3.16 Celluloses

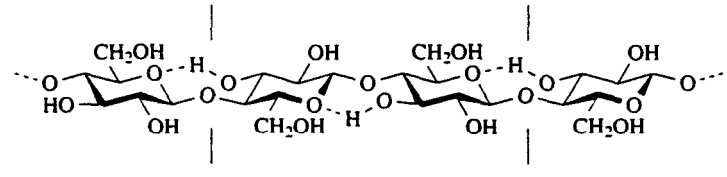
RAJAI H. ATALLA
USDA Forest Service and University of Wisconsin, Madison, WI, USA

3.16.1 INTRODUCTION

Cellulose is, in many respects, among the most challenging of the polysaccharides. Although it is a dominant component in the vast majority of plant forms and has a number of vital biological functions, and although cellulose based materials have been part of daily life for many millennia, our understanding of its nature remains incomplete. This circumstance is not the consequence of

lack of interest in investigations of cellulose, for the effort to develop a scientific characterization of cellulosic matter is well over 150 years old. The constraints on our understanding of cellulose and its phenomenology are rather rooted in the complexity of its behavior and the inadequacy of our conceptual frameworks and methodologies for the characterization of so unusual a substance; perhaps more than any other common chemical species, its chemistry and its entry into biological processes are as much a function of its state of aggregation as of its primary structure.

In the modern era, the beginning of the search for scientific understanding can be associated with the work of the French agriculturalist Anselm Payen, who introduced the term cellulose in 1842 to denote the neutral structural polysaccharides of plant tissue: he viewed the isolated matter. which had an elemental composition similar to that of the saccharides, as an aggregated form of "dextrose." By the turn of the century, when methods for discrimination between the monosaccharides were established, cellulose was recognized to be an aggregated, oligomeric form of anhydroglucose that remained the subject of controversy. Within the context of the polymer hypothesis. developed in the early decades of the twentieth century, cellulose has been recognized as the linear *b*-(1-4)-linked homopolymer of anhydroglucose (1). More recently, it has been noted that it is more accurate to define it as the homopolymer of anhydrocellobiose, since such definition explicitly incorporates the identity of the linkage. Given this definition, it might be anticipated that the nature and the chemical behavior of cellulose could be understood in terms of the chemistry of its monomeric constituents together with considerations arising from its polymeric nature. The reality, however, is that the chemical or physical transformations of particular samples of cellulose are as much reflections of their history and their states of aggregation as they are of the chemistry of their monomeric constituents. Indeed, this is the basis for the title *Celluloses* given to this chapter in reflection of the great diversity of the forms of cellulose. The unusual nature of cellulose in its aggregated states is perhaps best represented by its relationship to water; a primary structure with three hydroxyl groups per pyranose ring would, in the normal course of events, be expected to be quite soluble in water and aqueous-media. But it is in fact essentially insoluble, and this characteristic is but one of the many unusual patterns of behavior of the aggregated states of cellulose.



(1)

In this chapter, the author is concerned with delineating the frontiers of our understanding of cellulose, particularly with respect to its native forms and the biological processes into which they enter. While the presentation is relevant to the industrial utilization of cellulose, because it addresses the nature of the native forms of many of the feedstocks used as well as the effects of processes of isolation on structure, particular attention will be given to questions confronting the research community undertaking studies of cellulose in biological contexts, including biogenesis, biological function, and biodegradation.

Progress in the characterization of cellulose has been intimately intertwined with advances in methods of chemical analysis and the development of the molecular hypothesis, with the accumulation of additional observations on celluloses from an ever-expanding variety of sources, and with the development of new conceptual frameworks for interpreting the behavior of macromolecules. The evolution in understanding the nature of cellulose has continued during recent decades and will likely be ongoing for some time to come. One objective of this chapter is to capture the central features of this evolution in order to place recent studies of cellulose in a historical perspective and to suggest directions for future work.

One of the major challenges in any experimental program involving the use of cellulose as a substrate is the development of an adequate characterization of the tertiary structure of cellulosic samples. Careful specification of the history of particular samples of cellulose or precise definition of their states of aggregation are usually essential to reproducibility of experimental observations. Traditionally, this reality has been dealt with in experimental programs by reliance on the use of readily available types of pure celluloses such as cotton. or filter paper made from cotton linters. or, frequently in later years, microcrystalline celluloses prepared by acid hydrolysis of different cellulosic substrates. While this provides a point of reference for other investigators who are likely to have access to similar celluloses, it begs the question of an adequate definition of the state of

aggregation of the cellulose; the morphology of these celluloses is complex and quite variable at supramolecular levels in ways that are not easily detected or characterized. Another important objective of this chapter is therefore to address the question of aggregation and tertiary structure and the methodologies for their characterization. In addition to seeking to clarify issues of structure in the context of experimental studies of cellulosic substrates, the methodologies will also be discussed in relation to characterization of the wide variety of native forms of cellulose that are as diverse as the species of organisms that produce them.

We begin with a historical perspective, with some reflection on the different stages in the progress towards better understanding of the nature of cellulose. This is not intended to provide a duplication of some of the very comprehensive discussions of the history of ideas concerning the identity and structure of cellulose. Rather it is intended to illustrate how often the inadequacy of the paradigms available from the chemistry of a particular period has resulted in constraints that have limited progress and often resulted in major detours along the path to clearer understanding of the phenomenology. The discussion is intended to alert us to the reality that the unusual character of cellulose requires innovation in concept as well as in methodology, and to establish a basis for analyzing the conceptual barriers that remain to be dealt with in the next cycle of studies of the nature of celluloses.

The discussion of structure will focus on the results of investigations carried out during the past three decades, though some of the important contributions from earlier periods will not be neglected. Among the important developments has been the introduction of spectroscopic methods to complement information derived from traditional crystallographic methods for characterization of the aggregated states of cellulose. In this arena, the use of solid-state ^{13}C NMR spectroscopy and Raman spectroscopy have resulted in significant new insights concerning differences in the states of aggregation of different celluloses. The application of lattice imaging methods in electron microscopic studies of the algal celluloses has also been an important development during the last decade. Nevertheless, a number of questions concerning the structure of cellulose remain outstanding.

The challenge of defining structure emerges most clearly in an examination of the wide range of native celluloses that occur in nature. They occur in the cell walls of the vast majority of plant forms and are highly ordered structures that are elaborately integrated into biological structural tissues. They are also produced by a number of other organisms, including bacteria and some classes of marine life. Yet the types of order that prevail do not lend themselves to description in terms of the traditional concepts that have been developed for the description of organization in the solid states of inorganic or inanimate organic substances. Although cellulose is often regarded as crystalline, review of observations of its many different forms will reveal that the classical definitions of crystalline order for both molecular and polymeric materials are not well suited for the characterization of cellulose, although they may provide a basis for some useful measures of its organization.

The biogenesis of cellulose has also presented a major challenge to investigators. Early studies focused on the biosynthesis of bacterial celluloses as these systems lent themselves more readily to examination by the available methodologies. A significant number of electron microscopic examinations of the process was carried out. Considerable effort has also been invested in studies of biogenesis in selected higher plants, cotton being the primary one among these. The pathways for biosynthesis were established by administration of radiolabeled precursors and following their entry into the structure of cellulose. A significant effort has been invested in pursuing genetic encoding of the biosynthetic enzymes. But much remains unknown in this arena as well.

The diversity of celluloses and the difficulty in characterizing their states of aggregation have been complicating factors in studies directed at understanding the action of different agents on cellulosic substrates whether the agents be chemical or biological. Because of the heterogeneity of the substrates, issues of accessibility arise, and these, in turn, are intimately related to states of aggregation. The literature on biodegradation of cellulose and on the action of cellulolytic enzymes has many examples of the effects of sample history and structure on its response. The literature on chemical modification in many ways parallels that on the action of cellulolytic enzymes. Indeed, it goes further in the sense that differences in responses to the action of chemical modifying reagents have been used to characterize the differences between states of aggregation of cellulosic samples with different histories.

A wide range of methodologies of modem polymer science rely on the measurement of properties of the polymers in solution. The considerable difficulty in solubilizing cellulose has been a major hindrance to developing adequate methodologies. Almost all of the solvent systems for cellulose are multicomponent solvents, many of them acting by forming metastable derivatives. The majority of them are aqueous solutions of reagents that enter into strong associative interactions with cellulose.

Some are used in some of the established analytical procedures for characterizing celluloses. Others are used in industrial processes for regeneration of celluloses to form a variety of products. A number of organic solvent-based systems have also been developed. Some are used in analytical procedures and as media for some chemical derivatization reactions. One is also used for the regeneration of cellulose in fiber form. It is also of interest that a number of solutions, some aqueous and others nonaqueous, are capable of swelling cellulose without resulting in dissolution. They are of particular interest because many of them can penetrate the highly ordered semicrystalline domains sufficiently to facilitate a solid-state transformation that can result in a different type of lattice order in the semicrystalline domains. However, the solvent systems and their behavior are beyond the scope of this chapter.

The utilization of cellulosic materials is as old as civilization and predates by millennia any effort to understand its nature at a constitutional level. A number of uses established in past millennia remain at the heart of major industries. Those most commonly cited are those components of the textile industry that are based on cotton, and the pulp and paper industry. Of course the utilization of wood remains central also in the daily life of most cultures, and cellulose is the key structural component in wood. Here again, the technology is beyond the scope of a treatise on natural products.

In the years ahead it is likely that advances in methodologies for the characterization of biological order will be important. New instrumental methods will no doubt be developed allowing more detailed characterization of the diversity of celluloses. Advances in molecular biology related to both biogenesis and biodegradation are also likely to play key roles in advancing understanding of the nature of celluloses and their many entries into the biosphere. The author looks ahead and attempts to anticipate the pre-occupations of cellulose science that are likely to be dominant in the decades a head.

3.16.2 HISTORICAL PERSPECTIVE

Progress in understanding the nature of cellulose can be viewed as occurring in three distinct though overlapping stages. The first can be defined as the period between the early work of Payen, recognizing the carbohydrate-like character of the key constituents of plant cell walls, and establishment of the polymer hypothesis in the second and third decades of the twentieth century. Efforts during this period were focused on establishing the primary structure of cellulose, that is, the pattern of covalent bonds in the cellulose molecule. Though modem day cellulose scientists take for granted knowledge of this structure, the limitations of both methodologies and conceptual frameworks resulted in extension of this period from the time of Payen to the 1920s, when the polymer hypothesis received confirmation from a number of sources, and methods for distinguishing the different monosaccharides were well developed.

The second period can be regarded as the interval between the general acceptance of the polymer hypothesis, in the mid- 1920s. and the 1970s when a number of new methodologies for characterizing cellulose were introduced and provided a basis for rationalizing the great diversity of native celluloses. During this second period the general thrust of most structural studies was towards establishing the secondary and tertiary structures of cellulose and was based, in large measure, on application of the methods of crystallography. Much of the effort was implicitly based on the premise that a single crystal structure, identified as cellulose I, was basic to all native celluloses. However, there was very little consensus as to the fine points of this structure. A second crystallographic structure, identified as cellulose II, was regarded as the basic constituent of both regenerated and mercerized celluloses: this latter structure did not receive as much attention as that of cellulose I, in part because it was of limited significance to biological questions and in part because specimens of cellulose II of relatively high crystallinity could not be prepared.

The second period also witnessed the broad application of the principles of polymer science to the study of cellulose and its derivatives. Many aspects of the phenomenology of cellulose were interpreted by analogy with the phenomenology of its derivatives and that of some of the relatively more simple synthetic polymers that had become of considerable interest, both commercially and scientifically, early in this period.

In addition to advances in understanding the nature of cellulose as a substance, the second period saw the beginning of serious inquiries and many important contributions to new understanding of the role of cellulose in the biological systems within which it occurs. There were many investigations of the nature of its organization at the microscopic level within the cell walls of many plant

forms. Important studies of the processes of biosynthesis and biodegradation of cellulose were also undertaken. These will not be included in this historical perspective which is primarily concerned with issues of structure and organization in cellulose in its many aggregated states, and with the conceptual frameworks available for their description.

The third and current period in studies of the structure of native celluloses can be viewed as beginning with the discovery that all plant celluloses are composites of two closely related crystalline lattice forms, I_α and I_b that occur in distinctive proportions characteristic of the species producing the particular cellulose under observation. The uniqueness of the blend to each particular native form provides a basis for resolving some of the earlier conflicts in interpretations of crystallographic data which had been acquired from celluloses derived from different organisms. In addition to the advances in understanding issues of structure, the current period has witnessed progress in understanding both the biosynthesis and biodegradation of cellulose, as well as in characterizing some of the most interesting of cellulose derivatives. This third stage in investigations of the many aspects of the nature of cellulose and its entry into biological and industrial processes is likely to continue for many decades.

In parallel with the progress in understanding the nature of cellulose, there were, of course, many advances in the industrial technologies for the utilization of cellulose, the majority of which predate knowledge of its molecular structure. The two primary industries based on the utilization of cellulose in its native forms are those concerned with the production of textiles from different plant fibers, and with the production of paper from cellulosic fibers isolated primarily by pulping of woody species. In addition, the latter half of the nineteeth century and the first half of the twentieth century witnessed significant industrial utilization of cellulose derivatives in the manufacture of a wide array of products; although a number of the products fabricated from cellulose derivatives have been displaced by ones based on synthetic polymers, many cellulose derivatives remain important products with a wide range of applications. These three major industrial sectors have provided, over the years, considerable motivation for many of the research efforts that have resulted in advancing the understanding of cellulose. The technologies based on the utilization of cellulose and their histories are covered by broad categories of literature dedicated to each of them.

3.16.2.1 The Early Period (1840-1920)

An excellent account of the development of the early ideas concerning the nature of cellulose is presented by Purves.[1] It is very well complimented by the historical introduction in the treatise by Hermans,[2] and, to a more limited degree, at least with respect to cellulose, by the historical introduction in the classical treatise on polymer science by Flory.[3] The analysis presented here is based on these sources as well as some of the early texts on cellulose chemistry.[4,5]

Payen's seminal contribution arose from recognition that plant tissues from a number of different sources. when adequately "purified," resulted in substances that, when submitted to elemental analyses, appeared to have proportions of carbon, oxygen, and hydrogen that were similar to those of the sugars and starch. This led him to the view that another polysaccharide related to starch was one of the key constituents of plant cell walls. He was the first to argue that the different elemental compositions found in-other plant tissues, particularly lignified ones, were the consequence of inadequate removal of other substances that encrusted the primary cell wall constituent. His view was further reinforced by the observation that the purified cell wall substances, when subjected to acid hydrolysis, produced dextrorotatory reducing sugars not unlike those obtained upon hydrolysis of starch. He proposed that all plant cell walls contain a uniform chemical constituent made up of glucose residues; he viewed it as isomeric with starch. Since starch stained readily with iodine, while the cell wall residue gave a similar color only upon swelling with sulfuric acid, he regarded the cell wall substance as a more highly aggregated isomer and named it cellulose. Payen is likely to have based his conclusions on studies of cellulose from sources such as cotton or flax, wherein the cellulose is not heavily encrusted with constituents that would be more difficult to remove. Though he clearly was not aware of the possibility that other polysaccharides could co-exist with cellulose in the cell wall, Payen's fundamental insight concerning the existence of a primary structural component in all plant cell walls has proved to be an accurate one.

The next major advance in understanding the constitution of cell walls came in the decade between 1880 and 1890, when methods for discriminating between the monosaccharides were developed, and many celluloses that were comprehended within Payen's definition were shown to contain other sugars. Soon thereafter Schulze, who observed that the majority of the polysaccharides that contain

residues other than glucose are more readily hydrolyzed, suggested that the more resistant component of the cell wall be identified as the cellulose. Thus, the approach adopted by Schulze followed that of Payen; both were influenced in this direction by the adoption of cotton as the standard for cellulose, a standard that continues in many laboratories to this day. The other polysaccharides in the cell walls of most plants were then categorized as the hemicelluloses. In due course, the noncarbohydrate components of the cell wall were recognized to be chemically distinct and to include lignin in addition to a variety of minor constituents that were extractable with organic solvents.

It should be noted that the advances made during this period occurred in an environment wherein a number of investigators in plant science resisted the categorizations of the constituents of plant cell walls outlined above and gave preference to the view that the lignocelluloses derived from different species represented different chemical individuals. This view was reinforced by the observations that cell wall preparations from different sources responded differently in staining reactions that have long been used to distinguish between chemical constituents of biological tissues. This controversy is in many respects similar to others that have arisen over the years and that are rooted in the differences between patterns of aggregation of cell wall constituents at the supramolecular level. Resolution of the issue of the distinctness of cellulose as a constituent was finally accomplished when X-ray diffractometery demonstrated that diffraction patterns derived from samples of wood and other plant tissues before and after delignification were essentially the same.

Other major influences during the early period were the ongoing developments in organic chemistry and the many investigations associated with the industrial utilization of cellulose and its derivatives. The situation is best described in an excerpt from Purves:[1]

> Concurrently, structural organic chemistry in general was being built up with brilliant success by arguments based on the analyses of substances which were chemical individuals. But cellulose and its derivatives were generally fibrous, insoluble materials which had no sharp melting or boiling points and which gave negative or indefinite results in molecular-weight determinations. Deprived of the usual criteria of homogeneity and purity, investigators in the cellulose field very frequently made the mistake of assuming these qualities in their preparations on insufficient evidence and accordingly attached molecular meanings to their analytical data.

In these circumstances it was not surprising that different investigators, using preparative procedures that may have differed in detail, would ascribe different formulas to what would have been similar chemical entities.

The other group of scientific activities that had an important influence on the interpretations of the phenomenology of cellulosic preparations were those associated with the development of colloid chemistry, which was one of the key areas of physical chemical investigation. The similarities between the behavior of cellulose and its solutions and solutions of its derivatives, on the one hand, and those of colloidal systems, on the other, led many to conclude that the cellulosics were also colloidal. Among the many similarities were observations of colligative properties including osmotic pressures, the abnormal viscosity behavior. and the manner in which quantitative data can depend on the previous history of a particular sample. This led to the proposal of many different structures for cellulose that were based on the premise that it is an aggregate of one or another of a number of oligomeric structures that were consistent with both the elemental analysis and the results of hydrolytic action by acids.

The matter of the primary structure of cellulose was brought to closure in the early part of the twentieth century when crystallographic studies suggested an extended linear molecule and the methodologies of carbohydrate chemistry had advanced to the point that permethylation analyses supported the view that the primary linkage between the anhydroglucose units is of the *b*-1,4 type. In the closure of this brief overview of the early period it is well to include the closing paragraph from the essay by Purves.[1] He noted:

> It may seem at first sight surprising that the association theory retained its grip over the intellects of so many workers for so long a period. But it must be remembered that their intuition, although wrong in one dimension, was certainly right in two. Since the time when the thread-like nature of the cellulose molecule was placed beyond question, many researchers have been concerned with the association and interplay of the secondary valence forces that radiate laterally from the cellulose chain

As will be clear from the following discussions, the question of the lateral associations of the cellulose chains remains one of the key issues in studies of the secondary and tertiary structures of cellulose to this day.

3.16.2.2 The Middle Period (1920-1975)

The middle period in the effort to understand the nature of cellulose was focused on exploration of the lateral organization of the long chain molecules that were now almost universally accepted. It also included the beginning of important investigations of the role of cellulose in the biological contexts in which it occurs. The exploration of structure proceeded at two levels. The first was the solution of the crystallographic problem, the second was the effort to apply some of the ideas of colloid science to understanding the lateral organization of native celluloses at levels that were an order of magnitude greater than those addressed by the solution of the crystallographic problem. The work of this period is covered in a number of important reviews. Among the most comprehensive reviews of the crystallographic work are those by Jones[6] and by Tonessen and Ellefsen.[7] The matter of lateral organization at the fibrillar level has been discussed most extensively in the treatise by Hermans[2] and in the classics on cell wall structure by Preston[8] and Frey-Wyssling.[9] although all three have also presented extensive discussions of the crystallographic problem. Atalla[10] has also presented an overview of the same structural questions in light of the early developments of the applications of spectroscopic methods. The reader is referred to these sources for much more detailed discussions of the many issues addressed in the effort to reconcile the wide range of diverse observations during this period. The classic treatises by Preston[8] and Frey-Wyssling[9], also include extensive discussions of the biological dimensions of cellulose science. In the following, the author first focuses on the crystallographic problem. He then examines issues of lateral aggregation at the next higher level of organization.

3.16.2.2-1 Crystallographic studies

From the present perspective, it would appear that exploration of structure during the middle period was dominated by the newly developed methods of crystallography. Though it was rarely stated explicitly, analyses of the diffractometeric data implicitly sought definition of both the secondary and tertiary structures of the molecules. For clarity it is well to note here that the secondary structure represents definition of the relative organization of the repeat units in an individual chain, that is, the conformations of the chain, while the tertiary structure describes the arrangement of the molecules relative to each other in a particular state of aggregation. It is in the nature of crystallographic methods that definition of the structure of a unit cell in a crystal lattice of a polymeric material implies knowledge of the conformations of the individual molecules as well as their arrangement relative to each other. The distinction between these two levels of structure was of little note until spectroscopic methods became important.

The procedures for structural studies on cellulose have much in common with investigations of structure in polymers in general, in most instances diffractometric data are, not sufficient for a solution of the structure in a manner analogous to that possible for lower molecular weight compounds which can be made to form single crystals. It becomes necessary, therefore, to complement diffractometric data with structural information derived from studies carried out on monomers or oligomers.

Kakudo and Kasai[11] have summarized the central problem well:

> There are generally less than 100 independently observable diffractions for all layer lines in the x-ray diagram of a fibrous polymer. This clearly imposes limitations on the precision which can be achieved in polymer structure analysts, especially in comparison with the 2000 or more diffractions observable for ordinary single crystals. However, the molecular chains of the high polymer usually possess some symmetry of their own, and it is often possible to devise a structural model of the molecular chain to interpret the fiber period in terms of the chemical composition by comparison with similar or homologous substances of known structure. Structural information from methods other than x-ray diffraction (e.g., infrared and NMR spectroscopy) are also sometimes helpful in devising a structural model of the molecular chain. The majority of the structural analyses which have so far been performed are based on models derived in this way. This is, of course, a trial and error method.

Similar perspectives have been presented by Arnott,[12] Atkins,[13] and Tadokoro.[14,15]

An acceptable fit to the diffractometric data is not the ultimate objective, however. Rather it is the development of a model that possesses a significant measure of validity as the basis for organization, explanation. and prediction of experimental observations. With respect to this criterion, the models of cellulose which were developed in the middle period leave much to be desired, for their capacity to integrate and unify the vast array of information concerning celluloses was limited indeed.

Quite early in the X-ray diffractometric studies of cellulose, it was recognized that its crystallinity is polymorphic. It was established that native cellulose, on the one hand. and both regenerated and mercerized celluloses, on the other, represent two distinct crystallographic allomorphs.[16] Little has transpired since the early studies to change these perceptions. There has been. however, little agreement regarding the structures of the two forms. For example, Petitpas *et al.* [17] have suggested, on the basis of extensive analyses of electron-density distributions from X-ray diffractometric measurements, that chain conformations are different in celluloses I and II. In contrast, Norman[18] has interpreted the results of his equally comprehensive X-ray diffractometric studies in terms of similar conformations for the two allomorphs.

At a more basic level than comparison of celluloses I and II, the structure of the native form itself has remained in question. Among studies in the 1970s, for example, Gardner and Blackwell,[19] in their analysis of the structure of cellulose from *Valonia ventricosa*, assumed a lattice belonging to the $P2_1$ space group, with the twofold screw axis coincident with the molecular chain axis. Hebert and Muller,[20] on the other hand, in an electron diffractometric study of a number of celluloses including *Valonia*, confirmed the findings of earlier investigators who found no systematic absences of the odd order reflections forbidden by the selection rules of $P2_1$, and concluded that the cellulose unit cells do not belong to that space group.

Even when $P2_1$ is taken to be the appropriate space group, the question of chain polarity remains. As noted by Jones,[21] and by Howsmon and Sisson,[22] the structure initially proposed by Meyer and Mark[23] assumed that the chains were parallel in polarity. The structure later proposed by Meyer and Misch[24] was based on the reasoning that the rapidity of mercerization and its occurrence without dissolution required that the polarity of the chains be the same in both celluloses I and II. It was reasoned further that regeneration of cellulose from solution is most likely to result in precipitation in an antiparallel form, and that the similarity between X-ray diffraction patterns of mercerized and regenerated cellulose required that they have the same polarity. It was thus inferred that native cellulose must also have an antiparallel structure.

