these changes have not been well recognized in most prior work. One of the major challenges yet to be overcome in future studies is the definition and characterization of the native state of such celluloses prior to the application of isolation procedures that are severe and disruptive of structure.

Until recently, it has generally been assumed that the cell wall structure in higher plants can be represented as a two phase system, one consisting of the microfibrils of pure cellulose, the other a blend of all of the other constituents (2). It is implicit in this view that the cellulose from such cell walls can be isolated by removing the other components, leaving behind the cellulose in a condition that approximates its native state. Furthermore, it is also generally assumed that the microfibrils, both in their native state and after isolation, are similar to those that occur in the pure native celluloses. There are now reasons to reassess the assumptions, and the reassessments are at two levels (3). The first follows from observations that it is not possible to remove the other cell wall constituents without altering the cellulose at the microfibrillar level. The second follows from evidence that raises questions concerning validity of the two phase model.

Much of the discussion of the pure celluloses has concentrated on secondary and tertiary structure at the level of the unit cell (with characteristic dimensions of the order of 1 nm), and on the equivalence or non-equivalence of different monomeric or dimeric units within the unit cell (3). What are more relevant to the species specific character of native celluloses, however, are the implications of organization at the level next above that of the unit cell; for native celluloses that level is the most elementary of fibrillar structures. And it is at this level that departures from the organization of an ideal lattice first manifest themselves. They are most clearly obvious in the curvature and twisting that is seen in most electron micrographic images of native celluloses. The key point is that because they represent departures from an ideal lattice they influence the results of all measurements that are sensitive to lattice order. Measurements carried out on a substance with an ideal lattice would result in diffraction patterns that are very sharp and spectral lines that are very narrow. Departures from such a lattice result in broadening of both diffraction and spectral measurements, and such broadening is frequently used in efforts to quantify order in materials or arrive at measures of the degrees of their crystallinity. Such measurements have also been adopted in attempts to quantify the degree of order in cellulosic samples and they have been used to explore correlations between order in the cellulosic substrates and their susceptibility to chemical reagents or enzymatic action. The difficulty with this approach is that the traditional interpretations of broadening in diffractometric patterns or spectral lines have been in terms of the occurrence of domains of homogeneously disordered matter. While assumptions that such domains occur are reasonably valid approximations in the case of most synthetic polymeric materials, and for regenerated cellulose, they are very questionable in the case of native celluloses.

The reality is that most native celluloses are not disordered but rather highly organized in a hierarchy of structures that are defined at different scales of observation of the native tissue within which the cellulose occurs. The most elementary level beyond the unit cell is the microfibril, with characteristic lateral dimensions of 4 to 6 nm, and with longitudinal dimensions considerably in excess of 10 nm. With the exception of selected algal and tunicate celluloses, which can have microfibrils with lateral dimensions of the order of 20 nm, all native celluloses depart from the ideal lattice at the level of the microfibril. These departures represent features that are unique to a particular native cellulose, and are the primary basis for species specificity of structure.

The curvature and the twisting of the fibrils necessarily result in departures from ideal lattice order. In order to characterize these departures it is useful to consider the different levels of structure in cellulose. The primary structure has been well established for some time and is no longer in question. The secondary level of structure, that is, the conformation, reflects the internal organization of individual monomeric units within a molecule. The tertiary level then is concerned with the arrangements of the molecules relative to each other in a particular aggregated form. In much of the literature on cellulose structure, which is based on diffractometry, the distinctions between primary, secondary, and tertiary levels are not considered because specification of the coordinates of the atoms within a unit cell implicitly defines all three levels of structure. The recent findings mentioned above are based on spectroscopic observations, and for their interpretation the distinctions between primary, secondary and tertiary levels of structure is important. The distinction between the different levels of structure is also likely to be helpful in future studies of the interactions of cellulose with any agents of chemical transformation. The ability to distinguish between secondary and tertiary structures and to characterize them separately opens up the possibility of more precise interpretation of the results of observations. For example, mechanistic analyses of effects that are related to steric and conformational differences may be attributable to variability in secondary structures, while those that arise from differences in accessibility may be more directly related to tertiary structure.

RECENT DEVELOPMENTS

Spectroscopic studies

The characterization of structure by spectroscopic methods has progressed along two lines recently. Raman spectral studies revealed that there is greater diversity in the secondary structures of celluloses than previously recognized on the basis of diffractometric investigations. They also suggested that adjacent anhydroglucose units in the chain are not symmetrically equivalent and that the basic repeat unit of structure must be recognized as anhydrocellobiose. But perhaps of broader significance was the finding, on the basis of solid state ^{13}C NMR investigations, that native celluloses are composites of two distinct forms and that the organization of the composite structures is distinctive of the particular species or tissue within which the cellulose is produced.