The premise that regeneration from solution in the antiparallel mode of crystallization is more probable than the parallel mode was shown to be false within a decade of its first presentation.[25] In spite of this finding, the antiparallel organization of molecules suggested by Meyer and Misch remained the point of departure for most subsequent investigators.

When the models incorporating antiparallel arrangement of the chains are extended to native cellulose, they pose serious questions concerning proposed mechanisms for the biosynthesis of cellulose. Plausible mechanisms for simultaneous synthesis and aggregation of antiparallel chains are more difficult to envision. It is perhaps for this reason that proposals of parallel structures for native cellulose developed during the 1970s were embraced by investigators of the mechanism of biosynthesis.

As the last cycle of crystallographic studies during the middle period remains the basis of much discussion of the structure of cellulose in the wider arena of cellulose science, it is well to review the issues arising in greater detail. This will provide a clearer basis for assessing the implications of the new information concerning structure that has been developed since that time.

As noted by Kakudo and Kasai,[11] the primary difficulty in structural studies on polymeric fibers is that the number of reflections usually observed in diffractometric studies are quite limited. In the case of cellulose it is generally difficult to obtain more than 50 reflections. Consequently, it becomes necessary to minimize the number of structural coordinates to be determined from the data by adopting plausible assumptions concerning the structure of the monomeric entity. The limited scattering data are then used to determine the orientation of the monomer units with respect to each other. In the majority of diffractometric studies of cellulose published so far, the monomeric entity has been chosen as the anhydroglucose unit. Thus, structural information from single crystals of glucose is implicitly incorporated in analyses of the structure of cellulose. The coordinates which are adjusted in search of a fit to the diffractometric data include those of the primary alcohol group at C-6, those of the glycosidic linkage, and those defining the positions of the chains relative to each other.

In addition to selection of the structure of the monomer as the basis for defining the internal coordinates of the repeat unit, the possible structures are usually further constrained by taking advantage of any symmetry possessed by the unit cell. The symmetry is derived from the systematic absence of reflections which are forbidden by the selection rules for a particular space group. In the case of cellulose, the simplification usually introduced is the application of the symmetry of space group $P2_1$, which includes a twofold screw axis parallel to the direction of the chains. The validity of this simplification remained the subject of controversy, however, because the reflections which are disallowed under the selection rules of the space group are in fact frequently observed. In most

of the studies, these reflections, which are usually weak relative to the other main reflections, were assumed to be negligible. The controversy continued in part because the relative intensities can be influenced by experimental conditions such as the periods of exposure of the diffractometric plates. Furthermore, the disallowed reflections tend to be more intense in electron diffractometric measurements than in X-ray diffraction measurements. Thus, more often than not, investigators using electron diffraction challenged the validity of the assumption of twofold screw axis symmetry.

The key assumption with respect to symmetry, however, is not the existence of the twofold screw axis as an element of the symmetry of the unit cell, but rather the additional assumption that this axis coincides with the axis of the molecular chains of cellulose. This latter assumption has implicit in it a number of additional constraints on the possible structures which can be derived from the data. It requires that adjacent anhydroglucose units are related to each other by a rotation of 180 degrees about the axis, accompanied by a translation equivalent to half the length of the unit cell in that direction; it is implicit, therefore, that adjacent anhydroglucose units are symmetrically equivalent and, correspondingly, that alternating glycosidic linkages along the chain are symmetrically equivalent.

If the assumption concerning coincidence of the twofold screw axis and the molecular chain axis were relaxed, the diffractometric patterns would admit nonequivalence of alternate glycosidic linkages along the molecular chain, as well as the nonequivalence of adjacent anhydroglucose units. This possibility has been ignored, however, in large part because it requires expansion of the number of internal coordinates which have to be determined from the diffractometric data.

The assumptions concerning the symmetry of the unit cell noted above have been the basis of the last cycle of refinements of the structure of cellulose I. In one such refinement,[19] the forbidden reflections were simply assumed negligible, and the intensity data from *Valonia* cellulose were used to arrive at a final structure. In another study, the inadequate informational content of the diffractometric data was complemented with analyses of lattice packing energies;[26] the final structures were constrained to minimize the packing energy as well as optimizing the fit to the diffractometric data. Here the assumptions implicit in the weighting of the potential functions that are used in the energy calculations further complicate the interpretations; they can result in unacceptable hard sphere overlap of the two protons on the two carbons that anchor the glycosidic linkage. Furthermore, as has been noted by French,[27] the structures derived in these two studies, though both based on parallel chain arrangements, are nevertheless very different crystal structures. When the same convention is applied to defining the axes of the crystal lattice, the structure most favored in one analysis is strongly rejected in the other. Furthermore, neither of these is strongly favored over yet a third, antiparallel structure.[28]

It is fair to say that at the conclusion of the middle period, many of the complexities of the crystallographic problem had been explored but there was little consensus concerning the details of the structure, particularly for cellulose I. The publication during the 1970s of two crystallographic models with parallel chains was embraced by the community of research workers concerned with the biosynthesis of cellulose. However, there remained significant elements of the crystallographic data sets that could not be rationalized in terms of any of the models that were current.

In addition to the crystallographic studies focused on the structures of celluloses I and II, the application of X-ray diffractometry in investigations of cellulose revealed the occurrence of two other polymorphic forms of cellulose that are generally recognized, namely celluloses III and IV.[29] Cellulose III is most commonly prepared by treating cellulose with anhydrous ammonia, while cellulose IV is prepared by treating cellulose in glycerol at temperatures well above 200 °C. These received less attention than forms I and II, by and large because they were not encountered in any of the major commercial applications of cellulose.

3.16.2.2.2 Lateral aggregation

As noted earlier, in concluding his essay, Purves wrote that “the secondary valence forces that radiate laterally from the cellulose chain” had become the subject of investigation by many researchers. These secondary forces are now understood to be the hydrogen bonding associated with the hydroxyl groups and the van der Waals interactions involving the hydrophobic faces of the pyranose rings. Together with the distinctive conformational characteristics of the cellulose molecule itself, the secondary forces are responsible for the diversity of patterns that occur in the aggregated forms of cellulose, particularly in the native state. The diversity of these patterns had emerged from microscopic examination of the native forms of cellulose, first, at the resolution levels of optical microscopes and, later at levels made accessible by electron microscopes.

The application of electron microscopy, in particular, revealed that the aggregation of celluloses seemed to be subject to additional organizing influences at a scale that is an order of magnitude higher than that of the unit-cells of the crystallographic structure. It was recognized that this next higher level of organization, the supermolecular level, was the one at which the distinctiveness of the different native forms manifests itself. Aggregation at this level had become the subject of intensive study in two areas not unrelated to the study of cellulose; these were colloid science and the physical chemistry of macromolecules, both of which provided helpful concepts and methodologies that were welcomed by students of the nature of celluloses.

Studies of the organization of cellulose at the supermolecular level during the middle period have been the subject of a number of excellent comprehensive discussions to which the interested reader is referred. Those by Hermans[2] and by Ellefsen and Tonessen[30] focus on the physical, chemical and colloidal aspects of organization, while those by Frey-Wyssling[9] and Preston[8] address the issues from the perspective of plant biology.

A review of the research literature addressing questions of lateral organization in native cellulose suggests a continuing search for a suitable conceptual framework. Though it has not been explicitly stated, the challenge was to reconcile the fundamental premises of crystallographic models, that is, an infinite linear repetition of the unit cell in the direction of the chain axis, with the reality that at the level of 5-10 nm in lateral extension, most native celluloses were revealed to posses a fibrillar organization that is rarely linear. Rather, it most often had curvature in three dimensions usually reflecting the native morphology of the cell walls from which the cellulose was isolated. It was to bridge this gap that cellulose scientists turned their attention to the structural levels addressed within colloid science and the physical-chemical aspects of macromolecular science.

Within the context of colloid science, the concept of micellar organization had been developed to address the phenomenology of a number of classes of systems wherein substances were dispersed at a scale of 5-500 nm. Some of these systems were dispersions of particulate substances in liquid media, while others were dispersions of amphiphilic organic substances in aqueous media. They had in common the occurrence of large interfacial areas which were such that a significant fraction of the matter in the dispersed phase was at or near the interface. The term micelle was introduced to describe the individual units or domains of the dispersed phase. Since the dimensions of the micelles were usually dependent on the nature of the dispersed phase and on the microenvironment within a colloidal dispersion, the conceptual framework that was developed for the analysis of colloidal phenomena was thought to be applicable to characterization of the aggregation of cellulose.

With respect to the organization of native celluloses, particularly those from higher plant forms, the concept of micellar organization was adapted to describe organization in the two directions perpendicular to the chain axis. The organization of the fibrils seemed to be consistently the same for samples of cellulose isolated from a particular species. Thus, it appeared that the manner of assembly of the cellulose chains resulted in a pattern of aggregation that is determined by the microenvironment of the cell wall during biogenesis. A number of speculative models of the organization of cellulose chains within the fibrils were put forth; the majority of them were directed at devising patterns that were consistent with both the organization of chains in the unit cell and with the less linear organization at the next higher scale of organization. Most of the models included a component of disordered cellulose, in part to account for the difference between the geometry of the unit cells and that of the fibrils.

The models can be regarded as in two categories in that some of them focus on micellar organization in the lateral dimension while the others also include the longitudinal dimension. The first attempt to replicate a cross-section of the fibrils did not address the third dimension, which is parallel to the chain axis. The micellar organization is envisioned in two dimensions. Many studies were devoted to deriving estimates of the lateral dimensions of the crystalline micellar domains in native celluloses on the basis of the broadening of the primary reflections in X-ray diffraction patterns. In most instances the dimensions derived from analysis of the line shapes in the diffraction patterns were of the same magnitude as those derived from electron microscopic observations of the most finely dispersed forms of the native celluloses obtained through sonication of the aggregates.

The second category of models was based on a longitudinal view of the fibrils. The manner in which these models reconciled the geometry of the unit cell with that of the fibrils was based on introduction of the concept of micellization in the longitudinal dimension. Here, the micellar dimensions adopted in the direction of the chains were assumed to be much larger than in the lateral dimensions. In this context, the fibrils were viewed as consisting of domains of extended order connected by disordered domains. The disordered domains were often characterized as amorphous. As such, they do not impose a geometric constraint at the next higher level of organization, thus allowing for the differences in geometry between the fibrillar level and the unit cell level. The term

"fringed micelle" was introduced to describe the individual ordered domains surrounded and connected by disordered domains.

The concept of fringed micelle was also introduced in descriptions of the structures of cellulose derivatives and synthetic polymers capable of partial crystallization. Such polymers, usually described as semicrystalline, aggregate as microcrystalline domains embedded within domains in which the molecules are less well ordered. In these systems the dimensions of the crystalline domains are dependent on the conditions prevailing during the process of aggregation and in this respect are not unlike colloidal systems. The crystalline domains in these polymers were thus also viewed as micellar in nature. These crystalline domains have many similarities to the crystalline domains in both regenerated and mercerized cellulose. Because of the relationship between native and mercerized cellulose, many of these similarities were also ascribed to native cellulose; the degree to which this is appropriate remains in serious question.

3.16.2.2.3 Infrared spectroscopic studies

The contribution of infrared spectroscopy to studies of structure in the middle period was complementary to the crystallographic observations; it has been reviewed by Blackwell and Marchessault.[31] The information derived was in two areas. In studies where the dichroism of infrared absorption of oriented specimens was measured, proposals of particular hydrogen-bonding schemes within the crystalline domains were made. The differences between the spectra of celluloses I and II were explained in terms of differences in the packing of molecular chains and associated variations in the hydrogen-bonding patterns.

The second and perhaps more significant arena in which infrared spectroscopy provided important new insights was in the finding that, based on their absorption in the OH stretching region, native celluloses could be classified into two categories.[32] The first, consisting of algal and bacterial celluloses, was identified as group I_A, while the second, including celluloses from such standards as cotton and ramie, which are representative of higher plant celluloses, were identified as I_B. Though this finding pointed to a higher degree of complexity in the structure of native celluloses, the crystallographic studies continued to seek a unique unit cell assumed characteristic of all native celluloses.

Infrared absorption measurements were also used to explore the partition between crystalline and noncrystalline domains through the use of exposure to D_2O vapor to determine the accessible portion of the cellulose which was deemed noncrystalline.[33] IR spectroscopy was used in yet another more practical application as the basis of a crystallinity index by Nelson and O'Connor.[34,35]

Since closure of the middle period, the development of new information from analyses of both the Raman and solid-state ^{13}C NMR spectra have resulted in reassessments of the crystallographic problem; these will be considered in a subsequent section (3.16.3.2.1).

3.16.2.3 The current period (1975-Present)

Studies during the current period will be discussed in greater detail in the following sections. It is useful, however, to place them in perspective in relation to the work reported in the middle period. Much of the work in the middle period incorporated a criterion long honored in scientific studies, namely, William of Ockham's principle of economy, which requires that the most simple hypothesis consistent with observations should always be adopted. This is perhaps best represented by the crystallographic studies where the most simple structures that account for the majority of the diffractometric data were the ones adopted. In biological systems, where individuality may be reflected in very subtle differences, such as those noted in the infrared spectra, this approach can divert attention from important observations. Such has been the case in the study of native celluloses.

The beginning of the current period can be associated with the first application of some new spectroscopic methods to the characterization of cellulose, and the findings that the observed spectra could not be reconciled with the crystal structures derived from the diffractometric data. These led to reassessment of the degree of precision that can be ascribed to the structures based on the limited amount of diffractometric data, and the questions outlined above concerning the degree to which the reported structures are constrained by and in fact reflect assumptions with respect to symmetry introduced at the outset.

The first questions arose on the basis of Raman spectral observations that pointed to distinctive differences between the skeletal conformations of celluloses I and II. The differences could not be

accounted for on the basis of the conformations prescribed by the twofold screw axes of symmetry coincident with the chains in both celluloses I and II. The first efforts to rationalize the Raman spectra suggested that the true conformations may be represented by small departures from the twofold screw axis, with the departures from the twofold screw axis symmetry being a right-handed departure in the instance of cellulose I and a left-handed departure in the case of cellulose II. More careful consideration of this proposal led to the conclusion that it would not be consistent with the rapidity of the transformation from cellulose I to cellulose II in the course of mercerization of native celluloses.

Further examination of the Raman spectra and comparisons with the spectra of dimeric structures led to an alternative interpretation of the departures from twofold screw axis symmetry. It was proposed that the departures were small alternating right-handed and left-handed departures centered at dihedral angles of alternating glycosidic linkages along the chains. Such small departures from the twofold screw axis symmetry would result in diffraction patterns that are predominantly consistent with the $P2_1$ the selection rules, with disallowed reflections appearing quite weak under most conditions of observation. An important implication of this proposal, however, is that the basic repeat unit in the crystal structures must be viewed as the anhydrocellobiose unit rather than the anhydroglucose unit; all earlier studies had been based on the latter premise.

The proposal that anhydrocellobiose is the basic repeat unit of physical structure and the implicit corollary that alternating glycosidic linkages are not symmetrically equivalent were one of the motivations for turning to the method of solid-state ^{13}C NMR, which was then becoming available for application to polymeric systems. Although the solid-state ^{13}C NMR spectra could not be interpreted as conclusive evidence of nonequivalence of alternating glycosidic linkages, they were consistent with such an interpretation. But introduction of the use of solid-state ^{13}C NMR had even more profound implications with respect to a deeper understanding of the diversity of native celluloses.

Based on studies of a wide range of native celluloses from plant sources, it was proposed that the crystalline domains of all of these forms were composites of two different forms. I_α and I_b, that occur in proportions that are distinctive to the particular species producing the cellulose. Though the proposal of alternating glycosidic linkages remains controversial, the finding that native celluloses, are composites of the I_α and I_b forms has been widely accepted and has become the basis for many investigations directed at more detailed characterization of native celluloses.

This discovery of the I_α and I_b forms provided a deeper understanding of the source of the categorization developed earlier on the basis of infrared spectral observations in the OH stretching region. More fundamental perhaps, it also provided a basis for resolution of a paradox first posed by Cross and Bevan[4] almost a century ago when the crystalline aspects of the character of cellulose were first recognized on the basis of its birefringence. They had noted, "The root idea of crystallography is identical invariability, (while) the root idea of the world of living matter is essential individual variation," clearly recognizing that the uniformity of structure characteristic of crystalline order may be inconsistent with the diversity of celluloses from different biological sources. The uniqueness of the blend or integration of the I_α and I_b forms in the cellulose of a particular organism thus provides the basis for the "essential individual variation" alluded to by Cross and Bevan.[4]

The discovery of the occurrence of the two forms of cellulose, I_α and I_b became the point of departure for studies that led to other important advances during this period. Some involved further refinements of the solid-state ^{13}C NMR spectral analysis. Others were in the area of electron microscopic characterization of cellulose. The lattice imaging technique provided important evidence of the homogeneity of order within crystalline domains that were composites of the I_α and I_b forms of cellulose. Special methods for staining the reducing end groups of cellulosic chains addressed the question of parallel vs. antiparallael alignment of the chains in the crystalline domains. Greater sophistication in application of the methods of electron diffraction has opened up the possibility of a better understanding of the nature of the two forms of native cellulose, and Raman spectral evidence pointed to the similarities and differences between them.

The progress in understanding the nature of structure and aggregation in celluloses during the current period will be described and discussed in greater detail in the following section.

3.16.3 STRUCTURES

As noted earlier. the beginning of the current period of studies on the structure of cellulose was marked by the reintroduction of unit cell models based on parallel alignment of the cellulose molecular chains.[19,26] not unlike those abandoned by Meyer and Misch[24] in the 1930s, but also

incorporating bending of the glycosidic linkage to allow the intramolecular hydrogen bond, as suggested by Hermanns.[2] The new models were not consistent with each other, however, apart from the fact that both were based on parallel alignment of the cellulose chains. As French[27] pointed out, they were also not strongly preferred over an antiparallel structure. In the analysis by French,[27] it was recognized that the source of the inconsistency was not so much that the different laboratories were using different computational approaches as it was that the different diffractometric data sets were gathered from different samples and represented different intensities for the same reflections. All of these studies were undertaken before the variability of the crystalline forms of native celluloses was revealed through high-resolution solid-state ^{13}C NMR investigations.

The new crystallographic models also remained in question because the analyses on which they were based incorporated a level of symmetry in the unit cell that was inconsistent with some of the diffractometeric data. Some of the reflections that are consistently observed in electron diffraction patterns and are disallowed by the selection rules for the space group $P2_1$[20] were ignored in these crystallographic analyses. In addition to the disallowed reflections in the electron diffraction patterns that placed the crystallographic models in question, new spectral evidence was developed pointing to the need for further refinement of the structural models, particularly for native celluloses. The models derived from the crystallographic studies could not rationalize many features of the spectral data known to be quite sensitive to structural variations.

On the other hand, electron microscopic studies based on new staining techniques, specific to the reducing end groups of the polysaccharides, confirmed the parallel alignment of molecular chains within the microfibrils in native celluloses. These findings were confirmed further by the manifestation, at the electron microscopic level, of the action of cellulases specific to the nonreducing end group; they were clearly active at only one end of each microfibril. These observations were regarded as confirmation of the most recent crystallographic models. The remaining questions at the time, therefore, were concerned with the degree to which the symmetry of space group $P2_1$ is consistent with the other structure-sensitive observations.

It is well to revisit the issue of levels of structure at this point and clarify the levels at which the different investigative methods are most sensitive. The crystallographic models, which represent coordinates of the atoms in the unit cell, represent the most complete possible specification of structure because they include primary, secondary, and tertiary structures. Indeed, crystallographic studies of the monosaccharides and related structures provide the basis for much information concerning bond lengths and bond angles, as well as conformations in saccharide structures. As noted in the previous section, however, for polymeric systems the diffractometric data is far more limited than for a single crystal of a low molecular weight compound, so that diffraction data from a polymer must be complemented by information from other structure-sensitive methods. An acceptable model must rationalize not only the diffractometric data. which for cellulose is quite limited in comparison to the number of coordinates that must be specified in a definition of the unit cell, but it must also be such that it can be reconciled with information derived from measurements known to be sensitive to other levels of structure.

The new spectral evidence that must be rationalized by any acceptable structure came from two methodologies that are most sensitive to structure at the secondary and tertiary levels, These are Raman spectroscopy and solid-state ^{13}C NMR spectroscopy, both of which were applied to cellulosic samples for the first time during the 1970s. The exploration of spectra measurable by these two methods can provide significant information concerning both secondary and tertiary structures in the solid state. Because the spectral features observed are also sensitive to the molecular environment, they are influenced by the degree of symmetry of the aggregated state. Hence, they provide another avenue for exploration of the applicability of the symmetry of space group $P2_1$ to the structures of the solid state.

Though Raman spectroscopy was first developed in the 1930s, it could not be applied to optically heterogeneous samples like cellulose until lasers could be used as sources for excitation and until new generations of monochromators and detectors became available in the early 1970s. High-resolution solid-state ^{13}C NMR spectroscopy also first became available in the 1970s, and one of its most important areas of application has been in investigations of polymeric systems. Though these two methodologies were, in the first instances, applied separately to cellulose, they were eventually found to be complementary and facilitated development of the foundation for continuing investigations of the nature of native celluloses. The following sections will reflect this evolution. They begin with the early applications of Raman spectroscopy that focused on issues of molecular conformation, and with the high-resolution solid-state ^{13}C NM R investigations that led to discovery of the I_α and I_b forms of native celluloses. They will be followed by a discussion of more recent

applications of these methods together with infrared spectroscopy to further explore the nature of this duality in native celluloses.

In addition to the spectroscopic investigations, the use of electron microscopy to explore the structures of native cellulose benefited from the development and application of new techniques that were adapted to investigations of cellulose. The application of stains specific to the reducing end groups of cellulose noted above allowed the issues of parallel and antiparallel chain orientation to be explored in new ways. The method of lattice imaging was also applied to assessment of the homogeneity of microfibrils of native cellulose. Finally, electron diffraction from microdomains in individual microfibrils became an important tool in exploring the nature of I_α and I_b duality. A discussion of these developments will follow discussion of the spectral studies.

Yet another class of investigations that can shed light on the nature of cellulose are based on computational analysis of molecular models using the programs that have become increasingly available in recent decades. Such studies are frequently limited with respect to development of conclusive results because of assumptions that are implicitly or explicitly incorporated in the computational programs in order to make them more tractable, or because of constraints of symmetry that are frequently imposed in the course of the formulation of the models of cellulose that are chosen for analysis. However, they remain of value because they can provide insights concerning the classes of molecular interactions that are likely to be dominant in determining the structures of cellulose and the nature of its interactions with other constituents that occur with it in its native state. A brief overview of some work in this area will be presented.

To complement these studies of cellulose, some analyses of the structures of the oligomers have also been considered. Though these pertain primarily to the structure of cellulose II, they provide some insights with respect to the influence of packing in a lattice on a sequence of anhydroglucose units. As Kakudo and Kasai[11] have noted, the structures of the oligomers can provide important insights concerning the manner of organization of the monomers in the polymeric chains.

3.16.3.1 Spectroscopy

Spectroscopic observations are important in structural investigations because they can provide information which is complementary to that derived from diffractometric data. Such information is particularly important in instances such as that of cellulose, where the diffractometric data alone cannot provide sufficient information to serve as the basis for a structure determination. While the information derived from spectra is not directly related to the absolute values of the coordinates of atoms within the unit cell, it is particularly sensitive to the values of the internal coordinates, which in turn define molecular structures and conformations. In addition, in Raman spectroscopy and infrared spectroscopy, distinctive features associated with the crystal structure can provide critical information. In the case of solid-state ^{13}C NMR, the spectral information is also sensitive to nonequivalences in the environments of chemically equivalent atoms within adjacent monomeric repeat units of the primary structure. Thus, the spectral information provides a basis for testing the degrees of equivalence of structures. Very often also, specific spectral features can be identified with particular functional groups defined by distinctive sets of internal coordinates.

While spectral analyses cannot provide direct information concerning the structures, they establish criteria that any structure must meet to be regarded as an adequate model. The information from spectroscopic studies represents one of the major portions of the phenomenology that any acceptable structural model must rationalize. Reconciliation of the spectral observations with the model derived from crystallographic investigations provides tests of consistency of the proposed structures in the sense set forth by Kakudo and Kasai.[11]

3.16.3.1.1 Raman spectroscopic studies and conformational questions

Raman spectroscopy is the common alternative to infrared spectroscopy for investigating molecular vibrational states and vibrational spectra. It has enjoyed a significant revival since the development of laser sources for excitation of the spectra. In laser-excited Raman spectroscopy, a sample is exposed to the monochromatic laser beam and the scattered light is analyzed. While most of the scattering consists of Rayleigh scattered photons of the same frequency as that of the laser, a small fraction of the scattered photons undergo an energy exchange with the scattering molecules and are shifted in frequency relative to that of the exciting source; these are the Raman scattered photons. The magnitudes of the shifts in the frequencies of the Raman scattered photons correspond to the

vibrational frequencies of the molecules in the sample under investigation. Thus, the informrtion contained within a Raman spectrum, with respect to the vibrational frequencies of the scattering molecules, is similar to that contained in an infrared spectrum.

Both Raman and infrared spectroscopy provide information about chemical functionality, molecular conformation, and hydrogen bonding. Raman spectroscopy, however, has some important advantages in the study of biological materials. The key advantage arises from the different bases for activity of molecular vibrations in Raman and infrared spectra That is, whereas activity in the infrared region requires finite transition moments involving the permanent dipoles of the bonds undergoing vibrations, activity in the Raman spectrum requires finite transition moments involving the polarizabilities of the bonds. Thus, in infrared spectroscopy the exchange of energy between the molecules and the exciting field is dependent on the presence of an oscillating permanent dipole. In Raman spectroscopy, in contrast, the exciting field induces a dipole moment in the molecule and the induced moment then becomes the basis for exchange of energy with the exciting field. It is useful in this context to view bonds in terms of Pauling's classification along a scale between the two extremes of polar and covalent.[36] Bonds that are highly polar and possess relatively high dipole moments tend, when they undergo vibrational transitions, to result in bands that are intense in the infrared and relatively weak in Raman spectra. Conversely, bonds that are primarily covalent in character and have a relatively high polarizability generally result in bands that are intense in the Raman spectra but are relatively weak in the infrared. This is perhaps best illustrated by the fact that O_2 and N_2, which are homonuclear and without permanent dipoles, have very intense Raman spectra though they are inactive in infrared absorption, while H_2O, with a high permanent dipole moment, is a very strong absorber in the infrared but a very weak Raman scatterer. With respect to cellulose, the OH groups of cellulose and those of adsorbed water are dominant in many of the spectral features in infrared spectra. In contrast, the skeletal C—C bonds and the C—H bonds dominate the Raman spectra. A further simplification in the Raman spectra results from the circumstance that the selection rules forbidding activity of overtone and combination bands are more rigidly adhered to than is the case in infrared spectra so that the bands observed in Raman spectra are usually confined to the fundamental modes of the molecules under investigation.[37]

Raman spectroscopy possesses another important advantage in comparison to infrared spectroscopy in the study of biological systems; optical heterogeneities do not present serious problems. In infrared spectroscopy, the key observable is the degree of attenuation of an incident infrared beam in comparison to a reference beam. Thus, if any process other than absorption can result in attenuation of the infrared beam, interpretation of the spectra can be complicated. Since the refractive index of a substance can undergo large excursions in the neighborhood of strong absorption bands, the Rayleigh scattering losses will vary with frequency in the infrared absorption region of the spectrum, and they can cause anomalous features in the infrared spectra. Furthermore, in the case of cellulosic materials, the optical heterogeneities generally have dimensions in the same range as the wavelengths of the infrared absorption in the region of the fundamental vibrations. In consequence, Rayleigh scattering of the incident infrared beam can be significant and frequency dependent. In Raman spectroscopy, in contrast, variations in the refractive index do not present a difficulty since the excitation frequency and the frequencies of the Raman scattered photons are far removed from any absorption bands. It is therefore easier to record meaningful Raman spectra from samples such as cellulose, even though they may cause a high level of Rayleigh scattering.

In the context of studies on the structure of cellulose, the key advantage of Raman spectroscopy is the degree of its sensitivity to the skeletal vibrations of the cellulose molecule, with the mode of packing in the lattice having only secondary effects. This sensitivity is a consequence of the reality that most of the skeletal bonds are C—C and C—O bonds, both of which have relatively high polarizabilities and, hence, high Raman scattering coefficients. The minimal contribution of packing effects arises from the low scattering coefficients of the highly polar OH groups, which are the functionalities that are most directly involved in intermolecular associations. The result is that intramolecular variations such as changes in internal coordinates have a significantly greater influence on the Raman spectra than variations in intermolecular associations.

These considerations were paramount in the first detailed examination and comparison of the Raman spectra of celluloses I and II;[38] the spectra are shown in Figure 1. It was concluded that the differences between the spectra, particularly in the low frequency region, could not be accounted for in terms of chains possessing the same conformation but packed in different ways in the different lattices. As noted earlier, that had become the accepted rationalization of the differences between celluloses I and II developed from diffractometric studies of these two most common allomorphs. The analyses of the Raman spectra led to the proposal that two different stable conformations of the cellulose chains occur in the different allomorphs.

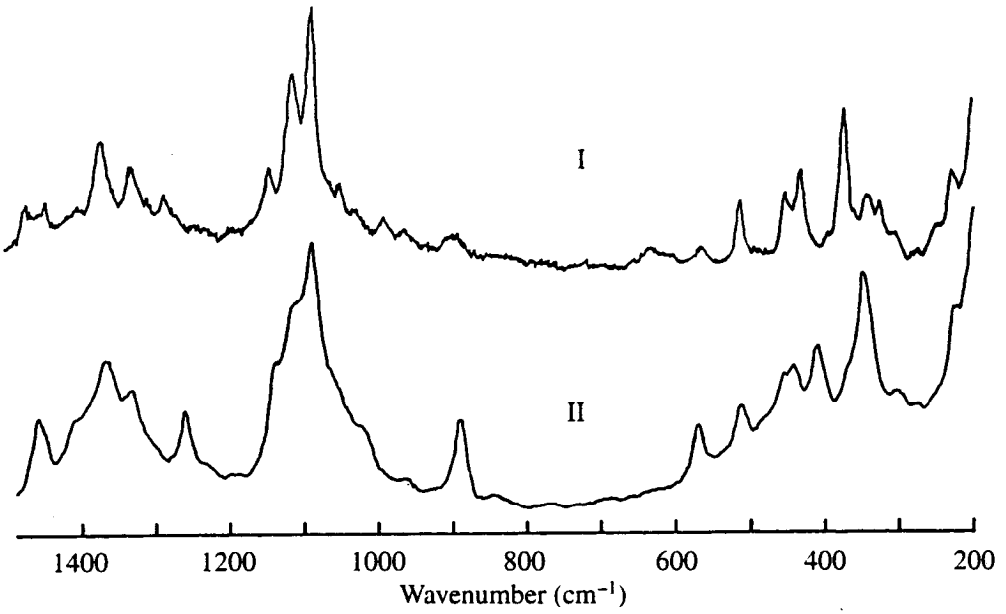


Figure 1 Raman spectra of high crystallinity celluloses I and II.[38]

Raman spectra of native cellulose had previously been reported and discussed by analogy with the spectra of glucose,[39] but the significant differences between celluloses I and II had not been addressed, nor did the assignments consider the complexity of coupling between the different fundamental modes. To establish a basis for assessing the differences between celluloses I and II, Atalla and co-workers undertook an extensive series of studies of model compounds of increasing complexity.[40-47] The studies included comprehensive normal coordinate analyses of the molecular vibrations of each of the groups of model compounds based on complementary infrared and Raman spectra. The objective of these analyses was to establish the degree to which the different classes of vibrational motions contribute to the spectral features in the different regions of the spectrum. Such a comprehensive approach was necessary because the skeletal bond systems that occur in the structures of carbohydrates are made up predominantly of C—C and C—O bonds, which possess similar reduced masses and vibrational force constants and, hence, have very similar vibrational frequencies. In consequence a high degree of coupling occurs between the vibrations, with the result that very few of the vibrational modes are localized within specific bonds or functional groups. Thus, the traditional group frequency approach common in the assignment of infrared and Raman spectra is of very limited use except in the case of vibrations localized in the bonds of hydrogen atoms bonded to much heavier atoms such as oxygen or carbon. On the other hand, the normal coordinate analyses allow identification of the degree to which the vibrations of each of the internal coordinates contributes to each of the observed bands. Since the coupling of the vibrations is very sensitive to changes in the bond angles and the dihedral angles associated with the bonds whose vibrations are coupled, the normal coordinate analyses allow detailed and systematic exploration of the effects of differences in skeletal conformations on the bands associated with particular vibrations.

The compounds chosen to establish the foundations for analysis of the spectra of cellulose began with the 1.5-anhydropentitols[40,41] and continued with the pentitols and erythritol,[42] the pentoses,[43] the inositols,[44,45] and three hexoses including glucose, mannose and galactose.[46,47] For each group of compounds, the force fields were refined against the observed frequencies of a subset of the group until a satisfactory fit was obtained. The force fields were validated by successful prediction of the spectra of compounds not included in the refinements. Furthermore, the calculated potential energy distributions were reasonable in comparison with the group frequency literature on carbohydrates. The analyses were thus successful in developing a physically meaningful force field specifically tailored to the carbohydrates. The hexose field was then used to extend the normal coordinate method to the cello-oligodextrins,[48] the vibrational spectra of which more closely resemble the spectra of cellulose. Because the number of vibrational degrees of freedom greatly exceeds the number of bands that can be resolved in the spectra of the cello-oligodextrins, it was neither possible nor meaningful to refine the force constants further in this context. The distribution of calculated frequencies was in qualitative agreement with the observed spectra. The force field derived from the

hexoses thus appears to provide a good basis for interpreting the spectra of the cello-oligodextrins and cellulose.

With respect to the question concerning the conformations of celluloses I and II, it is useful to consider first some of the pertinent information developed from the normal coordinate analyses, particularly with respect to the classes of molecular motions associated with the different spectral features. The region below 1500 cm^{-1} was the primary focus of the early exploration because the intense bands clustered at about 2900 cm^{-1} can be identified with the C—H stretching vibrations and the region beyond 3000 cm^{-1} is clearly associated with the O—H stretching vibrations. In addition to the C—H and O—H stretching vibrations, the internal deformation of the methylene group on C-6 is the only vibration which closely approximates a group or local mode in the usual sense implicit in discussions of assignments of vibrational spectra; the HCH bending vibration usually occurs above 1450 cm^{-1}. In all other bands at frequencies below 1450 cm^{-1}, the normal coordinate analysis indicated that the vibrations are so highly coupled that, in most instances, no single internal coordinate contributes more than 20% of the potential energy change associated with any particular frequency, though in a few instances contributions were as high as 40%. Thus, the traditional group frequency approach to assignment of vibrational spectra, which is based on the concept of local modes, is generally not applicable in this region in the spectra of saccharides. It is necessary instead to focus on the classes of internal motions that are associated with the different frequency ranges and to interpret the spectra in terms of the influence that variations in the internal coordinates can have on the coupling between different types of vibrational deformations.

For analysis of the spectra of celluloses, it is possible to classify the groups of features in the different spectral regions in terms of the types of internal deformations that make their maximum contributions to bands in those regions. The bands between 1200 and 1450 cm^{-1} are due to modes involving considerable coupling between methine bending, methylene rocking and wagging, and COH in-plane bending motions; these are angle bending coordinates involving one bond to a hydrogen atom and the other to a heavy atom. Significant contributions from ring stretching begin below 1200 cm^{-1} and these modes, together with C—O stretching motions, dominate between 950 and 1150 cm^{-1}. Below 950 cm^{-1}, angle bending coordinates involving heavy atoms only (i.e., CCC, COC, OCC, OCO) begin to contribute, though ring and C—O stretches and the external bending modes of the methylene group may be major components as well. The region between 400 and 700 cm^{-1} is dominated by the heavy atom bending, both C—O and ring modes, although some ring stretching coordinates still make minor contributions. In some instances, O—H out-of-plane bending motions may make minor contributions in this region as well. Between 300 and 400 cm^{-1}, the ring torsions make some contributions, and below 300 cm^{-1} they generally dominate.

In addition to the above generalized categorization concerning modes that occur in one or another of the model compound systems used in the normal coordinate analyses, the spectrum of cellulose will have contributions due to modes centered at the glycosidic linkage. The computations based on the cello-oligodextrins indicate that these modes are strongly coupled with modes involving similar coordinates in the adjacent anhydroglucose rings. The contributions of the different classes of internal coordinates to the different bands are presented in greater detail elsewhere.[49]

As noted above and shown in Figure 1, the differences between the Raman spectra of celluloses I and II are quite significant, particularly in the region of the skeletal bending modes of vibration. In the region above 800 cm^{-1} the differences are most obvious with respect to the relative intensities of the bands and the broadening of some of the bands upon conversion from cellulose I to cellulose II. In the region below 700 cm^{-1}, in contrast, the main features are quite different in the two spectra; these differences are even more evident in the spectra of single fibers which will be presented later.

In the analyses of the spectra of model compounds, changes of the magnitude indicated in Figure 1 were associated exclusively with the occurrence of differences in conformations. It seemed very probable therefore that the differences between the spectra of celluloses I and II reflect a change in molecular conformation accompanying the transition from one form to the other. Since the basic ring structure is not expected to change,[50] it would appear that variations in the dihedral angles at the glycosidic linkages provide the only opportunity for conformational differences. Because of the controversy surrounding similar conclusions based on crystallographic studies carried out in the early 1960s,[17,18] a number of experimental and theoretical avenues for validating this interpretation were pursued.

The first consideration was whether a multiplicity of stable conformations is consistent with the results of conformational energy calculations that were available at the time.[50,51] In both studies, the potential energy surfaces were found to possess multiple minima. When the additional constraint of a repeat length of approximately 0.515 nm per anhydroglucose unit was added, two minima representing both left and right-handed departures from the twofold helix appeared to be likely loci

of the stable conformations. It was noted in this context that these two minima were close to the positions of the dihedral angles of the glycosidic linkages in cellobiose and methyl-*b*-cellobioside, respectively, as these were determined from crystallographic studies.[52,53]

Next, inquiry was made into the degree to which changes in the dihedral angles about the bonds in the glycosidic linkage could influence the modes of vibration responsible for the spectral features in the different regions of the spectra. Two approaches were adopted for this purpose. The first was based on examining the Raman spectra of polysaccharide polymers and oligomers that were known to occur in different conformations. The second was a theoretical one based on an adaptation of the matrix perturbation treatment used by Wilson *et al.* [54] to discuss the effects of isotopic substitution on infrared and Raman spectra.

The polysaccharide systems chosen for investigation were among those most closely related to cellulose in the sense that they are the α-(1-4)-linked polymers and oligomers of anhydroglucose. They included amylose and two of its cyclic oligomers, with primary emphasis on the latter, the α– and *b* -Schardinger dextrins, often also known as cyclohexa- and cyclohepta-amylose. The structures of the two oligomers differ in that the values of the dihedral angles about the bonds of the glycosidic linkages have to change to accommodate the different number of monomer units. Comparison of the Raman spectra of the cyclic dextrins showed that the differences between them were quite minor in the regions above 800 cm^{-1}, but they were quite significant in the lower frequency region dominated by the skeletal bending and torsional modes. The differences were similar in kind and distribution to the differences between celluloses I and II. It was also noted that in earlier studies of the Raman spectra of amylose,[55] it had been observed that forms V_a and V_h, which are very similar in conformation but had different levels of hydration, had almost identical spectra. In contrast, form B, which is known to have a distinctly different helix period, was found to have a spectrum that differs from those of forms V_a and V_h in a manner approximating the differences between the two cyclic oligodextrins. Taken together, the observations of the Raman spectra of the amyloses support the interpretation of the differences between the Raman spectra of celluloses I and II as pointing to differences in the chain conformations localized at the glycosidic linkages.

In the theoretical analysis, the method of Wilson *et a1.* [54] was adapted to explore the consequences of variations in the dihedral angles about the bonds in the glycosidic linkage. Changes in the dihedral angles were found to influence skeletal stretching and bending modes primarily through changes in some of the corresponding off-diagonal terms in the kinetic energy matrix **G**. Examination of the general expressions for these terms[56] reveals that only one of the four classes of terms that influence stretching is sensitive to the value of the dihedral angle, and it is a class representing stretch-bend interactions. The interactions influencing the bending modes, in contrast, are more sensitive to the dihedral angles. Among these, three of the four classes of bend-bend interactions change with the dihedral angles; these are in addition to the stretch-bend interaction cited earlier, which would also influence the bending modes. Finally, the majority of the terms involving torsional coordinates are sensitive to variation in the dihedral angles. These considerations led to the conclusion that the skeletal bending and torsional modes are altered to a greater degree than the skeletal stretching modes when the dihedral angles associated with the glycosidic linkage undergo variations. When translated to spectral features in the Raman spectra, these observations point to major differences in the low-frequency region below 700 cm^{-1}, and minor ones in the fingerprint region between 900 and 1500 cm^{-1}. These are indeed precisely the types of differences observed in comparisons of the spectra of celluloses I and II.

One final consideration that was addressed is the possibility that rotations of the primary alcohol group at C-6 could account for the spectral differences seen in the spectra of celluloses I and II and in the spectra of the amyloses. The normal coordinate analyses of the hexoses showed that rotations about the C-5—C-6 bond can result in minor variations in the region below 600 cm^{-1} but that the major impact of such rotations is expected in the spectral region above 700 cm^{-1}.[46,47] With all of the above considerations in mind, it became clear that the only plausible rationalization of the differences between the Raman spectra of celluloses I and II had to be based on the possibility that differences between the skeletal conformations were the key.

The first effort to rationalize differences in conformation was based on the results of the conformational energy mappings that were available at the time.[50,51] The key points derived from those analyses, which have been confirmed by additional studies,[57,58] were that the two energy minima associated with variations in the dihedral angles of the glycosidic linkage correspond to relatively small left- and right-handed departures from glycosidic linkage conformations that are consistent with twofold helical symmetry. The minima also represented values of the dihedral angles that were very similar to those reported for cellobiose and methyl *b*-cellobioside on the basis of crystallographic

analyses.[52,53] The relationship between the different conformations is represented in Figure 2, which was adapted by Atalla[59] from a diagram first presented by Reese and Skerret.[50] It is a y f map presenting different categories of information concerning the conformation of the anhydrocellobiose unit as a function of the values of the two dihedral angles about the bonds in the glycosidic linkage. y is defined as the dihedral angle about the bond between C-4 and the glycosidic linkage oxygen and f as the dihedral angle about the bond between C-l and the glycosidic linkage oxygen. The parallel lines indicated by n = 3(L), 2, and 3(R) represent values of the dihedral angles that are consistent with a left-handed threefold helical conformation, a twofold helical conformation, and a right-handed threefold helical conformation, respectively; a twofold helical conformation inherently does not have a handedness to it. The dashed contours represent conformations that have the indicated repeat period per anhydroglucose unit; the innermost represents a period of 5.25 Å, corresponding to 10.5 Å, per anhydrocellobiose unit. The two dotted lines indicate conformations corresponding to values of 2.5 and 2.8 Å, for the distance between the two oxygen atoms anchoring the intramolecular hydrogen bond between the C-3 hydroxyl group of one anhydroglucose unit and the ring oxygen of the adjacent unit; the values bracket the range wherein hydrogen bonds are regarded as strong. The two domains defined by solid lines on either side of the twofold helix line (n = 2) represent the potential energy minima calculated by Rees and Skerret for the different conformations of cellobiose. Finally, the points marked by J and W represent the structure of cellobiose determined by Chu and Jeffry[52] and the structure of methyl-*b*-cellobioside determined by Ham and Williams.[53] The key point to be kept in mind in relation to this diagram is that structures along the twofold helix line and with a repeat period of 10.3 Å, per anhydrocellobiose unit possess an unacceptable degree of overlap between the van der Waals radii of the protons on either side of the glycosidic linkage.

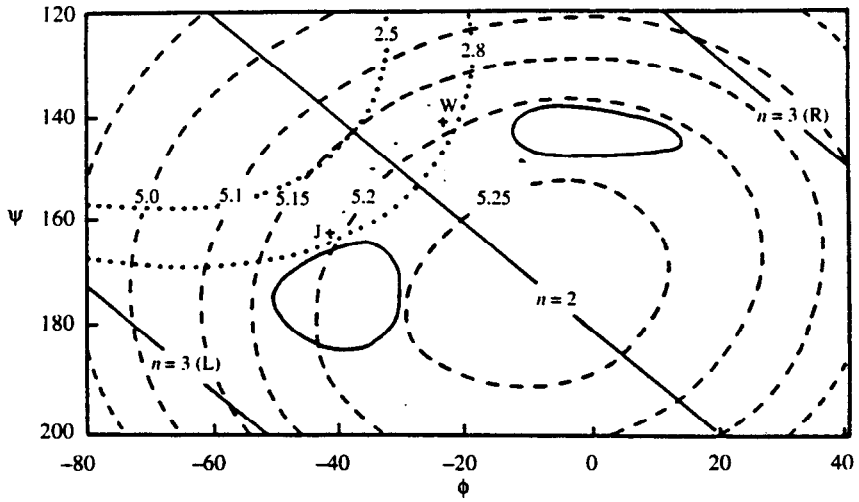


Figure 2 ψ/ϕ map. (— — —) Loci of structures with constant anhydroglucose repeat periods; (· · ·) loci of structures of constant intramolecular hydrogen bond O—O distances; (———) contours of potential energy minima based on nonbonded interactions in cellobiose. J, cellobiose; W, β-methylcellobioside. n = 2, the twofold helix line; n = 3 the threefold helix lines, (R) right-handed, (L) left-handed. The Meyer–Misch structure is at ψ = 180, ϕ = 0.[59]

Taking into consideration that it had earlier been speculated that the diffractometric data were consistent with a helix of much longer period that may be degenerate with the twofold helix[7] and that it was commonly assumed in polymer crystallography that chemically equivalent units were symmetrically equivalent,[3] Atalla proposed that the structure of cellulose I may be represented as a right-handed helix of long period with dihedral values similar to those in methyl-*b*-cellobioside and that the structure of cellulose II may be represented as a left-handed helix of long period with dihedral values similar to those of cellobiose.[60] Further observations and considerations led to revision of this proposal both with respect to the equivalence of successive units in the polymeric chains and with respect to the nature of the departures from twofold helix symmetry.

The first revision followed from observations and comparisons of the Raman spectra of cellulose II and of cellobiose in the O—H stretching region.[59] The latter showed a single sharp band superimposed on a broader background, and the band was identified with the O—H stretching vibration of the isolated intramolecular hydrogen bond revealed in the crystal structure;[52] it occurs between the hydroxyl group on C-3 of the reducing anhydroglucose unit and the ring oxygen of the

nonreducing unit. The spectrum of cellulose II revealed two such sharp bands in the same region; similar bands were observed in the spectra of the cello-oligodextrins.[48] Since the frequency at which such bands occur is very sensitive to the distance between the oxygen atoms that anchor the hydrogen bond, it appeared that the structure of cellulose II must incorporate intramolecular hydrogen bonds with two distinct values of the O—O distance. This led to the proposal that successive units in the structure are not equivalent, and that, as a consequence, alternating glycosidic linkages have different sets of dihedral angles defining their coordinates.[59] Thus, the dimeric anhydrocellobiose was regarded as the repeat unit of physical structure rather than the anhydroglucose unit. These conclusions, based on the Raman spectra in the O—H region, were confirmed when the solid-state $_{13}$C NMR spectra became available[61] as splittings were observed in the resonances associated with C-l and C-4, which anchor the glycosidic linkage. The occurrence of these splittings is indicative of the presence of nonequivalent glycosidic linkages within the structure; the NMR spectra will be considered in greater detail in the following section.

The second revision, which flowed from the first one, was rooted in the recognition that when alternating glycosidic linkages are admitted as an option, and when anhydrocellobiose is viewed as the repeat unit of the structure, the alternating glycosidic linkages no longer needed to have the same sense of departure from the twofold helix. That is, it was now possible to consider structures wherein the nonequivalent glycosidic linkages are alternating left- and right-handed departures from the twofold helix. Such structures would be ribbon like and could appear to approximate the twofold helix. The possibility of such structures would resolve one objection that was raised with respect to the earlier proposal,[62] namely, that the transition from cellulose I to cellulose II would require a remarkable degree of rotation of the molecular chains about their axes as they were converted from cellulose I to cellulose II during mercerization. The proposal incorporating the second revision also has the advantage that it can be reconciled with much of the diffractometeric data. That the departures from twofold helical symmetry are relatively small may explain the weakness of the reflections that are disallowed by the selection rules of space group $P2_1$.