It is beyond the scope of this report to review the Raman spectral findings in full; a brief overview, however, is helpful. The first key finding was that the crystallographic models propounded in the 1970s for cellulose could not rationalize the differences between the spectra of native celluloses (form I) and those of the regenerated or the mercerized celluloses (form II). The interpretation of the spectral observations included investigations of the vibrational spectra, both infrared and Raman, of many model systems. When the spectra of the celluloses were evaluated in this context it became clear that structures described as twofold helical conformations of the chains represented approximations of the true structures but they ignored some of the key differences detected from the spectral information. Although the departures from twofold helical organization were small, they were found to be at the heart of the distinctive features of the structures of many celluloses. In summary, the conclusions were that the adjacent anhydroglucose units in a cellulose chain are nonequivalent and that the nonquivalences take different forms in celluloses I and II. In cellulose II, they are centered at the glycosidic linkages, while in native celluloses the departures from equivalence are somewhat less at the glycosidic linkages but greater at the primary alcohol group at C6 (4).

The inadequacy of crystallographic models brought out by the Raman spectral studies were also encountered in the solid state ^{13}C NMR studies of the celluloses. The earliest of the NMR studies focused on the diverse forms of cellulose and attempted to resolve the question of nonquivalences among successive anhydroglucose units in the cellulose chain. It soon emerged, however, that the solid state ^{13}C NMR spectra revealed evidence of a structural diversity among native celluloses that had only been hinted at by earlier results (5,6). The spectra show a number of complex multiplet resonances associated with carbons that are chemically equivalent. After investigation of alternative interpretations of the diversity of the spectra of the different native celluloses, the possibility of composite structures was explored. It was found that the spectra of the native forms can be resolved into linear combinations of spectra corresponding to two forms of cellulose I; these have been designated I_α and I_β to distinguish them from earlier efforts to categorize the diversity of native celluloses.

When a wide variety of pure native celluloses were investigated, a number of patterns emerged. First it was clear that all of the celluloses from plant sources are blends or composites of the two forms of cellulose and that the particular blend is specific to the species and tissue from which the cellulose is isolated. The spectrum of a particular cellulose is distinguished by the relative proportions of the I_α and I_β forms and the degree of resolution of the resonances associated with each of the carbon atoms. Next, it was noted that among the pure celluloses from higher plants the I_β form always appears to be dominant. In contrast, the bacterial celllluloses and all the algal celluloses of high crystallinity are predominantly of the I_α form. Finally, it was obvious that the spectra recorded from samples that had been shown to be highly ordered on the basis of diffractometry exhibit the sharpest resonances in the multiplets and have the smallest shoulders associated with resonances of C4 and C6.

The differences between the I_α and I_β forms have been explored using two approaches. Electron diffractometric studies have indicated that the I_α form has a triclinic unit cell with one chain per unit cell and with nonequivalent anhydroglucose units. They also suggested that the I_β form has a monoclinic cell of space group $P2_1$ with two chains per unit cell with the twofold screw axis of symmetry coincident with the chains, requiring that adjacent anhydroglucose units be symmetrically equivalent (7). The other line of inquiry has been based on vibrational spectroscopy including both infrared (8) and Raman spectral measurements (9). Both sets of spectral measurements point to similar secondary structures in the two forms but with distinctly different hydrogen bonding patterns. These findings have recently been further confirmed by comparison of the Raman spectra of *Valonia*, an algal cellulose, which is predominantly of the I_α forms, and *Halocynthia*, a tunicate cellulose, which is predominantly of the I_β form. Their comparison is particularly useful because the lateral dimensions of their fibrils are both of the order of 20 nm, and they, therefore, allow the generation of spectra of equal resolution. It is clear from the comparison that the spectra are essentially identical in the low frequency region, which is most sensitive to secondary structure. In the OH stretching region, in contrast, there are significant differences between the band structure for the two forms. The indication here is that the secondary structures of the two forms are the same but the tertiary structures are different. This finding is in sharp contrast to the conclusions derived from the diffractometric data. Another important observation with respect to the Raman spectra is that there is no evidence of correlation field splitting in either of the spectra; such splitting is expected in the spectra of polymeric systems that have more than one chain per unit cell. This observation is not consistent with the interpretation of the diffractometric data that ascribes a two chain unit cell to the I_β form. Thus the results of the infrared and Raman spectral studies appear to be at variance with those of the diffractometric observations. It is clear that the nature of the differences between the I_α and I_β forms is not fully understood at the present time.

Model System

A number of important advances with respect to understanding native celluloses have resulted from studies *Acetobacter xylinum* and the celluloses it can produce. Two aspects will be reviewed briefly. These are the hierarchic assembly of native celluloses and the regulation of cellulose aggregation by hemicelluloses.

As noted earlier, native celluloses must be regarded as self assembling biological molecules that are organized hierarchically. This is obvious from examination of the microfibrillar structure of bacterial celluloses produced under different conditions that can perturb the process of assembly at the different levels. The most common form of the cellulose is produced in fibrillar form with the individual fibrils of the order of 6 nm in width. They have the form of a ribbon with a regular twist to it (10); recently it was shown to be right handed (11). Perturbations of growth which have been explored include the effects of carboxymethyl cellulose (CMC) or those of the fluorescent brightening agent (FBA) calcofluor. The CMC results in reduction of the coherence of the ribbons which appear like braids of elementary fibrils; it points to the capacity of molecules similar to cellulose to associate with it and to limit the degree of self-assembly. The effect of the FBA, which can associate even more strongly with cellulose, is even more dramatic than that of CMC. When the FBA is washed away with dilute acid, micrographs reveal that it has limited the capacity of cellulose to self-assemble so that the substructures are of the order of 1.5 nm in lateral dimension.