Based on the considerations outlined, the model that was adopted as a basis for continuing explorations of the spectra of cellulose was based on the proposal that the glycosidic linkages alternated between small left- and right-handed departures from the twofold helical conformation. Thus, the differences between the conformations of celluloses I and II now had to be understood in terms of differences in the internal organization of the anhydrocellobiose units that were the basic units of structure.[63,64]

In search of a rationalization of the changes in the internal organization of the cellobiose unit associated with the transition from cellulose I to cellulose II, Atalla drew on the analogy with the structures of cellobiose and methyl-*b*-cellobioside, which are represented in Figure 3. The methyl-*b*-cellobioside, which has values of dihedral angles corresponding to a right-handed departure from the twofold helix, also has a bifurcated intramolecular hydrogen bond in which the proton from the C-3 hydroxyl group appears to be located between the ring oxygen and the primary alcohol oxygen at C-6 of the adjacent unit. This bifurcation is in part responsible for the absence of a sharp OH band in the OH region of the spectrum of the methyl-*b*-cellobioside. Atalla suggested that such bifurcated intramolecular hydrogen bonds may occur in connection with every other glycosidic linkage in a molecule of native cellulose; these bifurcated hydrogen bonds would be associated with those glycosidic linkages that have values of dihedral angles representing right-handed departures from the twofold helix in a manner not unlike those in methyl-*b*-cellobioside. The action of mercerizing agents was seen as resulting in the disruption of the bifurcated OH bonds, thus allowing the glycosidic linkages to relax to slightly greater departure from the twofold helix.[63,64] Such an explanation would also be consistent with the observation that the two HCH bending bands in the Raman spectra of native celluloses collapse into a single band upon mercerization, suggesting a nonequivalence of the two primary alcohol groups in native cellulose and a shift closer to equivalence upon mercerization. It is also consistent with the greater splitting of the resonances associated with C-l and C-4 seen in the solid-state ^{13}C NMR spectra of cellulose II to be discussed in the following section. While the evidence supporting this proposal is strong, it is not conclusive and thus awaits further confirmation. Atalla also introduced the terms k_{I} and k_{II} to designate the conformations corresponding to celluloses I and II; the term k_{o} was introduced to describe cellulose in a disordered state.[65]

More detailed characterizations of the Raman spectra of a number of celluloses were undertaken at a later point. These are best considered, however, after presentation of some of the important conclusions derived from the CP/MAS solid-state ^{13}C NMR studies that were undertaken soon after some of the early studies of the Raman spectra.

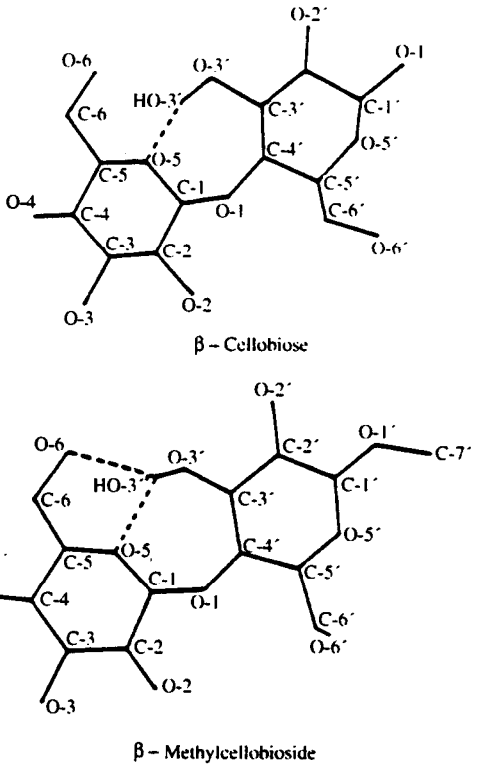


Figure 3 Structures of β-cellobiose and β-methylcellobioside.

3.16.3.1.2 Solid-state ^{13}C NMR spectroscopic studies and the two forms of native cellulose I_a and I_b

Though applied to cellulose later than Raman spectroscopy, high-resolution solid-state ^{13}C NMR has provided perhaps the most significant new insights regarding the structures of cellulose, particularly in its native state. Earlier applications of NMR to the study of solids had been based on exploration of dipolar interactions between pairs of magnetic nuclei and did not afford the possibility of acquisition of spectral information characteristic of molecular architecture in a manner that was parallel to that possible with liquid-state NMR spectra. The development of high-resolution solid-state NMR spectroscopy and its application to polymeric materials grew from complementary application of a number of procedures that had been developed in NMR spectroscopy. The first is proton carbon cross-polarization (CP) that is used to enhance sensitivity to the low-abundance ^{13}C nucleus. This was combined with high-power proton decoupling to eliminate the strong dipolar interaction between the ^{13}C nuclei and neighboring protons. Finally, the angular dependence of the chemical shift, or chemical shift anisotropy, is overcome by spinning the sample about an axis at a special angle to the direction of the magnetic field, commonly referred to as the magic angle, the procedure denoted by (MAS). The combined application of these procedures, usually designated by (CP/MAS), results in the acquisition of spectra that contain isotropic chemical shift information analogous to that obtained from liquid-state ^{13}C NMR with proton decoupling. One key difference between the two is that molecular motion in the liquid state is sufficiently rapid that chemically equivalent carbons result in single resonances, whereas in the solid-state spectra, because the molecules are immobile, the isotropic chemical shift is sensitive to the molecular environment. As a result, chemically equivalent carbons that occur in sites that are not magnetically equivalent, within an aggregated solid state, may have differences between their chemical shifts.

In summary, the most important characteristic of the spectra acquired using the (CP/MAS) ^{13}C NMR technique is that, if they are acquired under optimal conditions, they can have sufficient resolution so that chemically equivalent carbons that occur in magnetically nonequivalent sites can be distinguished. In the present context, the corresponding carbons in different anhydroglucose units would be regarded as chemically equivalent. If they are not also symmetrically equivalent, that is, if they occur in different environments or if the anhydroglucose rings possess different conformations, within the rings, at the glycosidic linkage, or at the primary alcohol group, the carbons will not have magnetically equivalent environments and will therefore result in distinctive resonances in the NMR spectrum. The fundamental challenge in the application of this method is to

achieve a level of resolution sufficient to distinguish nonequivalences between chemically equivalent carbons, because the magnetic nonequivalence can result in variations in the chemical shift that are small relative to the shifts determined by the primary chemical bonding pattern.

Another important feature of the (CP/MAS) ^{13}C NMR technique is that for a system such as cellulose, which consists of rather rigid hydrogen-bonded molecules and in which all carbons have directly bonded protons, the relative intensities of the resonances are expected to correspond to the proportion of the particular carbons giving rise to them. Thus, the intensities arising from each of the six carbons in the anhydroglucose ring are expected to be equal. This is an important characteristic that is central to the analysis and interpretation of the information contained within the spectra.

The first applications of the new technique to cellulose[61,66] demonstrated resolution of multiple resonances for some of the chemically equivalent carbons in the anhydroglucose units. It became clear that rationalization of the spectra that were observed would provide valuable additional information concerning the structure of the celluloses investigated. The first step in such a rationalization was the assignment of the resonances which appear in the spectra. The assignments, which have been discussed in a number of reports,[61,66-71] were based on comparisons with solution spectra of cello-oligosaccharides and of a low DP cellulose.[72] They are indicated in Figure 4, which shows a spectrum of cotton linters.[73] Beginning at the upfield part of the spectrum, the region between 60 and 70 ppm is assigned to C-6 of the primary alcohol group. The next cluster of resonances, between 70 and 81 ppm, is attributed to C-2, C-3, and C-5, the ring carbons other than those anchoring the glycosidic linkage. The region between 81 and 93 ppm is associated with C-4 and that between 102 and 108 ppm with C-1. the anomeric carbon.

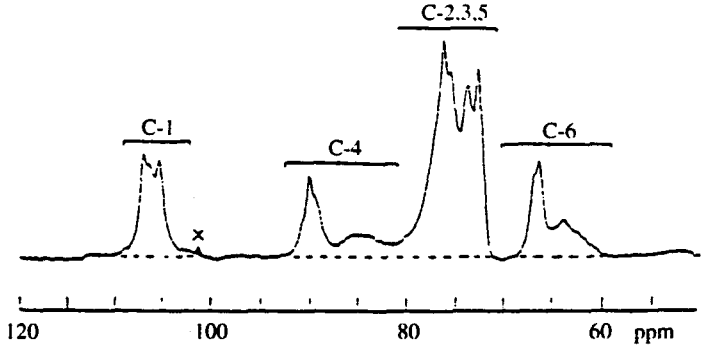


Figure 4 ^{13}C CP-MAS spectrum of cotton linters. The horizontal bars indicate the spectral ranges of the corresponding carbon sites in the anhydroglucose monomer unit of cellulose. The "X" marks the position of the small first spinning sideband of linear polyethylene, which was added as a chemical shift reference. The polyethylene center band (not shown) occurs at 33.63 ppm; the zero of reference for chemical shifts is liquid tetramethylsilane. Note the existence of both broad and narrow resonance features.[73]

In one of the first reports on application of the technique to studies of different celluloses, the splittings of the resonances of C-4 and C-1 in the spectrum of cellulose II (Figure 5) were regarded as confirmation of the occurrence of nonequivalent glycosidic linkages that had earlier been proposed on the basis of comparison of the Raman spectra of cellulose II and cellobiose in the O—H stretching region.[61] These splittings were also observed in the CP/MAS spectra of the cello-oligodextrins, which crystallize in a lattice very similar to that of cellulose II. In that context the splittings were attributed to the occurrence of nonequivalent cellulose molecules in the same unit cell.[69] However, such an interpretation leaves open the question as to why the resonances for carbons 2, 3, and 5 do not display similar splittings. If the splittings were indeed due to nonequivalent molecules it would be anticipated that those carbons nearest to the boundaries of the molecule would be the most affected. The carbons anchoring the glycosidic linkage. that is, C-1 and C-4, are the ones most removed from adjacent molecules, yet they also display the greatest splittings.

Interpretation of the spectra of native celluloses presented an even more challenging task. In the spectrum of cotton linters (Figure 4), the two resonance regions associated with C-6 and C-4 include sharper resonances overlapping broader upfield wings. After excluding the possibility that the broader wings could arise entirely from molecular mobility, [66,67] the wings were attributed to cellulose chains in two categories of environment. The first includes all chains located at the surfaces of cellulose microfibrils, which, because of their occurrence at the boundary, are less constrained with respect to the conformations they can adopt. The surfaces are regarded as regions of limited two-dimensional order. The importance of this category of order had earlier been demonstrated in a

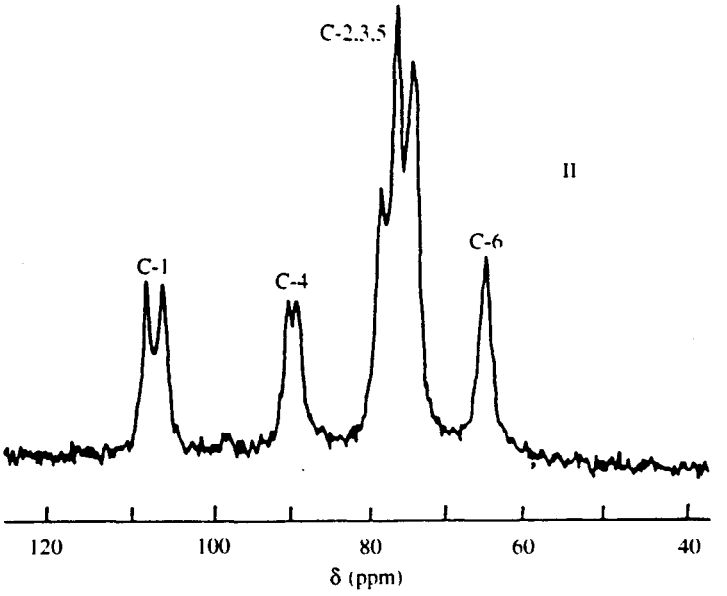


Figure 5 The CP/MAS ^{13}C spectrum of high crystallinity cellulose II recorded at relatively low resolution. Chemical shifts are shown in parts per million relative to Me_4Si. Assignment of the C-1, C-4, and C-6 resonances are based on pertinent liquid-state spectra.[61]

study of different native celluloses undertaken by Earl and VanderHart.[67] The celluloses had natural fibril diameters varying between 3.5 and 20 nm and it was shown that the areas of the upfield wings of C-4 and C-6 declined as the surface to volume ratio declined. The second category of environments contributing to the upfield wings is that of chains in regions within which the incoherence of order is not limited to two dimensions. Here, the dispersion of the frequencies at which resonances occur may arise from conformational differences, variations in bond geometries, changes in hydrogen bonding patterns, and nonuniformities in neighboring chain environments. These possibilities arise because in such regions the molecular chains are free to adopt a wider range of conformations than the ordering in a crystal lattice or its boundaries would allow.

Although the obvious upfield wings of the C-4 and C-6 resonances are the most direct evidence for the cellulose chains in less ordered environments, it is expected that the chains in these environments make similar contributions to the resonance regions associated with the other carbons. In the region of C-1, the contribution appears to be primarily underneath the sharper resonances, though a small component appears to extend towards 104 ppm. Similarly, it is expected that the contribution from chains in the less ordered environments underlies the sharper resonances of the C-2, C-3, and C-5 cluster.

The relative contributions of the two categories of environment to the intensity of the upfield wings was assessed in a careful analysis of the C-4 wing.[73] It was demonstrated that part of the wing could be correlated with the range of the C-4 resonance in amorphous cellulose prepared by ball milling. It was therefore assigned to cellulose chains occurring in the second type of environment, that is, domains wherein the incoherence of order is extended in all three dimensions. The other part of the wing was attributed to chains at the surfaces of the fibrils and, on the basis of these comparisons, it was concluded that approximately 50% of the wing is contributed by cellulose chains in each of the two types of less ordered environments described in the preceding paragraph. Though the upfield wing of C-4 is the basis of this allocation of intensities, it can be assumed that the relative contributions are similar for the upfield wing of C-6 and for the component that appears to underlie the sharper resonances at C-1. It is also expected that these domains contribute to the total intensity of the C-2, C-3, and C-5 cluster between 70 and 81 ppm.

The sharper resonances in the C-6 and C-4 regions, centered at 66 and 90 ppm, respectively, each appear to consist of more than one resonance line even though the resolution is not sufficient to distinguish the components well. The C-6 resonance seems to include at least two components while the C-4 resonance appears to include three closely spaced component lines. These multiplicities were interpreted as arising from carbons in cellulose molecules within the interior of crystalline domains and therefore taken as evidence of the occurrence of chemically equivalent carbons in different magnetic environments within the crystalline domains.

The region between 102 and 108 ppm, attributed to C-1, also reveals multiplicity and sharp

resonance features. Here, however, the shoulder is very limited. It appears that the resonances associated with the two categories of disordered domains described above lie underneath the sharp resonances associated with the interior of the crystalline domains. It can be concluded that, in most instances, the dispersion of frequencies associated with the disorder is small relative to the shift associated with the character of the anomeric carbon C-1, while that is not the case for the shifts associated with C-4 and C-6. Alternatively, it may be evidence that, because of the anomeric effect, the internal coordinates surrounding C-1 are much less flexible within the range of possible conformational variations than the other internal coordinates.

In search of a rationalization of the splittings observed in the sharp resonances, (CP/MAS) ^{13}C NMR spectra of a wide variety of samples of cellulose I were recorded. Some of these are shown in Figure 6. They include (a) ramie fibers (b), cotton linters (c), hydrocellulose prepared from cotton linters by acid hydrolysis (d), a low-DP regenerated cellulose I (e), cellulose from *Acerobacter xylinum* (f), and cellulose from the cell wall of *Valonia ventricosa*, an alga. While similar observations were reported in a number of studies,[61,63,67-71] their implications with respect to structure were more fully developed in the work of VanderHart and Atalla,[73,74] which provides the basis for the following discussion.

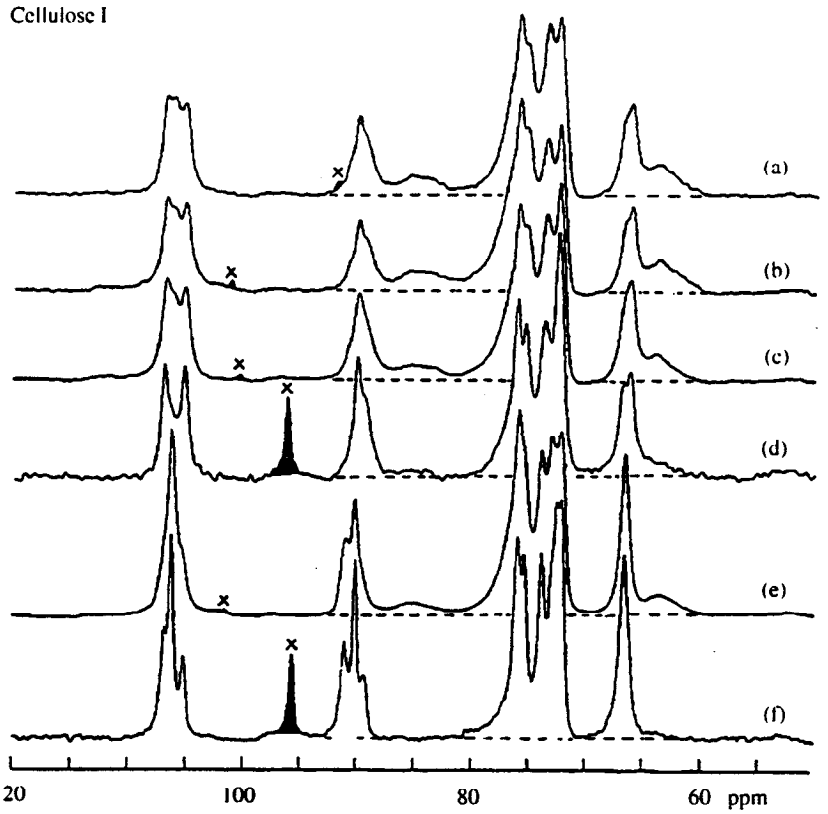


Figure 6 The ^{13}C CP/MAS spectra of several cellulose I samples: (a) ramie; (b) cotton linters; (c) hydrocellulose from cotton linters; (d) a low DP regenerated cellulose I: (e) *Acetobacter xylinum* cellulose; (f) *Valonia ventricosa* cellulose. Note the varied fine structure particularly at C-1 and C-4. Signal-to-noise variation due to limited amount of some samples. In that instance more polyethylene was added so the side band intensity increased. No line broadening or resolution enhancement techniques were applied in the acquisition of the spectra (after VanderHart and Atalla[73]).

All of the spectra shown in Figure 6 (a)-(f) are of celluloses that occur in relatively pure form in their native states and require relatively mild isolation procedures. The most striking feature in these spectra, when viewed together, is the variation in the patterns of the multiplets at C-1, C-4, and C-6. These resonances, which are viewed as arising from chains in the interior of crystalline domains, appear to be unique to the particular celluloses; among the native forms they appear to be distinctive of the source species. The first attempt to rationalize the spectra was in terms of information that they might provide concerning the unit cell of the structure of cellulose I. However, it soon became obvious that such a rationalization was not possible because the relative intensities within the multiplets were not constant nor were they in ratios of small whole numbers as would be the case if

the same unit cell prevailed throughout the crystalline domains. The conclusion was that the multiplicities were evidence of site heterogeneity within the crystalline domains and that, therefore, native celluloses must be composites of more than one crystalline form.

Further rationalization of the spectra required a careful analysis of the mutliplets at C-1, C-4, and C-6 and the variations of the relative intensities of the lines within each multiplet among the spectra of the different celluloses. In addition to excluding a single crystal form on the basis of the considerations noted above, it was also possible to exclude the possibility of three different forms with each contrbuting a line to the more complex multiplets. Thus, a decomposition of the spectra on the basis of two distinct crystalline forms was pursued. The results of the decomposition are shown as spectra (b) and (c) in Figure 7, and were designated as the I_a and I_b forms of native cellulose; this designation was chosen in order to avoid the possibility of confusion with the I_A and I_B forms that had earlier been defined in terms of differences in the appearance of the O-H bands in different types of native celluloses. [22,31] Spectrum (a) was acquired from a high-crystallinity sample of cellulose II and is included so as to distinguish the heterogeneity of crystalline forms occurring in the different forms of cellulose I from the long-known polymorphic variation of the crystallinity of cellulose.

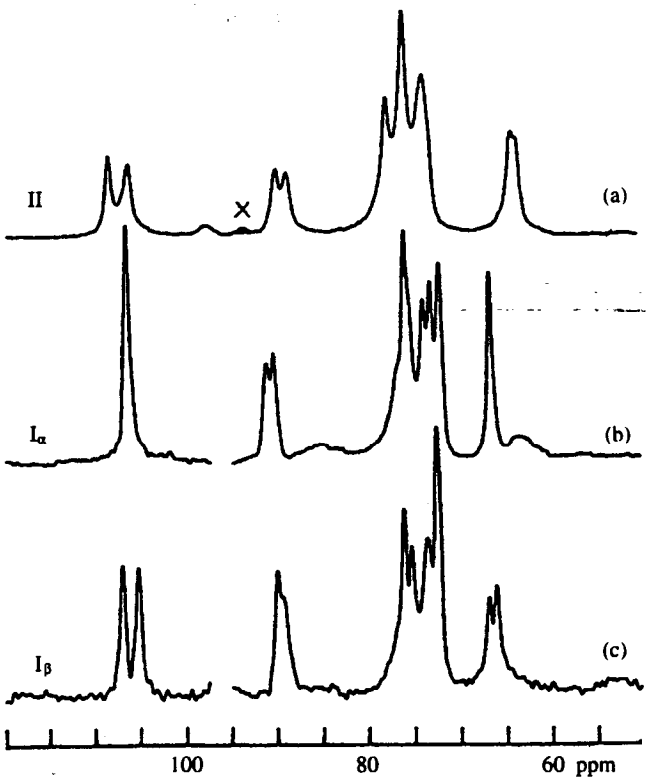


Figure 7 Comparison of the ^{13}C CP-MAS spectrum (a) of a low-DP cellulose II sample and the spectra (b) and (c) corresponding, respectively, to the two proposed crystalline forms of cellulose I, namely I_α and I_β. Spectra (b) and (c) were obtained by taking linear combinations of the low-DP and *Acetobacter* cellulose spectra. Discontinuities in spectra (b) and (c) occur where the polyethylene sidebands would have appeared. The I_α spectrum still contains a significant amount of non-I_α resonances as shown by the visible C-4 and C-6 upfield wings. Multiplicities of the C-1, C-4, and C-6 narrower resonances ought to indicate unit cell inequivalences.[73]

Spectra (b) and (c) in Figure 7 were in fact derived from appropriate linear combinations of the spectra of the low-DP cellulose I (d) and of the *Acetobacter xylinum* cellulose (e) in Figure 6. Though they represent the best approximations to the two forms of cellulose postulated, they cannot be regarded as representative of the pure forms as they do not adequately reflect the component of the cellulose at the surfaces of the crystalline domains. Spectrum 7(b) does have some intensity in the upfield wings of C-4 and C-6, but spectrum 7(c) has very little evidence of such wings. There is very little question, however, that the sharp components of spectra 7(b) and 7(c) include the key features in the spectra of the I_α and I_b forms. It is of interest to note here that among the distinct resonances of the I_α form at C-1, C-4, and C-6, only the one at C-4 appears to be split, while for the I_b form all

three resonances associated with these carbons show splitting, with the one at C-1 the most pronounced.

In an effort to further validate the proposal that the I_α and I_b forms were the primary constituents of native celluloses, VanderHart and Atalla undertook another extensive study to exclude the possibility that experimental artifacts contributed to the key spectral features assigned to the two forms.[75] A number of possible sources of distinctive spectral features were explored. The first was the question whether surface layers associated with crystalline domains within particular morphological features in the native celluloses could give rise to features other than those of the core crystalline domains. The second was whether variations in the anisotropic bulk magnetic susceptibility associated with different morphologies could contribute distinctive spectral features. Exploration of the spectra of higher plant celluloses with different native morphologies revealed very little difference in the essential features of the spectra, even after the samples had been subjected to acid hydrolysis. Furthermore, it was concluded that the I_α, component of higher plant celluloses was sufficiently low that some question was raised as to whether it occurs at all in these higher plant celluloses. In this context, it was also concluded that, in higher plant celluloses, the lineshapes of the I_b form at C-4 could only be reconciled with a unit cell possessing more than four anhydroglucose residues per unit cell.

Attention was then directed to analysis of the spectra of algal celluloses wherein the I_α component is the dominant one. Relaxation experiments confirmed that the essential spectral features identified with the two crystalline forms of cellulose were characteristic of the core crystalline domains; when measurements were conducted such that magnetization of the surface domains was first allowed to undergo relaxation, very little change in the spectral features was observed. The relaxation experiments suggested that domains consisting of both the I_α and I_b forms have equal average proximity to the surface. One possible interpretation of these observations, that the two forms are very intimately mixed, was ruled out at that time on the basis of hydrolysis experiments, the results of which are now in question.

Two groups of modifying experiments were carried out with the algal celluloses. In the first, the algal celluloses were subjected to severe mechanical action in a Waring blender. In the second, the algal celluloses were subjected to acid hydrolysis, in 4 N HCl for 44 h at 100 °C. While the mechanical action resulted in some reduction in the proportion of the I_α form, acid hydrolysis resulted in a dramatic reduction, sufficient indeed to make the spectra seem like those of the higher plants, except that resolution of the spectral lines was much enhanced relative to that observed in the spectra of even the purest higher plant celluloses. The samples subjected to hydrolysis wherein the recovery varied between 12 and 22%, were examined by electron microscopy and shown to have lateral dimensions not unlike those of the original samples. These observations were interpreted to imply that the I_α form is more susceptible to hydrolysis than the I_b form. An earlier study of the effect of hydrolysis, under similar conditions but for only 4 h, had been carried out with cellulose from *Rhizoclonium heiroglyplicum* with no discernible effect on the spectra.[76] The difference in duration of the hydrolysis may well have been the key factor. Both of these observations and their interpretations had been presented, however, before it was recognized that exposure of celluloses with relatively high contents of the I_α form to elevated temperatures can result in its conversion to the I_b form.[77] When the possibility that the I_α content of the algal cellulose had been converted to the I_b form is taken into account, the results of the relaxation experiments of VanderHart and Atalla cited above can be reinterpreted as indicating intimate mixing of the I_α and I_b forms within the crystalline domains of the algal celluloses.

VanderHart and Atalla also took advantage of the spectra derived from the acid hydrolyzed samples of the algal cellulose to generate more highly resolved representative spectra of the I_α and I_b forms. These are shown in Figure 8 where it is clear that even in the spectrum representative of the I_α form the upfield wings of the C-4 and C-6 resonances are reduced to a minimum. With the completion of this study by VanderHart and Atalla, most of the questions about the possibility that the spectral features were the results of artifacts were put to rest, and the hypothesis that all native celluloses belong to one or to a combination of these forms was generally accepted.

With the above resolution of the questions concerning the nature of native celluloses in mind, it was possible to classify these celluloses with respect to the relative amounts of the I_α and I_b forms occurring in the celluloses produced by particular species. It emerged in these early studies that the celluloses from more primitive organisms such as *Valonia ventricosa* and *Acetobacter xylinum* are predominantly of the I_α form, while those from higher plants such as cotton and ramie are predominantly of the I_β form. As noted earlier, the nomenclature chosen was intended to avoid confusion with the I_A and I_B forms previously used to classify the celluloses on the basis of their infrared spectra in the OH stretching region. In relation to that classification, the NMR spectra

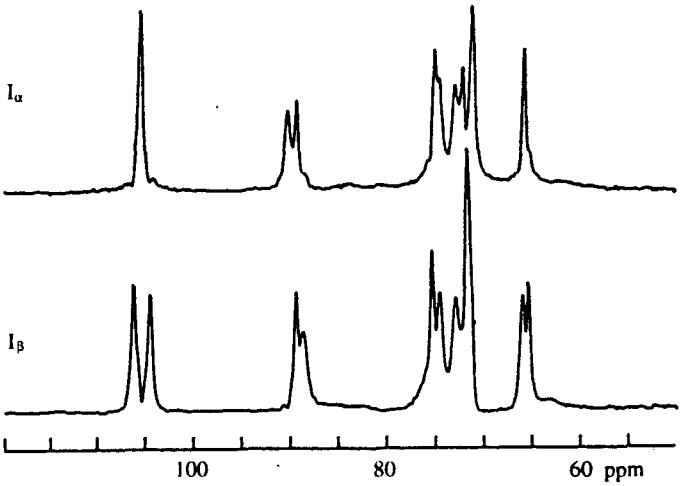


Figure 8 Alternative candidates for the spectra of celluloses I_α (top) and I_β (bottom) derived from linear combinations of the spectra of I_α rich *Cladophera glomerata*, before and after acid hydrolysis, which resulted in a I_β rich cellulose.[75]

suggest that the I_A group has the I_α form as its dominant component, while the I_B group is predominantly of the I_b form.

With the recognition that all plant celluloses are composites of the I_α and I_b forms, it was possible to rationalize many of the earlier difficulties in developing suitable structural models. It became clear that the efforts to reconcile the diffraction patterns in terms of a unique unit cell for native celluloses were frustrated by the reality that the celluloses were composites of two crystalline forms that were blended in different proportions in celluloses produced by different organisms. As noted earlier, the finding that celluloses from different biological sources represented unique blends of the two different forms also resolved the paradox posed early in this century by Cross and Bevan[4] when they contrasted the observation of individuality in biological systems with the invariance of crystalline forms.[4]

3.16.3.2 Further Studies of Structures in Cellulose

With the wide acceptance of the proposal of the two crystalline forms (I_α and I_b) came the challenge of understanding the differences between them and their relationship to each other within the morphology of native cellulosic tissues. A number of complementary approaches were pursued by different investigators in the search for answers to these questions. Some were based on further application of solid-state ^{13}C NMR to the study of different celluloses as well as to celluloses that had been subjected to different modifying treatments. Others were based on application of Raman and IR spectroscopy to new classes of cellulosic samples. Others still were based on refinement of electron microscopic and electron diffractometric methods. The results of these investigations will be presented in summary.

3.16.3.2.1 Raman and infrared spectroscopic studies

The categorization of native celluloses into the I_A and I_B group by Howsmon and Sisson[22] and Blackwell and Marchessault[31] on the basis of the appearance of the OH stretching region of their infrared spectra suggested that the hydrogen bonding patterns within the crystalline domains may be part of the key to the differences between the two forms of native cellulose. This was, in fact, confirmed in the course of more detailed investigations of the Raman spectra carried out on single oriented fibers of native celluloses[78] and in a comprehensive study of the infrared spectra of a number of celluloses of the two forrns.[79] The Raman spectral investigations were part of a broader study directed primarily at assigning the bands associated with the skeletal vibrational motions and at exploring the differences between celluloses I and II.[78] They differed from earlier Raman spectral studies in that the spectra were recorded with a Raman microprobe on which individual fibers could be mounted for spectral investigation. With this system it was also possible to explore the variation

of intensity of the bands as the polarization of the exciting laser beam was rotated relative to the axis of the fibers.

The observed spectra are shown in Figures 9 and 10, each of which includes six spectra. Figure 9 shows the region between 250 and 1500 cm^{-1}, while Figure 10 shows the region above 2600 cm^{-1}; the region between 1500 and 2600 cm^{-1} does not contain any spectral features. The spectra in Figures 9 and 10 are of native and mercerized ramie fibers and native *Valonia ventricosa,* and they are recorded with both parallel and perpendicular polarization of the exciting laser beam. Those identified as 0° spectra were recorded with polarization of the electric vector of the exciting laser beam parallel to the direction of the fiber axes, while those identified as 90° spectra were recorded with polarization of the electric vector of the laser perpendicular to the fiber axes. The ramie fibers are known to have the molecular chains parallel to the fiber axes; the *Valonia ventricosa* fibers were prepared by drawing the cell wall in order to align the microfibrils within it.

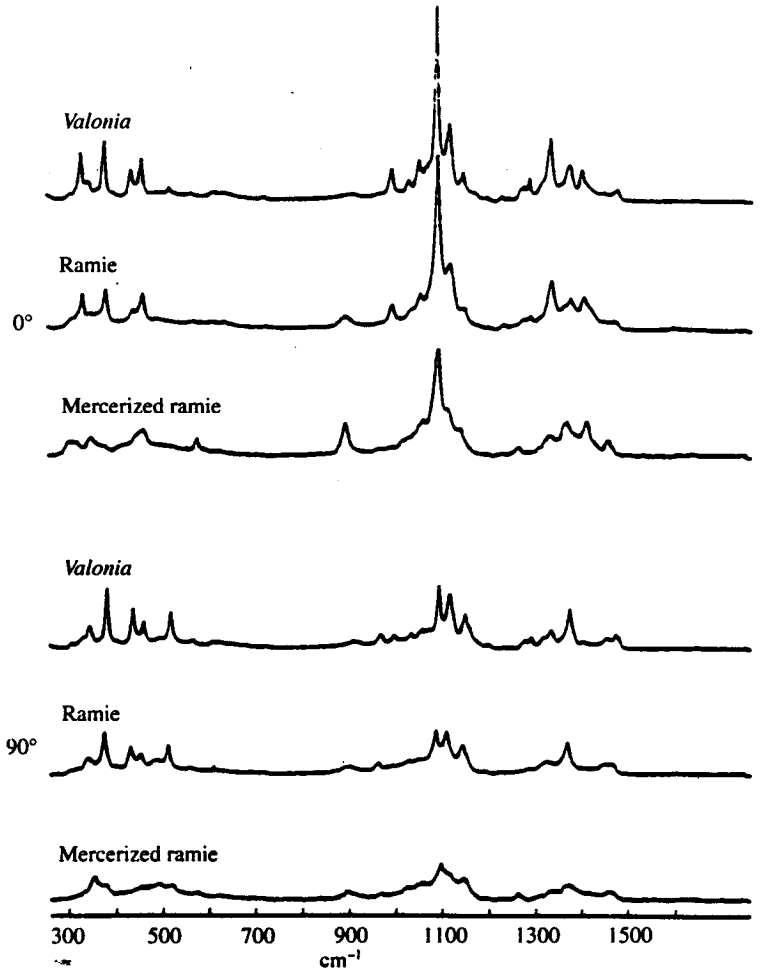


Figure 9 Comparison of the Raman spectra from *Valonia*, ramie, and mercerized ramie (low-frequency region). Spectra were recorded with the electric vector at both 0° and 90°.[78]

A number of features in the spectra are noteworthy with respect to earlier discussions. The first is a comparison of the spectra of *Valonia ventricosa* and ramie. It is clear that, apart from a broadening of the bands in the ramie spectra, because of the smaller lateral dimensions of the crystalline domains, the spectra are very similar except in the OH stretching region. This was interpreted as evidence that the chain conformations in both the I_α and I_b forms are the same but that the hydrogen bonding patterns between the chains are different within the two forms. This interpretation is more clearly demonstrated in a comparison of the spectra of *Valonia ventricosa* and *Halocynthia* presented below.

The second feature worthy of note is the dramatic difference between the spectra of native (cellulose I) and mercerized (cellulose II) ramie fibers, particularly in the low-frequency region. This was taken as further confirmation that the conformations of cellulose I and cellulose II must differ sufficiently to result in significant alteration of the coupling patterns between the internal vibrational

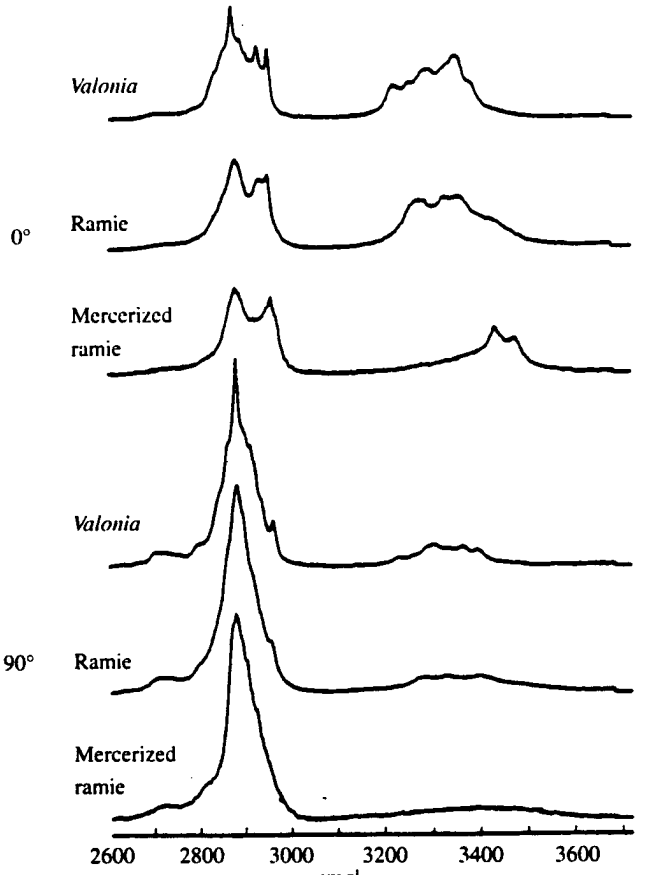


Figure 10 Comparison of the Raman spectra from *Valonia*, ramie, and mercerized ramie (high-frequency region). Spectra were recorded with the electric vector at both 0° and 90°.[78]

modes of the pyranose rings in the molecular chains. It is interesting in this connection to also compare the intensities of the band at 1098 cm^{-1} in the spectra of the two forms of ramie. The band is clearly less intense in the spectrum of the mercerized sample, suggesting that a conformational change which reduces coupling of the skeletal motions has occurred. The 1098 cm^{-1} band is the strongest skeletal band and it is the most intense feature in the spectrum when polarization of the exciting radiation is parallel to the chain direction. The sensitivity of the Raman spectra to the orientation of an intramolecular vibrational motion is also illustrated in the intensity of the methine CH stretching band at about 2889 cm^{-1}. It is most intense with the electric vector of the exciting radiation at 90° to the chain axis, an orientation which is parallel to that of the methine CH bonds of the pyranose rings.

Finally, in light of the discussion of the nonequivalence of adjacent anhydroglucose units and the corresponding nonequivalence of alternating glycosidic linkages, the OH region in the 0° spectrum of mercerized ramie is of particular interest. It shows the two distinct sharp bands that provide evidence of the presence of nonequivalent intramolecular hydrogen bonds in concert with the alternating glycosidic linkages along the chain; the hydrogen bonds are oriented parallel to the chain direction. This alternation clearly stands out most distinctly in cellulose II. These distinct bands cannot be attributed to nonequivalent chains as the difference in frequency implies a difference in the O···O distances between the oxygen atoms anchoring the hydrogen bond as well as a difference in the dihedral angles y and f of the associated glycosidic linkages. Nonequivalent chains would have different periods in the chain direction if they possessed twofold helical symmetry.

Infrared spectral studies of the I_α and I_b forms were carried out by Sugiyama *et al.* [79] on a number of different native celluloses of both forms. Furthermore, it included examination of a number of I_α rich celluloses that were converted to the I_b form through the annealing process first reported by Hori and co-workers.[80] In order to complement the infrared spectra, Sugiyama *et al.* [81] recorded electron diffraction patterns for the samples, which allowed classification of the celluloses through comparison with the diffraction patterns acquired in an earlier electron diffractometeric study which will be discussed in greater detail in a subsequent section (3.16.3.2.3).

The key finding emerging from examination of the infrared spectra of the different forms was that the only differences noted were in bands clearly associated with the OH group. This was also true of the changes observed upon conversion of the I_α form to the I_b form through annealing. The bands associated with both the differences in native forms and with the effects of transformation were observed in both the O-H stretching region above 3000 cm^{-1} and the O-H out-of-plane bending region between 650 and 800 cm^{-1}. It was reported that spectra of the I_α form had distinctive bands at 3240 and 750 cm^{-1}, while the spectra of the I_b form had distinctive bands at 3270 and 710 cm^{-1}. Furthermore, it was observed that the band at 3240 cm^{-1} appears to be polarized parallel to the direction of the fibril orientation, while the band at 3270 cm^{-1} is not polarized. Among the low-frequency bands, the one at 710 cm^{-1} appears to be polarized perpendicular to the fibril direction, while the one at 750 cm^{-1} is not polarized. It was also observed that upon transformation of the I_α rich celluloses to the I_b form through annealing, the corresponding bands changed accordingly. The authors concurred with the interpretation of the differences between the two forms suggested by Wiley and Atalla and concluded that I_α to I_b transformation corresponded primarily to a rearrangement of the hydrogen bond system within the structures and that the two structures appeared to have very similar conformations. The infrared spectral studies by Sugiyama *et al.* are particularly interesting because they included the spectra of both *Valonia* and *Halocynthia,* the Raman spectra of which have been investigated at high resolution.[82]

The Raman spectra of *Valonia macrophysa* and *Halocynthia* (tunicate) celluloses acquired by Atalla *et a1.* [82] are shown in Figure 11. These particular spectra are of interest because *V. macrophysa* is known to be predominantly the I_α form, while *Halocynthia* is predominantly of the I_b form. Comparison of their spectra can be more rigorous than was possible in the earlier work of Wiley and Atalla[78] because the lateral dimensions of the fibrils of both forms are of the order of 20 nm. with the result that their spectra show equal resolution of the bands in all regions of the spectrum. It is to be noted that their spectra are essentially identical in all of the regions associated with skeletal vibrations of all types as well as regions associated with most of the vibrations involving C—H bonds, whether in the bending or stretching regions. Indeed, the primary differences between the two spectra are in the broad complex bands that occur in the O—H stretching region, and these differences are not unlike those noted in the earlier Raman spectral studies described above. In addition, the weak band at about 840 cm^{-1} in the spectrum of *V. macrophysa* has no corresponding band in the spectrum of *Halocynthia;* this is the band attributed to the out-of-plane bending vibrations of hydrogen bonded OH groups. There are also minor differences in the relative intensities of the methylene C—H stretches and H—C—H bending vibrations, but these are the natural consequences of different hydrogen bonding patterns for the hydroxyl group at C-6. Comparison of the two spectra reinforces the interpretation presented earlier, on the basis of spectra in Figures 9 and 10, leading to the conclusion that the only difference between the I_α and I_b forms is in the pattern of hydrogen bonding. Thus, Raman spectral comparison of the two forms is entirely consistent with that reported for the infrared spectra of these highly crystalline celluloses. It must be kept in mind, of course, that the bands associated with OH group vibrations are not expected to coincide in the Raman and infrared spectra; because of the different bases for activity in the two different spectral approaches to measurement of vibrational frequencies, Raman-active vibrational modes are frequently silent in the infrared and vice versa. This, of course, is also true for the skeletal bands.

In view of the considerable variation observed in the Raman spectra of celluloses as a result of changes in molecular conformations, there can be little question that the spectra in Figure 11 indicate that the conformations of the cellulose molecules in *Valonia* and *Halocynthia* are essentially identical. It is also important to note that the Raman spectra of the celluloses from *V. macrophysa* and *Valonia ventricosa,* both of which have been used in different studies as representative of the I_α form, are effectively indistinguishable in all regions of the spectra. This is also true of the Raman spectra of celluloses from the algae *Cladophera glomerata* and *Rhizoclonium heirglyphicum,* which have also been used in many studies as representative of celluloses that are predominantly of the I_α form.

In summary then, the Raman and infrared spectral studies undertaken after discovery of the composite nature of native celluloses point to the conclusion that the only difference between the two forms is in the pattern of hydrogen bonding between chains that possess identical conformations. Yet electron microscopic and electron diffractometric studies to be described in greater detail in a following section (3.16.3.2.3) have led to conclusions that the two forms represent two crystalline phases with different crystal habits.[81] It is therefore important to consider what information may be developed from the vibrational spectra with regard to this question.

The key conclusion drawn from the electron diffractometeric data was that the I_α form represents

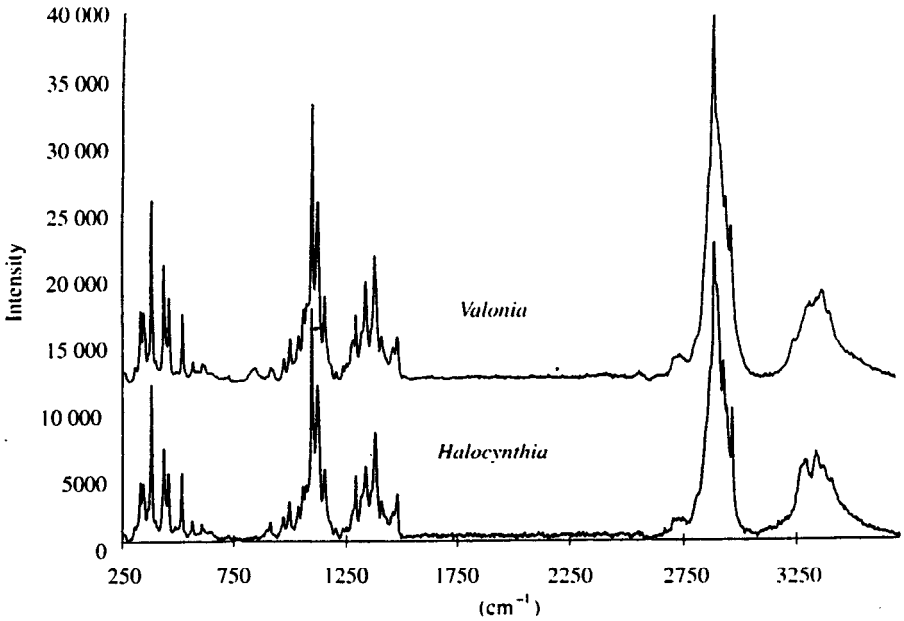


Figure 11 Raman spectra of I_α rich *Valonia* and I_β rich *Halcynthia* celluloses. Because both have fibrils of large lateral dimensions, the spectra of both are well resolved and provide a better basis for comparison of the spectra of I_α rich and I_β rich celluloses.[82]

a triclinic phase with one chain per unit cell, while the I_β form represents a monoclinic phase with two chains per unit cell. Furthermore, the symmetry of the monoclinic phase appeared to be that of space group $P2_1$. It has been recognized recently[82] that such a proposal is not consistent with the vibrational spectra. While it was not possible to have full confidence in this conclusion on the basis of the earlier spectral data because of the differences in the level of resolution between the spectra of ramie and *Valonia* celluloses, the spectra shown in Figure 11 are of sufficiently high resolution and sufficiently similar that the comparisons can indeed be made with confidence.

The key issue is that when crystal structures possess more than one molecule per unit cell, and the molecules have the same vibrational frequencies, the vibrational modes of the unit cell become degenerate. Under these circumstances, couplings will arise between equivalent modes in the different molecules, and it is generally observed that such couplings result in splittings of the bands associated with key vibrational modes. The type of coupling that is relevant in the case of cellulose is that described as correlation field splitting.[83] This effect arises because, as a result of the coupling, the vibrations of a particular mode in the two molecules will now occur at two frequencies that are different from those of the isolated molecule; one of the two new frequencies will have the modes in the two different molecules in phase with each other, while the other will have the modes out of phase with each other. Such correlation field effects result in doublets with a splitting of 10-15 cm^{-1} in some modes of crystalline polyethylenes having two chains per unit cell. Since no evidence of such splittings occurs in the *Halocynthia* spectrum shown in Figure 11, it must be concluded that the I_b form cannot have more than one molecule per unit cell. Nor can it be suggested that the two molecules in a monoclinic unit cell are nonequivalent and may have modes that are at different frequencies because the skeletal bands in the *Halocynthia* spectrum are essentially identical to those in the *Valonia* spectrum. Furthermore, this similarity was also reported in the infrared spectra observed by Sugiyama *et al.*[81] Thus, it is clear that the vibrational spectra, both Raman and infrared, point to the conclusion that both the I_α and I_b forms have only one molecule per unit cell. This conclusion of course raises the question as to why the crystallographic data has been viewed for so long as pointing to a two-chain unit cell with the symmetry of space group $P2_1$. This is an issue that is best addressed after the results of electron diffractometric studies have been described in greater detail.

3.16.3.2.2 Solid-state ^{13}C NMR spectroscopic studies

It is not surprising that the methodology which first provided the basis for understanding the composite nature of native celluloses in terms of the I_α and I_b duality has continued to be the one

most often used for seeking deeper understanding of the differences between native celluloses derived from different biological sources. This has been facilitated by the greater availability of solid-state ^{13}C NMR spectrometer systems and by the relative simplicity of the procedures for acquiring the spectra from cellulosic samples. The studies undertaken on the basis of further examination of the solid-state ^{13}C NMR spectra of celluloses are in a number of categories. The first group are focused on further examination of the spectra of different native celluloses, in part aided by mathematical procedures for deconvolution of the spectra or for resolution enhancement. Another group rely on exploring the spectral manifestations of native celluloses that have been modified in different ways. Yet a third approach is based on investigation of celluloses subjected to different but well-known procedures for inducing structural transformations in the solid aggregated state of cellulose. Since the approaches adopted by some groups of investigators in their acquisition and analysis of the spectra have been different, the following discussions will be structured to reflect these differences. The work of the different groups with regard to native cellulose and its response to a variety of treatments will be explored first to the extent that it illuminates questions concerning the nature of native celluloses. This will be followed by an examination of the manifestations of a broader category of structural changes induced by different treatments known to alter the states of aggregation of cellulose. In selecting the investigations to be emphasized in this discussion, the author focuses on studies that provide insight into the variations of the states of aggregation with the history of particular celluloses, both with respect to source and processes of isolation and transformation.