The patterns of aggregation displayed by the ribbons of cellulose formed under different levels of perturbation indicate that the most basic building blocks are the 1.5 nm fibrils emerging from a single assembly complex. These then would assemble with similar fibrils emerging from adjacent complexes. In the absence of agents capable of associating with the nascent cellulose this would result in the normal ribbons. If the pattern of self assembly is an inherent characteristic of the cellulose molecule in its nascent state, it can be anticipated that the same influences would prevail during the formation of cellulose in higher plants, and that the basic building blocks would have similar characteristics. This is in fact what has been observed in much of the electron microscopic examination of celluloses from higher plants.

In studies of the *Acetobacter xylinum* model system cited above, the modifying agents were added to the culture medium. In higher plants, in contrast, it is expected that during the formation of the cell wall matrix many agents will be present that can interact with cellulose to modify its assembly. The possible effects of matrix components were investigated in studies of the *Acetobacter xylinum* system where different hemicelluloses were added to the cultures at relatively low levels (12,13,14). It was observed that the hemicelluloses are indeed incorporated into the structures of the celluloses and that each of the hemicelluloses elicited a different pattern of modification of the cellulose produced in its presence. Furthermore, the changes induced by the presence of the hemicelluloses resulted in aggregation of the celluloses in a manner similar to that of higher plants. In particular, the solid state ^{13}C NMR spectra revealed a predominance of the I_β form as a result of the influence of the hemicelluloses. It can be anticipated on this basis that the patterns of self-assembly and the final tertiary structures of the native celluloses of higher plants are likely to be as diverse as the plants within which they are formed.

On the basis of the observations of the model systems it may be concluded that the celluloses in the category defined as pure celluloses are formed in environments where the other cell wall components cannot associate with the nascent cellulose sufficiently to modify its tertiary structures, while the complex celluloses are formed in cell wall matrices that include constituents that can modify the tertiary structure. A clearer definition of the diversity of the tertiary structures is essential to advancing our understanding of the species specificity of native celluloses. Further elaboration of the diversity of native celluloses is beyond the scope of this report; the author has addressed many aspects in greater detail elsewhere (3).

REFERENCES

1. Atalla, R. H., Whitmore, R., *J. Polym. Sci., Polym. Lett. Ed.*, **16**, 601 (1978).
2. Preston, R. D., "The Physical Biology of Plant Cell Walls", Chapman and Hall, London (1974).
3. Atalla, R. H., In "Comprehensive Natural Products Chemistry", Vol. 3, Ed. by Pinto, B. M., Chapter 16, pp. 529-598 (1999).
4. Atalla, R. H., In "The Structures of Cellulose", Ed. by Atalla, R. H., ACS Symp. Ser. No. 340, Am. Chem. Soc., Washington, D.C., p. 1 (1987).
5. Atalla, R. H., VanderHart, D. L., *Science*, **223**, 283 (1984).
6. VanderHart D. L., Atalla, R. H., *Macromolecules*, **17**, 1465 (1984).
7. Sugiyama, J., Okano, T., Yamamoto, H., Horii, F., *Macromolecules*, **23**, 3196 (1990).
8. Sugiyama, J., Vuong, R., Chanzy, H., *Macromolecules*, **24**, 4168 (1991).
9. Atalla, R. H., Agarwal, U. P., Hackney, J. M., Isogai, A., to be published.
10. Haigler, C. H., In "Biosynthesis and Biodegradation of Cellulose", Ed. by Haigler, C. H., Weimer, P. J., Marcel Dekker, New York, pp. 99 (1991).
11. Hirai, A., Tsuji, M., Horii, F., *Sen'i Gakkaishi*, **54**, 506 (1998).
12. Uhlin, K. I., Atalla, R. H., Thompson, N. S., *Cellulose*, **2**, 129 (1995).
13. Atalla, R. H., Hackney, J. M., Uhlin, K. I., Thompson, N. S., *Int. J. Biol. Macromol.*, **15**, 109 (1993).
14. Hackney, J. M., Atalla, R. H., VanderHart, D. L., *Int. J. Biol. Macromol.*, **16**, 215 (1994).

Problem 3 **Increased understanding of the biogenesis and molecular architecture of wood cell walls, their response to environmental stresses, and their transformation in the course of industrial processing.**

FY 1999 Research Attainments

Publications

Research Unit

Atalla, R. H. 1999. The individual structures of native celluloses: Series specificity of secondary and tertiary structures and its implications with respect to processes of biogenesis. In: '99 Cellulose R & D, 6th annual meeting of the cellulose society of Japan: 1999 JUNE 15-16; Sapporo, Japan. Sapporo, Japan: '99 Cellulose R & D, 6th Annual Meeting of the Cellulose Society of Japan: 9-12.