The group at the Kyoto University Institute for Chemical Research carried out important studies that were complementary to those undertaken by VanderHart and Atalla.[73-75] A number of other groups have made valuable contributions. Since a number of questions concerning the nature of the I_α and I_b forms remain outstanding, it is useful to begin with an overview of the findings of different groups. These will then make it possible to view results of studies using other methods in a clearer perspective.

The early studies by the Kyoto University group have been well summarized in a report that addresses the key points that were the focus of their investigation.[84] In a careful analysis of the chemical shifts of the C-1, C-4, and C-6 carbons in the CP/MAS spectra of monosaccharides and disaccharides for which crystallographic structures were available, Horii *et al.* [84] recognized a correlation between the chemical shifts and the dihedral angles defined by the bonds associated with these particular carbons. In particular, with respect to C-6, they demonstrated a correlation between the chemical shift of the C-6 resonance and the value of the dihedral angle χ defining the orientation of the OH group at C-6 relative to the C-4—C-5 bond in the pyranose ring. This correlation is of value in interpretation of the solid-state ^{13}C NMR spectra with respect to structure as well as discussion of the implications of splittings of the C-6 resonances observed in some of the spectra.

Of even greater interest, in light of the discussions of deviations from twofold screw axis symmetry in some of the structures, it was observed that the chemical shifts of C-1 and C-4 are correlated with the dihedral angles about the glycosidic linkage. In particular, there was a correlation between the shift of C-1 and the dihedral angle 4 about the C-1—O bond and a correlation between the shift of C-4 and the dihedral angle y about the O—C-4 bond. As the spectra published in the earliest studies did not have sufficient resolution to reveal the splittings of the resonances of C-1 and C-4, the possibility of occurrence of nonequivalent glycosidic linkages was not addressed at that time.

In addition to analysis of the correlation between the chemical shifts and the dihedral angles, the Kyoto group investigated the distribution of cellulosic matter between crystalline and noncrystalline domains on the basis of measurements of the relaxation of magnetization associated with the different features of the spectra. By measurement of the values of the spin-lattice relaxation times $T_1(C)$ associated with the different spectral features, they developed a quantification of the degree of crystallinity in the different celluloses. They also undertook analysis of the lineshapes of the different resonances, particularly that of the C-4 resonance. The lineshape analysis was based on deconvolution of the spectral features into combinations of Lorentzian functions centered at the assigned shifts for the particular resonances. It is to be noted that the use of Lorentzian functions, which can be justified at a fundamental level in the case of spectra from molecules in solution, has no basis in any fundamental understanding of the phenomenology of acquisition of the solid-state ^{13}C NMR spectra. However, since deconvolution into Lorentzians has been found to be a useful tool in assessing the spectral features in the spectra of cellulose, its use has continued. The qualifications that must be kept in mind when it is used have been addressed by VanderHart and Campbell.[85]

In the early studies by Horii *et al.*, [84] all of the upfield wing of the C-4 resonance was attributed to molecules in noncrystalline domains. On this basis, lineshape analysis of the C-4 resonance of different native celluloses did not seem consistent with the model proposed by VanderHart and

Atalla[73] with respect to the composite nature of native celluloses. In later studies, when Horii *et al.* took note of the fact that, in the study by VanderHart and Atalla, approximately half of the upfield wing of C-4 in the spectra of higher plant celluloses was attributed to the surface molecules of crystalline domains, Yamamoto and Horii,[86] indicated that their results confirm the proposal of VanderHart and Atalla. It is to be noted that in their early reports in this area. Hori *et al.* used the designations I_b and I_a to designate the different groups of celluloses in which the I_α and I_β forms were dominant. However, in their later studies they have adopted the I_α and I_β designations that are designed to avoid confusion with the categories first introduced by Howsmon and Sisson[16] discussed earlier.

In pursuit of further understanding of the I_α and I_β duality, Horii et al. explored the effects of transformative treatments on the solid-state ^{13}C NMR spectra. The first group of studies were directed at the effects of annealing, first in saturated steam[77] and later in aqueous alkaline solutions (0.1 N NaOH) selected to avoid hydrolytic decomposition of the cellulose.[87,88] In summary, the key findings were that the celluloses wherein the I_α form is dominant are substantially transformed into the I_β form when conditions are established so as to allow the transformation to be complete. The cellulose representative of the I_α form that was used for these studies was *Valonia macrophysa*. The effects of the annealing treatment are demonstrated in Figure 12 which shows the progression in the degree of conversion as the treatment temperature is increased. Each of the treatments were for 30 min in aqueous alkaline solution. These results, of course, point to the susceptibility of the I_α form to conversion to the I_β form, suggesting that the latter is the more stable form. To test this hypothesis, a sample of tunicate cellulose, which had earlier been shown to be of the I_β form by Belton *et al.*,[89] was also annealed in an aqueous alkaline solution at 260 C; it showed little change as a result of the annealing.[88]

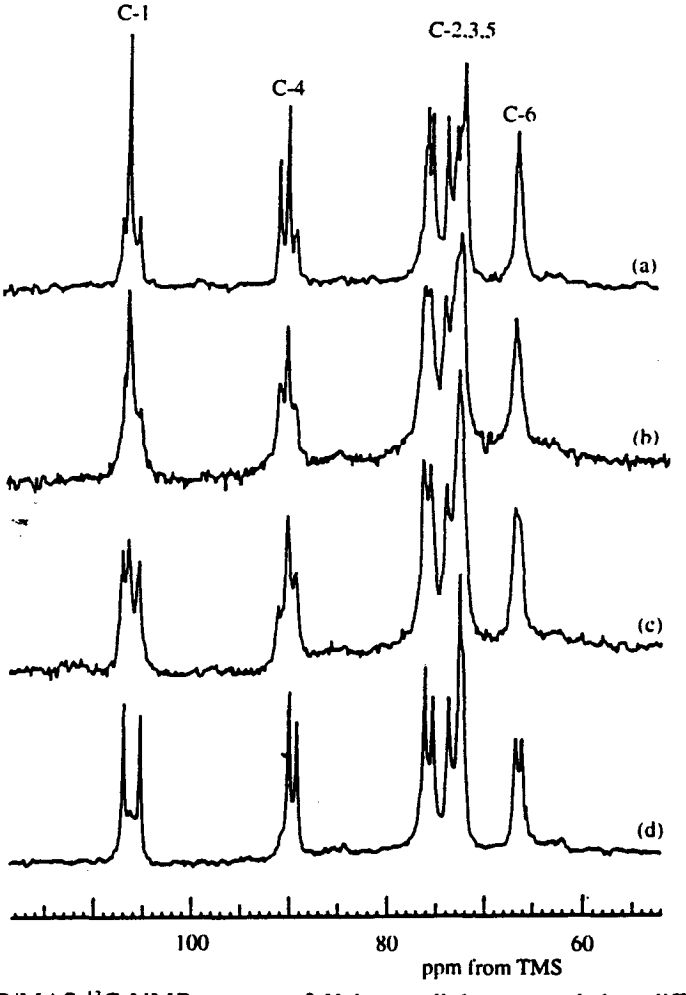


Figure 12 50 MHz CP/MAS ^{13}C NMR spectra of *Valonia* cellulose annealed at different temperatures in 0.1 N NaOH solution: (a) original; (b) 220 °C; (c) 240 °C; (d) 260 °C.[87]

Additional studies by the Kyoto group relied on solid-state ^{13}C NMR to explore the effects of different variables on the structure of cellulose.[90] They will be considered in a later section. It is in order, in the present context, to note briefly the results of one study in which celluloses from *Acetobacter xylinum* cultures were investigated. One of the variables explored was the temperature of the culture; it was observed that lower temperatures favored formation of the I_α form at the expense of the I_b form. This finding raises a fundamental question regarding the possibility that the variation of the balance between the two forms is, in part, an adaptive response to changes in the environment.

In 1990, Newman and Hemmingson[91] began to combine some additional methods of processing the ^{13}C NMR spectral data with those that had been used previously such as monitoring the value of T_1 (C) associated with the different spectral features. While these procedures incorporate a significant degree of empiricism, they have facilitated rationalization of the spectral features of a number of native celluloses and are therefore valuable contributions to the repertoire of methods available for interpreting the ^{13}C NMR spectra of native celluloses. It must be noted, however, that the application of these methods has been complemented in the work of Newman and Hemmingson by a considerable degree of awareness of the complexity of the structures of both native and processed celluloses, so that their application by others needs to be approached with this awareness in mind.

Two innovations were introduced into studies of cellulose in this work. The first was the adoption of an alternative resolution enhancement protocol that was justified primarily on the basis of its success in isolating certain features of the spectra. The second and perhaps more helpful innovation was the application of a procedure that relies on the variation in the proton relaxation times to edit the spectra in order to separate contributions from domains wherein the molecules had different degrees of molecular mobility. In lignocellulosic materials that have not been chemically fractionated through the application of appropriate isolation procedures, the possibility of separation of the subspectra from domains with different degrees of molecular mobility provides a distinct advantage. Thus, the procedure affords the opportunity to examine native tissues that include cellulose without the need to isolate the cellulose, and thereby to develop some useful information concerning the nature of the cellulose under conditions that more closely approximate its native state.

Resolution enhancement was accomplished by convoluting the free induction decay with a function having the form:

$$\mathbf{f}(t) = \exp\{\mathbf{a}t^2 - \mathbf{b}t^3\}$$

This function differs from that used by other investigators in earlier studies in that both exponents have been incremented by 1. The rationalization was that the function used in earlier studies was the one usually used for resolution enhancement in liquids and that there is no fundamental basis for its application to solids. Furthermore, the function adopted was found to be better suited to the studies of Parameters **a** and **b** were selected to enhance resolution "without decreasing the signal-to-noise ratio to unacceptable levels." Here again, it should be noted that this procedure needs to be applied with caution as it is in the nature of the function adopted for the convolution that it can create artifacts in the spectra. Applied without supporting information from complementary methods, it can be misleading.

The procedure based on variability in rates of proton spin relaxation[92] was used to generate spectra that were described as "proton spin-relaxation edited" or PSRE spectra Relaxation of the protons was allowed to occur for a particular period prior to the cross-polarization step in the CP-MAS protocol; these were referred to as "delayed-contact" experiments. Linear combinations of spectra acquired with and without the delay were then generated to describe the spectra of the domains within which the different populations of protons occurred. This procedure requires adjustment of parameters used in generating the linear combinations of the spectra. The guideline used for optimization was "to maximize mutual discrimination between signals at 89 ppm (C-4 in crystal interior cellulose) and 80 ppm (C-4 in noncrystalline cellulose) without allowing any signal to become inverted."

The results of this process are well illustrated in application of this approach to acquisition of spectra from Avicel microcrystalline cellulose illustrated in Figures 13 and 14. Figure 13 shows the normal CP-MAS spectrum together with one acquired with a proton relaxation time of 10 ms. Figure 14 shows the PSRE subspectra of the crystalline and noncrystalline domains generated through linear combination of the two spectra shown in Figure 13, with the linear combination optimized through trial and error to meet the criterion quoted above. It is clear that application of

this procedure reveals that a significant fraction of the upfield wings of C-4 and C-6 is associated with domains wherein the molecular mobility is lower than in the more disordered regions.

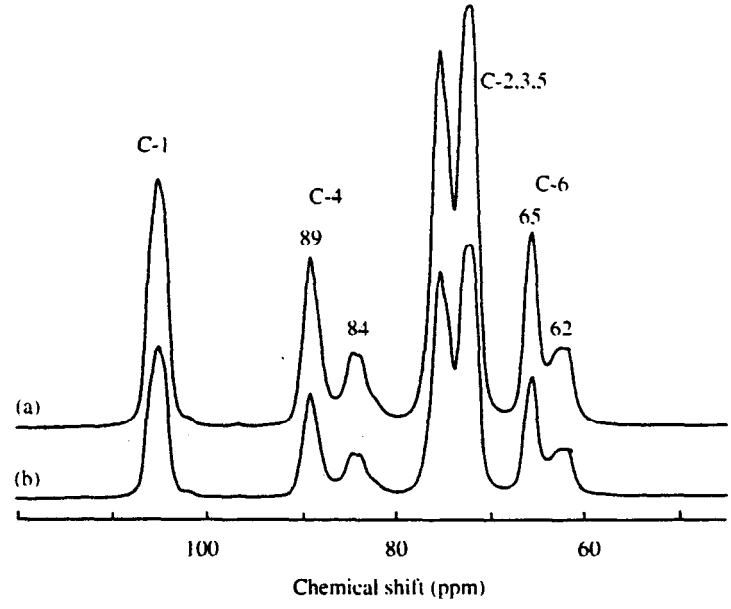


Figure 13 ^{13}C NMR spectra of microcrystalline cellulose I: (a) normal CP/MAS NMR spectrum, (b) delayed contact spectrum with $T = 10$ ms.[92]

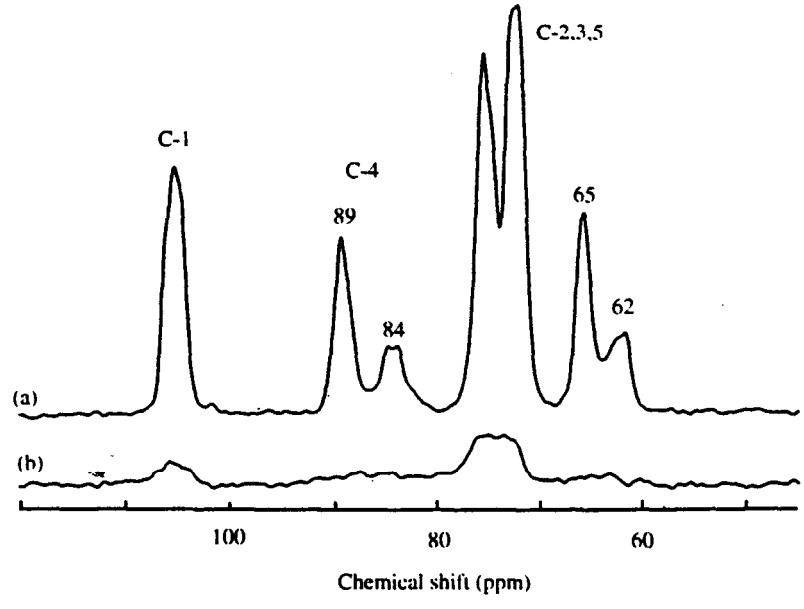


Figure 14 PSRE subspectra of microcrystalline cellulose I, obtained from the spectra shown in Figure 13, selected for domains with (a) $T_{1\rho}(H) = 20$ ms and (b) $T_{1\rho}(H) = 13$ ms.[92]

Application of the resolution enhancement procedure to the data for the more crystalline fraction in Figure 14 results in the spectra shown in Figure 15. Here, both the benefits and the risks of resolution enhancement procedures are demonstrated. The possibility of artifacts is suggested by the low-level peaks that occur in regions wherein no resonances occur in the unenhanced spectra. The magnitude of these artifacts is such that it raises uncertainty about the relative intensities of features that are recognized as real because they coincide with features known to occur in unedited and unenhanced spectra. On the other hand, resolution enhancement clearly demonstrates that the contribution of the less mobile component of the upfield wing of C-4 is a doublet. This doublet has been attributed to molecules on crystalline surfaces representing different faces of the crystalline

domains. The resolution enhanced spectrum also gives credibility to the suggestion of a doublet in the upfield wing of C-6.

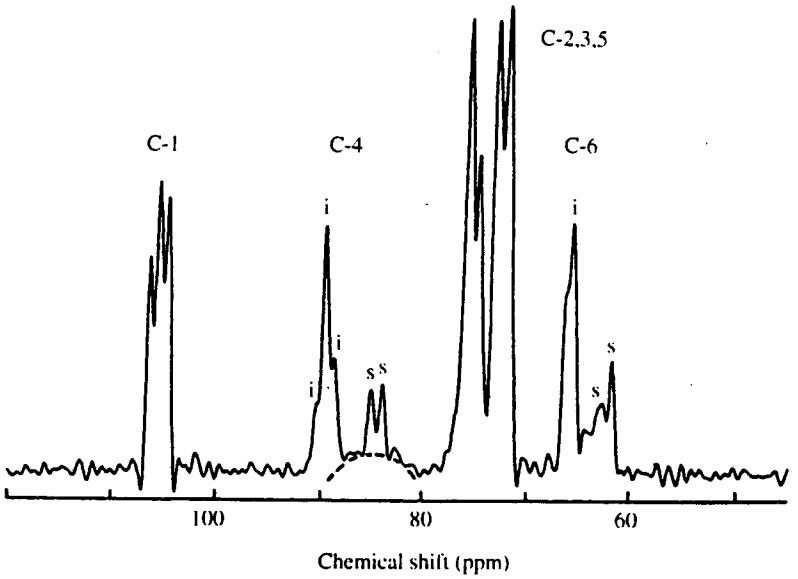


Figure 15 Resolution-enhanced PSRE subspectrum of crystalline domains in Avicel microcrystalline cellulose I. Signals assigned to crystal interior and crystal surface cellulose are labeled i and s, respectively. A broken line outlines a broad signal assigned to poorly-ordered surfaces.[92]

Newman *et al.* [93-96] have used resolution enhanced spectra in attempts to quantify the relative amounts of I_α and I_b forms in a number of native celluloses.[93-96] In such instances the results must be viewed with caution precisely because of the uncertainties concerning the manner in which the enhancement procedure can distort the relative intensities of the spectral features associated with each of the forms and because in the spectral regions characteristic of both C-l and C-4, there is significant overlap between the spectral features of the I_α and I_b forms. Nevertheless, when applied with care and on the basis of comparisons with spectra of well-characterized species, this approach can be useful for indicating trends.

A different approach to mathematical analysis of the solid-state ^{13}C NMR spectra of celluloses was introduced by the group at the Swedish Forest Products Laboratory (STFI).[97] They took advantage of statistical multivariate data analysis techniques that had been adapted for use with spectroscopic methods. Principal component analysis (PCA) was used to derive a suitable set of subspectra from the CP/MAS spectra of a set of well-characterized cellulosic samples. The relative amounts of I_α and I_b forms and the crystallinity index for these well-characterized samples were defined in terms of the integrals of specific features in the spectra. These were then used to derive subspectra of the principal components, which in turn were used as the basis for a partial least squares analysis of the experimental spectra. Once the subspectra of the principal components are validated by relating their features to the known measures of variability, they become the basis for analysis of the spectra of other cellulosic samples that were not included in the initial analysis.

Here again, it must be recognized that although the methodology has proved to be quite useful in characterizing both native and processed celluloses, it incorporates some arbitrarily defined measures of some of the variables investigated. Thus, quantification of the variables in the context of this approach must be viewed as an aid in comparison of different samples rather than as an absolute measure of these variables. For example, the crystallinity index was defined in terms of the ratio of the integral between 86 and 92 ppm to the integral between 80 and 92 ppm. This, of course, ignores the fact that some component of the integral between 80 and 86 ppm must be attributed to surface crystalline domains. In the same way, the relative amounts of the I_α and I_b forms are defined in terms of arbitrarily chosen partitions of the integral between 103.3 and 106.4 ppm. The part between 104.4 and 105.4 is taken as a measure of the I_α component and the sum of the parts between 103.3 and 104.4 ppm and between 105.4 and 106.4 ppm as a measure of the I_b component; this measure of the partition between the I_α and I_b forms clearly ignores the contribution of the noncrystalline component underlying all of the resonances associated with C-1.

The procedures developed were applied to the analysis of a variety of wood pulps.[98] While they indicate a higher level of the I_α form than had been suggested in previous studies of wood pulps, the

relative proportions of the I_α and I_b forms seemed to be the same in most of the pulps, with the major differences between them appearing to reside in the variation in the amorphous fraction.

In later studies[99,100] that are more significant to fundamental understanding of the nature of native celluloses, the spectra of some highly crystalline celluloses were analyzed by nonlinear least-squares fitting with a combination of Lorentzian and Gaussian functions. These analyses led to the conclusion that the *Halocynthia* celluloses do in fact contain a limited amount of the I_α form, contrary to earlier reports that it is exclusively the I_b form. Furthermore, the analyses detected within the envelope of the C-4 resonances give some indication of the occurrence of a "paracrystalline" component. On the basis of spin relaxation measurements, this component was judged to be somewhat more mobile than the I_α and I_b components but clearly less mobile than the amorphous fraction or the fractions of crystalline cellulose at the surface of microfibrils. On this basis it was described as a paracrystalline core structure.

The application of nonlinear least-squares decomposition into a combination of Lorentzian and Gaussian functions incorporates a significant element of empiricism as it requires judgments to be made with respect to which parameters are to be adjusted to achieve the best fit. In addition as noted earlier, the use of Lorentzian functions is not based on any fundamental understanding of the phenomenology of spectral acquisition in solid-state ^{13}C NMR. Yet, in the comparison of the native celluloses, it represents a useful tool that can provide insights into the differences between the types of order and organization that prevail in the different native forms. This is well illustrated in Table 1, taken from Larsson et al., [100] wherein they report the variation of the types of order among a variety of the naturally occurring forms of cellulose.

Table 1 Quantitations made by nonlinear least-squares fitting of the C-4 region in the ^{13}C NMR spectra. All values are relative intensities in percent. and values given in parentheses are standard errors.[100]

Cellulose source	*Cellulose* I_α	*Cellulose* I_β	*Paracrystalline cellulose*	*Surface cellulose*	*Amorphous cellulose*	*Hemicellulose and cellulose oligomers*	*Unassigned signal intensity*
Wood	9.1 (0.7)		31.1 (0.7)	6.0 (0.8)	53.7 (0.9)		
Cotton	4.2 (0.5)	27.6 (1.8)	33.1 (1.3)	7.8 (1.2)	24.1 (2.0)	3.1 (1.6)	
Halocynthia	9.6 (1.8)	61.2 (11.6)	12.5 (3.5)		16.7 (1.6)		
Cladophora	51.7 (1.1)	29.4 (1.6)	4.5 (0.8)		11.4 (1.0)		3.1 (0.4)
Valonia	55.9	25.1	7.0		7.0		5.0

The types of native celluloses represented cover the wide range of states of aggregation that occur naturally. The most highly crystalline are those of *Halocynthia, Cladophera,* and *Valonia,* all of which have microfibrils that have lateral dimensions of 20 nm or more. Among these it appears that *Valonia* and *Cladophera,* both algal celluloses, are the most crystalline in that they have the lowest content of paracrystalline and amorphous components, and for both of them the I_α form is the dominant component although the content of the I_b form ranges from 45 to 60% of that of the I_α form. The *Halocynthia* cellulose appears to have a somewhat higher level of the paracrystalline and amorphous fractions, and the partition between the I_α and I_b forms is clearly shifted much more toward the I_b form. For all three of the highly crystalline celluloses, the value of the surface cellulose content is not reported. It is clear that the resonances at 83.2 and 84.1 ppm, which are associated with the different ordered celluloses at the surfaces of the microfibrils, are within the limits of error.

The results for cotton are particularly interesting because almost half of the intensity of the C-4 resonance that had earlier been regarded as the crystalline component[73,75,84] is reported by Larsson *et al.* to be in the category of the paracrystalline form.[100] The surface celluloses are clearly measurable although a small fraction, and the amorphous fraction is almost as large as the paracrystalline fraction.

The results for wood perhaps show most clearly that this different approach to analysis of the spectra provides a different perspective. Here the I_α and I_b forms together represent a fraction that is one-third that of the paracrystalline form, which is the largest fraction of cellulose. The surface celluloses also are clearly measurable. No value is reported for the amorphous fraction, which is presumably included in the fraction labeled "hemicellulose and cellulose oligomers."

Though the partitioning of the cellulose content in Table 1 cannot be regarded as indicating absolute values of the different fractions, it represents an important step forward in recognizing the different categories of order with respect to the level of molecular mobility associated with them. Furthermore, it highlights the fact that the celluloses which are produced in different functional contexts in different organisms can have a wide range of variation in their states of aggregation and

the levels of mobility associated with these states. Thus, it is clear that the mechanisms for formation of cellulose can result in different types of organization at the nanoscale level and, as a consequence, a wide range of properties.

3.16.3.2.3 Electron microscopic studies

The use of electron microscopy in the study of celluloses. particularly in their native state, has resulted in important advances beginning with investigations that were undertaken at the time of introduction of the earliest electron microscopes. The early work has been ably reviewed by a number of authors.[101,102] Of particular note among these is the coverage of the subject in the treatise by Preston.[8] Here the author focuses on studies that have been important to advancing understanding of the structure of cellulose at the submicroscopic level.

The earliest and most significant observations, from a structural perspective. were those by Hieta *et al.* [103] where they applied a staining method incorporating a chemistry that requires the presence of reducing end groups. They observed that when whole microfibrils of *Valonia* were viewed, only one end of each microfibril was stained. This clearly indicated that the molecular chains were parallel as the reducing ends of the cellulose chains occurred together at one end of the fibrils. Had the structure been one with an antiparallel arrangement of cellulose chains, it would have been expected that the reducing end groups would occur with equal frequency at both ends of the microfibril with the result that both ends would be equally stained.

The conclusions of Hieta *et al.* [104] were independently confirmed by another method introduced by Chanzy and Henrissat wherein the microfibrils were subjected to the action of a cellobiohydrolase that is specific in its action on the nonreducing ends of the cellulose chains. They observed a clear narrowing of the tips of the microfibrils to a triangular form at only one end of each microfibril. In this instance the action was at the nonreducing ends but the observations were equally convincing evidence that the chains are aligned in a parallel arrangement in these microfibrils.

These early studies were focused on microfibrils from algal celluloses that, because of their larger lateral dimensions, could be more easily visualized in detail. More recently, the technique of specific staining of reducing end groups was adapted for application to cotton microfibrils by Maurer and Fengel.[105] In addition to application of the technique to examination of native cellulose, Maurer and Fengel applied the method to examination of microfibrils of mercerized cellulose (cellulose II) for which they also observed staining at only one end of the microfibrils. This last observation, which indicates a parallel chain structure in cellulose II, is very much in contrast to the crystallographic models that point to an antiparallel structure for this form of cellulose. It reinforces the view that the structure of cellulose II still has many uncertainties associated with it, in spite of the many theoretical analyses that have attempted to rationalize the antiparallel form.

In yet another important set of investigations by Sugiyama *et al.* [106-108] reported at approximately the same time, it was demonstrated that lattice images could be recorded from the microfibrils of *V. macrophysia*. The first images captured were based on lateral observation of the microfibrils;[106,107] an example is shown in Figure 16. Later, the techniques were refined to allow the acquisition of lattice images of cross-sections of microfibrils[108] as shown in Figure 17. The significance of these observations is that it is now possible to demonstrate conclusively that the microfibrils are uniform in formation and that there is no evidence that they are constituted of smaller subunits which aggregate together to form the individual microfibrils that are observed in the electron micrographs. Thus, the observations resolved some of the questions that had arisen earlier concerning the interpretation of electron micrographs of native celluloses;[8,9] the findings of Sugiyama *et al.* were the first direct evidence that the approximately 20 × 20 nm cross-sections were not composed of distinguishable smaller subunits. It should be noted, however, that the electron diffraction processes responsible for formation of the lattice images are dominated by organization of the heavy atoms in the molecular chains and would be insensitive to any nonuniformity in the hydrogen bonding patterns within the interior of the 20 × 20 nm fibrils. The homogeneity of the microfibrils revealed in the lattice images is an issue that needs to be revisited in the context of discussions of biogenesis, for in each instance the homogeneous crystalline domains clearly include a much larger number of cellulose chains than could possibly arise from the individual membrane complexes associated with the biogenesis of cellulose.

Later studies by Sugiyama *et al.* were based on electron diffraction and were directed at addressing questions concerning the nature of the differences between the I_α and I_b forms of cellulose. In a landmark study, [109] electron diffraction patterns were recorded from *Valonia macrophysia* in both its

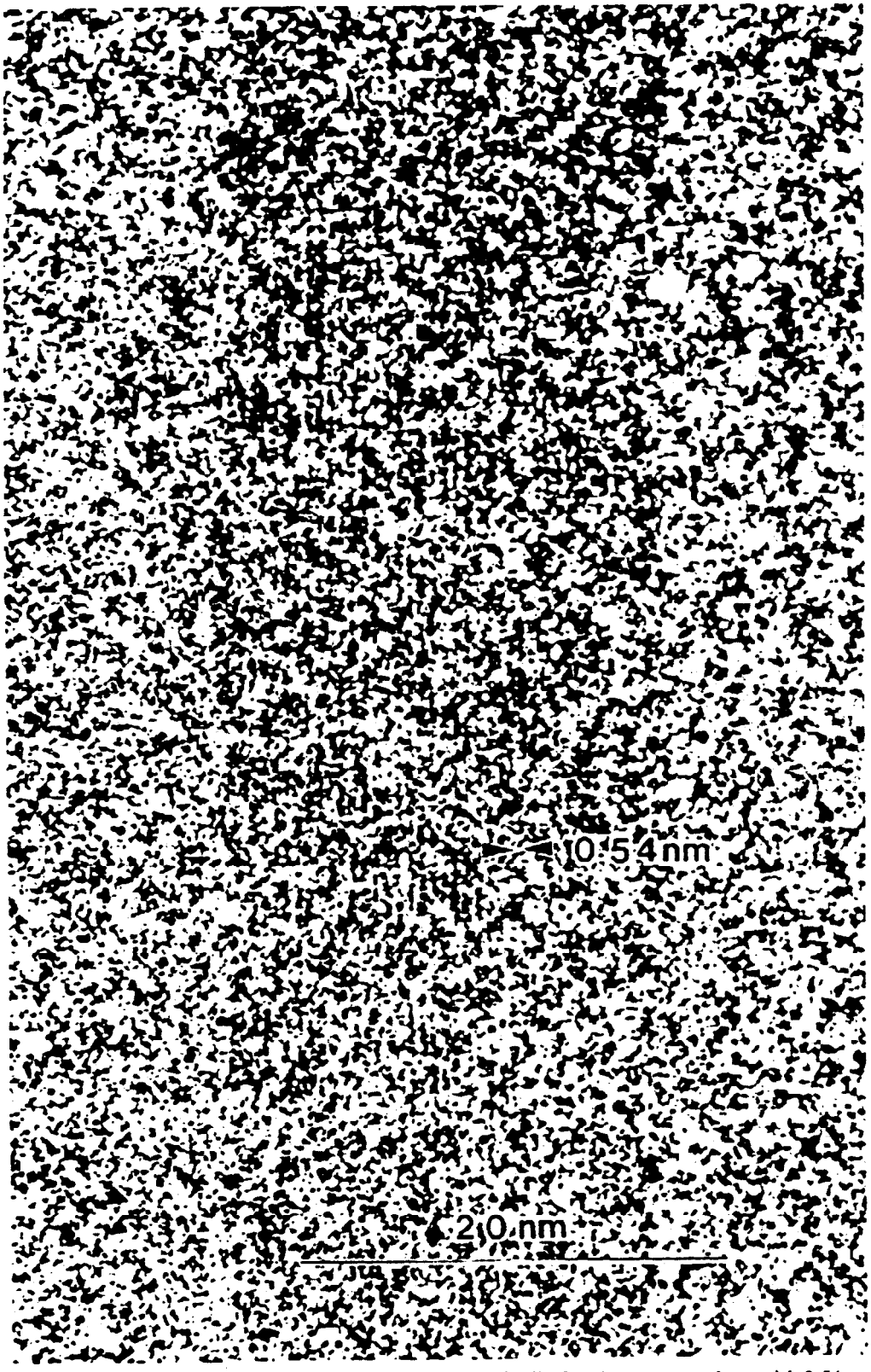


Figure 16 An enlarged photomicrograph of a cellulose microfibril of *Valonia macrophysa* with 0.54 nm lattice lines clearly visible (after Sugiyama[107]).

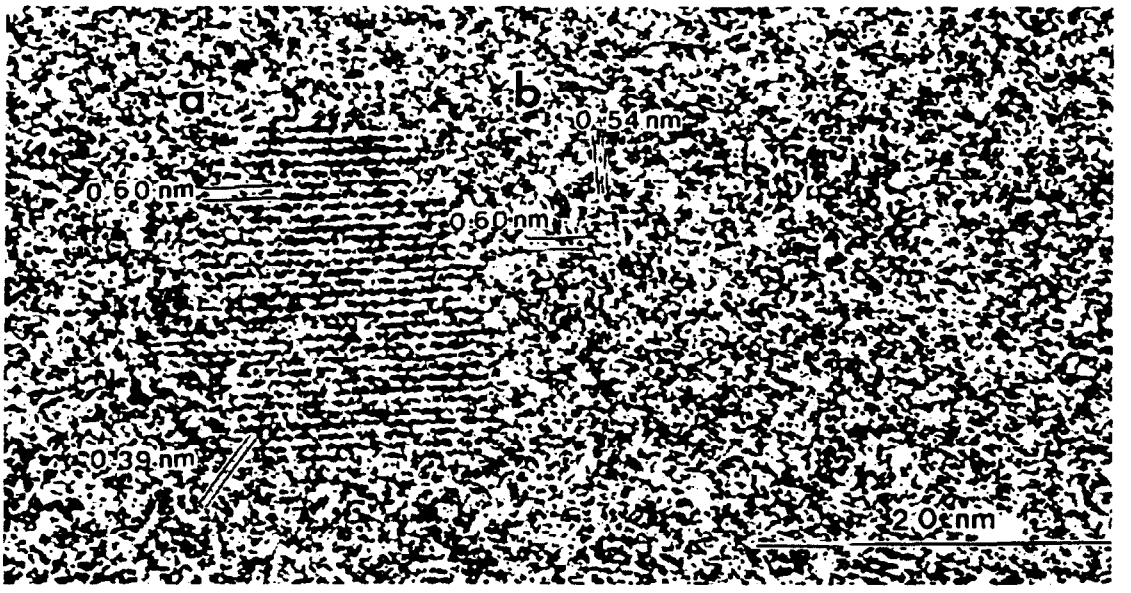


Figure 17 Lattice images od the lateral faces of microfibrils of *Valonia macrophysa* showing more than two sets of lattice fringes indicated (after Sugiyama[108]).

native state, wherein the I_α and I_b forms occur in their natural relative proportions, and after annealing using the process first reported by Horii and co-workers,[80] which converts the I_α form into the I_b form. The native material, which is predominantly the I_α form, was shown to produce a complex electron diffraction pattern similar to that which had earlier led Honjo and Watanabe to propose an eight-chain unit cell. In sharp contrast, the annealed sample, which is essentially all of the I_b form, produced a more simple and symmetric pattern that could be indexed in terms of a two-chain monoclinic unit cell. The observed patterns are shown in Figure 18. Figure 18(a) shows the diffraction pattern of the native forms, while Figure 18(b) shows how the diffraction pattern is transformed upon annealing. It is the latter that is identified with the I_b form and which has been interpreted to indicate a monoclinic unit cell. Figures 18(c) and 18(d) are schematic representations of the spots in the diffraction diagrams (a) and (b) and show more clearly how the diffraction pattern is transformed by annealing; the spots marked with arrows are those that disappear upon annealing. Upon separating the diffraction pattern of the I_b form from the original pattern, it was possible to identify the components of the original pattern that could be attributed to the I_α form, and it was found to correspond to a triclinic unit cell.

In this first report concerning the differences between the diffraction patterns of the I_α and I_b forms. the positioning of the chains within the monoclinic unit cell associated with the I_b form was left open. Two possibilities were regarded as consistent with the diffraction patterns, the first with the twofold screw axes coincident with the molecular chains, the second with the twofold screw axes between the chains. Both possibilities were consistent with the occurrence of nonequivalent anhydroglucose units. The triclinic unit cell associated with the I_α form was also viewed as consistent with two possibilities, the first a two-chain unit cell and the second an eight-chain unit cell similar to the one first proposed by Honjo and Watanabe.[110]

In a later study by Sugiyama *et al.*,[111] the possibilities were narrowed. It was stated that the monoclinic unit cell corresponding to the I_b form was viewed as one wherein the chains were coincident with the twofold screw axes. It was also indicated that the pattern of the triclinic unit cell corresponding to the I_α form appeared consistent with a unit cell with only one chain per unit cell. In both instances the rationale for these determinations was not presented.

Another interesting group of observations reported in the second electron diffraction study by Sugiyama *et al.*[111] was interpreted as evidence of the occurrence of the two forms of cellulose in separate domains within the same microfibrils. It was reported that the subsets of reflections associated with the two different forms could be observed separately or in combination along the length of an individual microfibril within domains 50 μm from each other. This set of observations was interpreted as indicating an alternation between the I_α and I_b forms along the length of an individual microfibril. Such an interpretation of course raises questions concerning the processes of biogenesis particularly since the relative proportions of the two forms of cellulose have been found to be invariant for a particular species as long as the procedures used for isolation of the cellulose

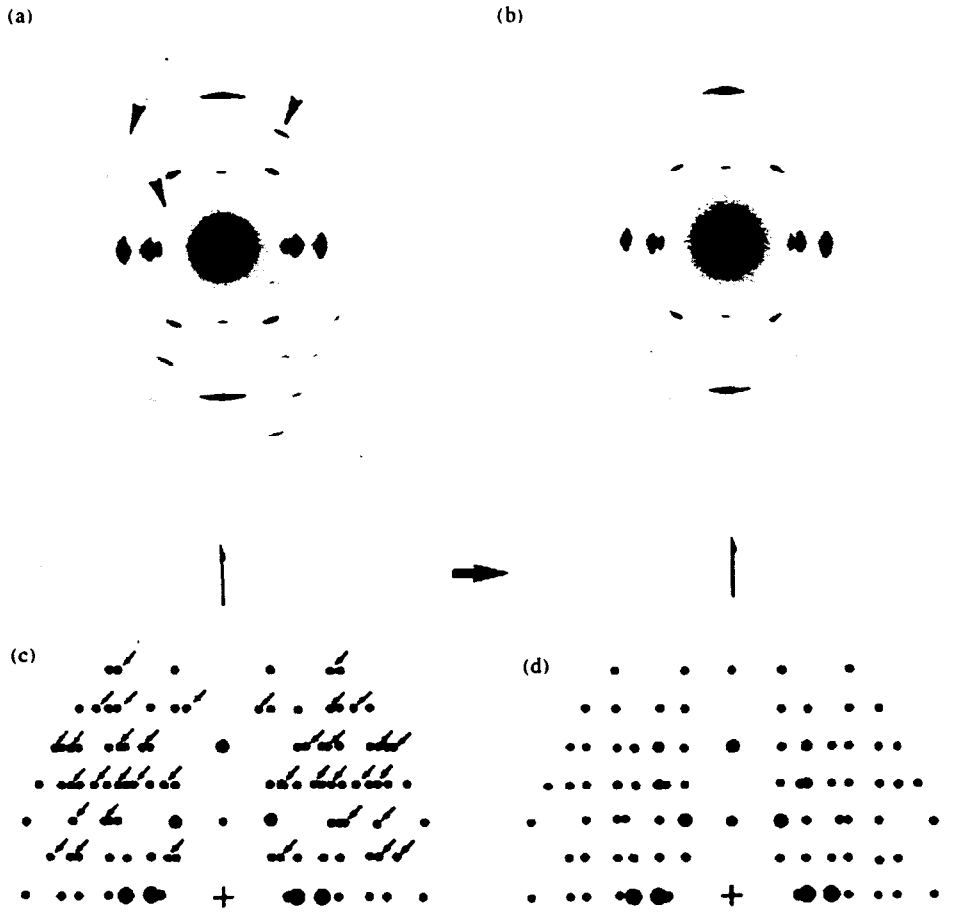


Figure 18 Typical electron diffraction patterns of *Valonia macrophysa* before (a) and after (b) annealing. In the schematic representations of the patterns the spots marked with arrows correspond to the reflections that disappear during the annealing treatment (after Sugiyama[109]).

do not incorporate exposure to conditions that can result in transformation of the I_α form into the I_b form.

The observation of different domains producing different diffraction patterns along the same microfibril can be envisioned as arising in two ways. The first is the possibility that the microfibril which was used to acquire the diffraction patterns had a limited amount of curvature or twist to it so that the angle between the electron beam and the unit cell axes was not constant. This could result in differences of relative intensities of diffraction spots from different planes and, given the short duration of the exposures, result in an unintended editing of the diffraction patterns. Thus, only those diffraction spots that are intense enough to be observed at a particular angle will be detected, while weaker ones go unseen. For example, if the lattice structure first suggested by Honjo and Watanabe[110] is the true one characteristic of the algal celluloses, diffraction patterns observed at different angles would result in different degrees of enhancement of the different subsets of the total diffraction pattern. This would also be true if the three-dimensional organization of the chains is more appropriately viewed as a superlattice. Indeed it is possible that the lattice structure first proposed by Honjo and Watanabe[110] represents the unit cell of such a superlattice.

Such an interpretation of these observations is consistent with earlier observations by Chanzy *et* al., [112] wherein an electron microscopic image of microfibrils of algal celluloses was formed by use of a technique based on diffraction contrast. It resulted in images of the algal microfibrils that had alternating dark and bright domains which appeared to be of the order of 50 nm in length. This suggests that the Bragg angle associated with a particular set of reflections is not likely to be coherently ordered relative to the electron beam in domains that are more than 50 nm in length. Given that the coherence of orientation relative to the electron beam was not found to extend beyond 50 nm, it would appear unlikely that it would remain invariant over a distance of 50 μm.

An alternative interpretation of the observation is that alternation of the I_α and I_b forms is real. as proposed by Sugiyama et *al.*, [111] and it reflects an assembly process that is not yet sufficiently well understood. It has been suggested that mechanical stress can facilitate transformation of the I_α form into the I_b form and that formation of the I_b form may arise from mechanical deformations of the fibrils in the course of deposition; as they emerge from the plasma membrane they are required, in most instances, to be bent to be parallel with the plane of the cell wall. If this is indeed the source of the reported alternation of the I_α and I_b forms along the microfibril, it would raise questions concerning the uniqueness of the balance between the I_α and I_b forms that seems to be characteristic of particular species.

3.16.3.3 Computational Modeling

Computational modeling has become an important aid in advancing understanding of complex molecular systems. Molecular modeling methods have allowed exploration of many different factors and interactions at the molecular level and the degree to which they may contribute to the phenomenology of different molecular systems. Computational modeling has found particular favor in the analyses of large molecules of biological origin, and of course, cellulose and its oligomers have attracted some attention in this arena. It is valuable to briefly review some of the efforts directed at advancing the understanding of cellulose because, in addition to providing insights regarding the contributions of different classes of interactions, they illustrate the reality that the results of analyses can often be the consequences of assumptions and premises introduced at the outset, rather than conclusions that can provide definitive answers to questions under exploration. It has been the author's experience that conclusions concerning the structures that are most favored for different celluloses change when new sets of potential functions are introduced into the computational programs and as different approaches to finding the most favored structures are adopted.

As alluded to earlier, the analysis by Rees and Skerret[50] was one of the first computational efforts to explore the constraints on the freedom of variation inherent in the structure of cellobiose. It relied on a potential function that is focused on van der Waals interactions in order to establish the degree to which domains within y/f space may be excluded by hard sphere overlap. The key finding was that approximately 95% of y/f space was indeed excluded from accessibility on the basis of hard sphere overlaps that were unacceptable in the sense that they required particular atoms associated with the region of the glycosidic linkage to be significantly closer to each other than the sum of the van der Waals radii. Upon mapping the energy associated with allowable conformations, they found that the two regions indicated by the solid line contours in Figure 2 represented energy minima close to the conformations defined by the twofold helical constraint. The boundaries of the acceptable region are not very far removed from the domains within the contours; the region between the two domains along the twofold helix line (n = 2) was not excluded by hard sphere overlap but it did represent a saddle point in the potential energy surface.

The next group of computational studies did incorporate hydrogen bonding energies as well as the van der Waals interactions. Whether they exhibited the double minima, and the degree to which the double minima were pronounced, depended in large measure on the relative weighting given to the different types of nonbonded interactions. In many, particularly those relying on the potential energy functions incorpcorated in the linked atom least-squares (LALS) programs, the weighting was based on fitting the potential functions to optimize the match between computed structures for small molecules for which the crystal structures were known from crystallographic studies. The results were that some showed very shallow minima off the twofold helix line;[51] the twofold helical structures were then rationalized on the basis that the departure represented a small difference in the energies that were regarded as within the error of the computation. When the criterion for quality of fit is chosen as a global minimum of the potential energy, without attention to the fact that it may incorporate unacceptable hard sphere overlaps, the results of the computational analysis can be misleading.

Later studies did not incorporate disproportionate weighting of the different types of nonbonded interactions,[57,58] and the result is perhaps best illustrated in the mapping of the potential energy for cellobiose shown in Figure 19, taken from the study of cello-oligomers by Henrissat *et al.* [58] In this instance, for the purposes of visualization, the y/f map presents the mirror image of the potential energy surface computed for cellobiose. While well more than two minima are shown in this mapping. it should be noted that only the two corresponding to the crystal structures of cellobiose and methyl-*b*-cellobioside, marked by arrows, are within the boundaries established in the analysis by Rees and Skerret[50] described earlier. The other minima correspond to conformations that are

more favorable to hydrogen bonding, but with relatively high energies associated with the van der Waals interactions pointing to severe hard sphere overlap.

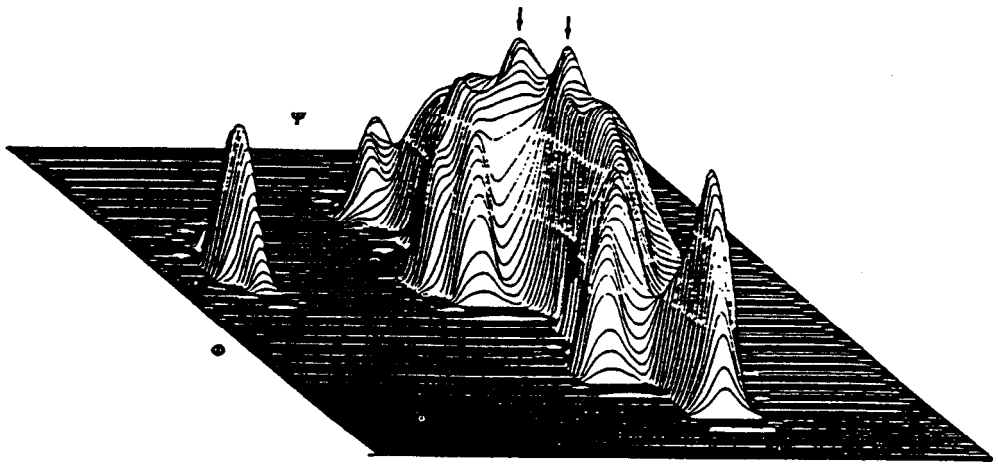


Figure 19 Perspective drawing of the three-dimensional shape of the mirror image of the conformational energy well for the full angular range of Φ and Ψ. The volume was constructed using the following scheme: $V(\Phi,\Psi) > 15$ kcal^{-1} mol^{-1}: $V_n(\Phi,\Psi) = 0$; $V(\Phi,\Psi) < 15$ kcal^{-1} mol^{-1}: $V_n(\Phi,\Psi) = -(V(\Phi,\Psi)-15)$, V being the energy expressed relative to the minimum. Proceeding from top to bottom of the three-dimensional shape, note the very low energy region (the arrows point towards the conformations observed for crystalline cellobiose and methyl-β-D-cellobioside). The 5–10 kcal^{-1} mol^{-1} energy contours correspond to the light gray region of the volume.[58]

There is little question that the results of modeling the cellobiose system reflect constraints on the freedom of rotation about the bonds of the glycosidic linkage. This is confirmed by the observation that the most favored structures indeed coincide with the crystal structures of cellobiose and methyl-*b*-cellobioside, both of which have been established on the basis of single crystal studies that are not in question. It is at the next higher level of structure in oligomers that the complexity of the potential functions and the multiplicity of adjustable and variable parameters begin to introduce considerable uncertainty about the results and the degree to which they reflect the real system. In the work of Henrissat *et al.*, [58] for example, when the analyses of the cellotriose and cellotetraose are carried out, the conclusions drawn are based on additional support from other sources of information on structure, and even with the additional information from other sources, the conclusions can only be presented as the most plausible, given the information available.

It is beyond the scope of this chapter to discuss the many computational modeling studies that have been undertaken wherein the subjects of the modeling exercise were chains of cellulose consisting of four or more anhydroglucose units. This is particularly so since a number of groups have entered the field of inquiry and are, in most instances, still refining their computational methods to factor in the eccentricities of cellulose as a molecule. Some observations on the different approaches are in order, however. The modeling programs are by and large of two types, although some investigators use combinations of the two approaches.

The first and the earlier of the two approaches adopted is the one identified as molecular mechanics, although it is in fact not a dynamic model but rather is based on a search for coordinates corresponding to potential energy minima. Its weakness is that it combines in one potential energy function interactions corresponding to all of the levels of structure alluded to earlier. Thus, it combines searches for equilibrium values of bond lengths together with the search for optimum packing of the molecules on the basis of different possible secondary and tertiary structures. In such computations it is to be expected that energy contributions of the interatomic bonds of the primary structure will be the dominant component of the potential energy. As a result, it is not unusual for the variations in energy associated with the hard sphere overlaps at the glycosidic linkage to be within the uncertainties of the calculations. In such cases, the search for global minima can point to structures with unrealistic local conformations.

The second approach, which is usually identified as molecular dynamics simulation, is based on integration of the equations of motion for all of the atoms in the system. Thus, they differ from the earlier studies in that the search is not for an equilibrium structure that represents a global minimum of potential energy but rather an exploration of the range of movement of the atoms that make up the structure under the influence of the prevailing forces of intramolecular and intermolecular

interactions, which are represented by complex potential functions that include terms appropriate for each of the interactions. The coordinates of the atoms are recorded at regular intervals and are used to determine the range of variation of the different internal coordinates within a molecule and the degree to which the molecules in an aggregate move relative to each other. Here again, the assumptions introduced at the outset can be primary determinants of the results. In some studies using this approach, for example, the force field used relies on the united-atom approximation wherein all aliphatic hydrogen atoms attached to a carbon atom are not included except as variations on the terms representing the carbon atoms in the potential function. While this is a useful and adequate approximation in the vast majority of instances where it is applied, its use with cellulose ignores the constraints of hard sphere overlaps on the glycosidic linkage. On the other hand, the programs designed to carry out the dynamic simulation have been modified to allow introduction of the constraint that the bond lengths are fixed in value, so that the energy associated with the primary structure no longer dominates the potential energy as is the case in the molecular mechanics calculations. It is therefore likely to be more successful in advancing understanding of the variability of the secondary and tertiary structures of cellulose.

Since many of the programs that have been used in the computational modeling studies of cellulose were, in the first instance, developed for a far wider range of molecular modeling studies, it is not surprising that they are not particularly well adapted to cellulose at the present time. If progress in computational modeling of complex molecules continues to advance as it has in the past decade, it is anticipated that within the next decade or two, some of these programs will be sufficiently flexible that allowances can be made for the special features in the structure of cellulose associated with the glycosidic linkage and that it will be possible to explore more widely the variability of secondary and tertiary structures associated with the primary structure that is fairly well established. It should then be possible to rationalize the wide variation in the patterns of aggregation of cellulose. The molecular dynamics simulation approach also has the advantage that it allows the exploration of interactions between cellulosic surfaces and associated liquids such as water.

The reader interested in exploring the different approaches to computational modeling of cellulose in greater depth can find them in the publications of Hardy and Sarko *et al.*, [113] Kroon-Battenberg *et al.*, [114,115] Heiner *et a1.*, [116,117] O'Sullivan and co-workers[118] and references cited therein.

3.16.3.4 Polymorphy in Cellulose

As was noted earlier, one of the discoveries from the diffractometric studies of cellulose undertaken during the middle period was that it can occur in a number of allomorphic forms in the solid state, each producing distinctive X-ray diffractometric patterns.[30] In addition to the cellulose II form, which has been discussed extensively, two other forms are well recognized; these are cellulose III and cellulose IV. It is of interest to consider them briefly because they reflect the capacity of cellulose to aggregate in a wide variety of secondary and tertiary structures, and because some of the higher plant celluloses produce diffraction patterns that are not unlike those of cellulose IV. Furthermore, they reflect the tendency for some of the celluloses to retain some memory of their earlier states of aggregation in a manner not yet understood.

Cellulose III is of little interest from a biological perspective except to the extent that its behavior may reveal some of the interesting characteristics of the native celluloses from which it can be prepared. It can be prepared from either native cellulose or from cellulose II by treatment with anhydrous liquid ammonia at temperatures near –30 °C. It produces distinctive X-ray patterns, Raman spectra, and solid-state ^{13}C NMR spectra. The most interesting characteristic it has is that it can be restored to the original form by treatment in boiling water. Because of this characteristic it is common to designate samples of cellulose III as either III_I or III_{II} to indicate both the source material and the form that will be recovered if the cellulose is boiled in water. In the case of native celluloses, transformation to the III_I form and back to the I form also has the unusual effect of converting those which have the I_α form dominant, such as those from algal sources, into forms in which the I_β form is dominant. This effect, first reported by Chanzy *et al.*, [119] is accompanied by partitioning of the algal microfibrils into smaller ones that are closer in lateral dimensions to those characteristic of higher plants. The solid-state ^{13}C NMR spectra also then appear more like those of the higher plants. No such changes have been reported for native celluloses in which the I_β form is dominant. These behaviors by cellulose III point to a memory effect with respect to the secondary and tertiary structures of cellulose that remains very much a mystery at the present time.

Cellulose IV is most often described as the high-temperature cellulose because it can be prepared by exposing the source cellulose to temperatures in the vicinity of 260 °C while it is immersed in

glycerol. In this preparation. its structure depends on whether it is prepared from cellulose I or cellulose II: hence the frequent designation as IV_I or IV_{II}. When prepared from cellulose I, it is first converted to the III_I form prior to treatment at high temperature in glycerol. When prepared from cellulose II, it can be produced directly from the II form or via the III_{II} form as an intermediate. However, in the instance of cellulose IV, there are no known procedures that allow restoration to the original form; the use of the different designations reflects some differences in the diffraction patterns observed from the two different forms. Furthermore, most of its reported preparations from native forms of cellulose have been from higher plant celluloses wherein the I_b form is dominant and the lateral dimensions of the native microfibrils are quite small; it is not at all clear that treatment of microfibrils of larger lateral dimensions such as those of *Valonia* or *Halocynthia* will result in such changes.

In addition to its preparation by heating at 260 °C in glycerol, cellulose IV has been recovered when cellulose is regenerated from solution at elevated temperatures. This has been observed with solutions in phosphoric acid regenerated in boiling water or in ethylene glycol or glycerol at temperatures above 100 °C. [120] It has also been observed upon regeneration from the dimethyl-sulfoxide-paraformaldehyde solvent system at elevated temperatures.[120] In yet another exploration of high-temperature effects on the aggregation of cellulose, it was found that when amorphous celluloses are prepared under anhydrous conditions and then induced to crystallize by exposure to water, exposure at elevated temperatures resulted in the formation of cellulose IV rather than cellulose II, which is the form usually obtained upon crystallization at room temperature.[121]

The samples of cellulose IV obtained through regeneration from solution were shown to have Raman spectra that could be represented as linear combinations of the spectra of celluloses I and II, suggesting that it may be a mixed lattice in which molecules with two different secondary structures coexist. This possibility is consistent with the earlier conclusion that both cellulose I and cellulose II have ribbon like structures that depart to a limited degree from a twofold helix but in different ways. It is not at all implausible that molecules so similar in shape could coexist in the same lattice.

One of the complications in interpreting observations of the occurrence of cellulose IV is that its X-ray diffraction powder pattern is very similar to that of cellulose I. The 020 reflection is nearly identical to that of cellulose I and the 110 and the $1\bar{1}0$ reflections collapse into a single reflection approximately midway between those of cellulose I. As a result, many of the less well-ordered native celluloses produce X-ray patterns that could equally well be interpreted as indicating cellulose I, but with inadequate resolution of the 110 and $1\bar{1}0$ reflections, or cellulose IV. They are usually characterized as indicating cellulose I because they represent celluloses derived from native sources. Indeed, when cellulose IV was first observed, it was thought to be a less ordered form of cellulose I.

The close relationship between cellulose IV and the native state is also reflected in reports of its observation in the native state of primary cell wall celluloses. These were observations based on electron diffraction studies of isolated primary cell wall celluloses.[122]

3.16.3.5 Oligomers and Structure at the Nanoscale Level

The oligomers of cellulose, which are essentially insoluble at the octamer level, have been the subject of a number of investigations because their secondary and tertiary structures seem to converge to structures similar to those of cellulose II. The X-ray powder patterns are quite similar to those of cellulose II,[123] and the vibrational spectra, both Raman and infrared, converge to that of cellulose at the tetramer.[48] The high-resolution solid-state ^{13}C NMR also converge to that of cellulose II, although at the tetramer the resonance of the anomeric carbon on the reducing end is still distinct.[58,69]

In one of the studies cited,[58] a multidisciplinary approach was used to explore the structures and led to the conclusion that the molecules of the celotetraose were arranged in an antiparallel manner, while another, based on X-ray diffractometry, concluded that a parallel structure is equally probable.[123] These studies also indicated that the dihedral angles of the glycosidic linkage departed systematically from those of a twofold helix. As noted earlier, the recent results of Maurer and Fengel[105] point to a parallel alignment of the molecular chains in the structure of cellulose II. Thus, in spite of many studies, the structures of both cellulose II and the oligomers remain the subject of many unanswered questions.

In relation to the structure of the polymer, the questions arising are the degree to which the observations on the oligomers can inform the interpretation of observations on the polymer. The

key question is the point at which convergence of the properties to those of the polymer can be anticipated. While it has been demonstrated for cellulose II only, it is anticipated that convergence of the vibrational spectra to those of the polymer at the level of the tetramer would also be true for celluloses I_α and I_b if it were possible to isolate tetramers in those forms. Thus, in terms of the basic structural units, the vibrational coupling patterns of four anhydroglucose units are sufficient to define the majority of bands in the polymer. These observations, when taken together with the observations of the X-ray powder patterns, suggest that the secondary and tertiary structures are well defined by the organization of eight chains over a length equivalent to that of the tetramer. That is, a volume defined by dimensions equal to twice the dimensions of the most simple unit cells in all three dimensions. The dimensions are of the order of 1.6-2.1 nm, and this level of structure will be defined as organization at the nanoscale level.

3.16.3.6 From Nanoscale to Microscale: Supramolecular Organization in Celluloses

Structure at the level next beyond that defined as nanoscale is the one at which most of the distinctive organization of native celluloses is most clearly expressed. It is also the level most readily accessible through electron microscopy. Here we define it as the microscale level and associate with it molecular organization at the scale ranging between 2 and 50 nm. We have already discussed some aspects of structure at this level in distinguishing the different classes of samples of native cellulose that have been the subject of the studies presented above in overview. We will discuss structure at the microscale level and its implications in greater detail in the next section.

Before summarizing our discussions of structure and considering those of native celluloses, it is important to examine a number of questions that arise at the transition point from nanoscale to microscale for any system, but which have not been adequately addressed in the case of cellulose. To the author's knowledge, these questions have not been addressed in prior reviews of the structure of cellulose. They are of two classes though not unrelated to each other. The first class deals with the degrees of departure from the approximation of the infinite lattice linearly extended in three dimensions implicit in the structural models at the nanoscale level. The second class is concerned with the validity of regarding the aggregated states of cellulose, particularly in its native state, as phases in the traditional sense.

In most native celluloses the departures from the infinite linearly extended lattice are evidently very close to the nanoscale level. It has already been noted and it is well established that lateral order rarely exceeds 6 nm except in the case of some algal celluloses and those of the tunicates, which can have lateral dimensions of up to 25 nm. In most instances curvature of the elementary fibrils begins to manifest itself at these levels. Both of these departures from the infinitely extended linear lattice are usually inherent in the nature of organization of the biological tissue of which the native cellulose is a constituent and an integral part. They are manifestations of the biological order of the particular native cellulose. Yet the effects of these departures from a linear lattice structure on both diffractometric and spectroscopic measurements are indistinguishable from the effects of random disorder as it might occur in systems that are not of a biological origin. In the context of cellulose science, this reality has most often been dealt with by defining that which is not ordered in a linear lattice as disordered or amorphous. Such an approach ignores the fact that in most states of aggregation, amorphous systems are usually assumed to be homogeneously disordered. It has been quite common in the literature on cellulose to attribute to the nonordered components of native cellulosic substances the properties of such amorphous substances. This is an issue that must be kept in mind as individual native celluloses are considered below.

The second class of questions is a logical consequence of the limitation of lateral dimensions. They are associated with the fact that when the lateral dimensions are as small as they are in most native celluloses, a significant fraction of the substance lies at the surface, so that the criterion for describing the aggregates as a separate phase is no longer adhered to. This criterion is that the amount of substance at the surface be vanishingly small in comparison with the amount of substance within the interior of the phase. [124] Given this circumstance, the application of stability criteria derived from traditional thermodynamic analyses must be approached with caution.

It is also appropriate at this point to address a fallacy commonly recurring in the literature with respect to the relative stability of celluloses I and II. Cellulose I is frequently regarded as the "metastable" form of cellulose because, upon regeneration from solution, cellulose II is usually the form recovered, and because upon swelling in strong caustic solution followed by deswelling, which is the process of mercerization, cellulose II is also recovered. Cellulose I, which has been prepared

by regeneration from cellulose solutions at elevated temperatures.[125,126] though the procedure is not easy to reproduce and there remain a number of factors involved in the regeneration that are not well understood.

3.16.3.7 Chemical Implications of Structure

It was noted earlier that an acceptable fit to the diffractometric data is not the ultimate objective of structural studies. Rather it is the development of a model that possesses a significant measure of validity and usefulness as the basis for organizing, explaining, and predicting the results of experimental observations. In the sections above, the new and evolving conceptual framework for describing the structures of cellulose was described in relation to spectral observations. It is also important to consider the degree to which the structural information that has been developed above may be useful as the basis for advancing understanding of the response of celluloses to chemical reagents and enzyme systems. It is useful first to review briefly past work directed at rationalizing the responses to such agents.

The vast majority of studies of the chemistry of cellulose has been directed at the preparation of cellulose derivatives with varying degrees of substitution depending on the desired product. Sometimes the goal is to prepare a cellulose derivative that possesses properties which differ significantly from those of the native form: some derivatives are water-soluble, others are thermoplastic, and others still are used as intermediates in processes for the regeneration of cellulose in the form of films or fibers. At other times the objective is to introduce relatively small amounts of substitution to modify the properties of the cellulosic substrate without it losing its macroscopic identity or form such as fiber or microcrystalline powder or regenerated filament or film. All such modification processes begin with a heterogeneous reaction system which may or may not eventually evolve into a homogeneous system as the reaction progresses. Thus, in all chemical investigations that begin with cellulose as one of the ingredients, issues associated with heterogeneous reaction systems arise. Understandably, the one that has been dominant in most investigations is the question of accessibility of the cellulose.

A variety of methods has been developed to relate accessibility to microstructure. Almost all of them begin with the premise that the cellulose can be regarded as having a crystalline fraction and a disordered or amorphous fraction. It is then assumed that the amorphous or disordered fraction is accessible while the crystalline fraction is not. In some instances, the portion of crystalline domains at the surface is regarded as accessible and it is therefore included as part of the disordered fraction. In other instances, the particular chemistry is thought to occur only in the disordered fraction and the surfaces of crystalline domains are not included. The different approaches have been reviewed by Bertoniere and Zeronian,[127] who regard the different approaches as alternative methods for measuring the degree of crystallinity or the crystalline fraction in the particular celluloses.

A number of different chemical and physical approaches is described by Bertoniere and Zeronian. The first is based on acid hydrolysis followed by quantification of the weight loss due to dissolution of glucose, cellobiose, and the soluble oligomers.[128] This method is thought to incorporate some error in quantification of the crystalline domains because chain cleavage upon hydrolysis can facilitate crystallization of chain molecules that had been kept in disorder due to entanglement with other molecules. Another method is based on monitoring the degree of formylation of cellulose when reacted with formic acid to form the ester.[129] In this method, progress of the reaction with cellulose is compared with a similar reaction with starch, which provides a measure of the possibility of formylation in a homogeneous system wherein the issue of accessibility does not arise.

In another method developed by Rowland and co-workers,[130-133] accessible hydroxyl groups are tagged through reaction of the particular cellulose with *N*,*N*-diethylaziridinium chloride to produce a DEAE-cellulose. This is then hydrolyzed, subjected to enzyme action to remove the untagged glucose, silylated, and subjected to chromatographic analysis. This method has the added advantage that it can be used to explore the relative reactivity of the different hydroxyl groups. It is usually observed that the secondary hydroxyl group on C-2 is the most reactive and the one at C-3, the least reactive, with the primary hydroxyl at C-6 having a reactivity approaching that of the group on C-2 under some conditions. Here, of course, steric effects are also factors in these substitution reactions.

Among the physical methods discussed by Bertoniere and Zeronianl[127] are those based on sorption and solvent exclusion. One of the earliest studies relying on the use of sorption as a measure of accessibility was the classical study by Mann and Marinan[33] wherein deuterium exchange with

protons was monitored. The cellulose was exposed to D_2O vapor for a period sufficient to attain equilibrium and then the degree of exchange was measured by observation of the infrared spectra. Comparison of the bands associated with the OD stretching vibration to those associated with the OH stretching vibration provided a measure of the relative amounts of accessible and inaccessible hydroxyl groups. Another approach to monitoring availability to adsorbed molecules is measurement of moisture regain upon conditioning under well-defined conditions as described by Zeronian and co-workers.[134]

The method of solvent exclusion has been used to explore issues of accessibility on a somewhat larger scale. An approach pioneered by Stone and Scallan[135] and Stone *et al.* [136] relied on static measurement using a series of oligomeric sugars and dextrans of increasing size to establish the distribution of pore sizes in different preparations of a variety of native celluloses. These, however, provide measures of accessibility beyond the nanoscale and were indeed designed to explore pores closer to the lower range of what we have defined as the microscale.

While methods for characterizing celluloses on the basis of their accessibility have been useful, they do not provide a basis for understanding the level of structure at which the response of a particular cellulose is determined. This follows from the rather simple categorization of the substrate cellulose into ordered and disordered fractions corresponding to the fractions thought to be crystalline and those that are not. This classification does not allow discrimination between effects that have their origin at the level of secondary structure and those that arise from the nature of the tertiary structure. Thus, in terms of chemical reactions, this approach does not facilitate separation of steric effects that follow from the conformation of the molecule as it is approached by a reacting species, from effects of accessibility which are inherently a consequence of the tertiary structure.

The possibility of advancing understanding of the chemical implications of structure is best illustrated in the context of hydrolytic reactions. Among the patterns that emerge fairly early in any examination of the published literature on acid hydrolysis and enzymatic degradation of cellulose are the many similarities in the response to the two classes of hydrolytic agents. In both instances a rapid initial conversion to glucose and cellodextrins is followed by a period of relatively slower conversion, the rate of conversion in the second period depending on the prior history of the cellulosic substrate. In general, the nonnative polymorphic forms are degraded more rapidly during this second phase. In addition, it is found that the most crystalline or highly ordered of the native celluloses are particularly resistant to attack, with the most highly crystalline regions converted much more slowly than any of the other forms of cellulose.

The relationship of the patterns of hydrolytic susceptibility to the range of conformational variation discussed above can be interpreted in terms of contrast between the states of the glycosidic linkage in cellobiose and *b*-methylcellobioside. The differences between the states that are likely to contribute to the differences in observed reactivity are of two types. The first is differences in the steric environment of the glycosidic linkage, particularly with respect to activity of the C-6 group as a steric hindrance to, or as a potential promoter of, proton transfer reactions, depending on its orientation relative to the adjacent glycosidic linkage. The second type of difference is electronic in nature and involves readjustment of hybridization of the bonding orbitals at the oxygen in the linkage. It is worthwhile examining the potential contribution of each of these effects.

The effect of steric environment emerges most simply from examination of scale models of the cello-oligodextrins. They reveal that when C-6 is positioned in a manner approximating the structure in *b*-methylcellobioside, the methylene hydrogens are so disposed that they contribute significantly to creation of a hydrophobic protective environment for the adjacent glycosidic linkage. If, however, rotation about the C-5—C-6 bond is allowed, the primary hydroxyl group can come into proximity with the linkage and provide a potential path for more rapid proton transfer.

If, as suggested earlier on the basis of spectral data, the orientation of some C-6 groups in native cellulose is locked in by their participation in a bifurcated hydrogen bond to the hydroxyl group on C-3, they may contribute to the higher degree of resistance to hydrolytic action. Access to the linkage oxygen would be through a relatively narrow solid angle, barely large enough to permit entry of the hydronium ions that are the primary carriers of protons in acidic media.[137] If, on the other hand, the C-6 group has greater freedom to rotate, as is likely to be the case in cellulose II, the hindrance due to the methylene hydrogens can be reduced and, in some orientations, the oxygen of the primary hydroxyl group may provide a tunneling path for transfer of protons from hydronium ions to the glycosidic linkage. This would result in greater susceptibility of nonnative celluloses to hydrolytic attack.

The hypothesis concerning steric effects in acid hydrolysis has as its corollary the proposal that the role of the C_1 component in cellulase enzyme system complexes is to disrupt engagement of

the C-6 oxygen in the bifurcated intramolecular hydrogen bond and thus permit rotation of the C-6 group into a position more favorable to hydrolytic attack.

The key role of C-6 in stabilizing the native cellulose structures is supported by findings concerning the mechanism of action of the dimethylsulfoxide-paraformaldehyde solvent system for cellulose, which is quite effective for solubilizing even the most crystalline of celluloses. The crucial step in the mechanism that has been established for this system is substitution of a methylol group on the primary hydroxyl at the C-6 carbon.[138,139]

The effect of conformation on the electronic structure of the linkage is also likely to be a factor with respect to its susceptibility to hydrolytic attack. Though there is no basis for anticipating the directions of this effect at this time, it is well to consider it from a qualitative perspective. First it is clear that the hybrids of oxygen orbitals involved in the bonds to carbon must be nonequivalent because the bond distances differ to a significant degree.[52,53] The angle of approximately 116 imposed on the linkage is likely to result in greater differences between the bonding orbitals and the lone pair orbitals than might be expected in a typical glycosidic linkage that is free from strain. Among themselves, the lone pair orbitals are likely to be different because of their different disposition with respect to the ring oxygen adjacent to C-l in the linkage; the differences may be small and subtle, but they are no less real. Given these many influences on the nature of the hybridization at the oxygen in the linkage; it seems most unlikely that they would remain unaltered by changes in the dihedral angles of the magnitude of the difference between cellobiose and *b*-methylcellobioside. Hence a difference in electronic character must be expected.

At present it is not possible to estimate the magnitude of the effects discussed, nor to speculate about the direction of the change in relative reactivity of the glycosidic linkage in the two different conformations. Yet it is clear that differences can be anticipated and they may be viewed, within limits of course, as altering the chemical identity of the glycosidic linkage as its conformation changes. It remains for future studies to define the differences more precisely.

The points raised with regard to the influence of conformation on factors that determine the pathways for chemical reaction have not been specific subjects of investigation because methods for characterizing secondary structure as apart from the tertiary structure have not been available. It has also been true that suitable conceptual frameworks have not been available for developing the questions beyond the levels of the order-disorder duality. With the development of the approaches outlined above for exploring and distinguishing between matters of secondary and tertiary structure, it is quite likely that in the years ahead it will be possible to achieve a higher level of organization of information concerning the chemistry of cellulose.

With respect to questions of tertiary structure, the key issue introduced by the new structural information, and one that has not been explored at all to date, is whether the different hydrogen bonding patterns associated with the I_α/I_b duality have associated with them differences between the reactivity of the hydroxyl groups involved. It is not clear at this time how experiments exploring such effects might be carried out so as to separate issues associated with the differences between the hydrogen bonding patterns from issues associated with differences in accessibility.

3.16.3.8 Cellulose Structures in Summary

Before turning to the discussion of biological aspects in the next section, it is helpful to summarize where studies at the nanoscale level stand at the present time, and to assess the degree of confidence with which one can use their conclusions as the basis for further discussion.

From crystallographic studies, based on both X-ray and electron diffraction measurements, it can be concluded that the secondary structures of native celluloses are ribbon-like conformations approximating twofold helical structures. Their organization into crystallographic unit cells remains uncertain, however. The monoclinic space group $P2_1$, with two chains per unit cell, has been proposed for both the earlier studies prior to discovery of the I_α/I_b duality in native forms, and also for the I_b form. The I_α form is thought to possess a triclinic unit cell structure. Some important questions remain regarding the degree to which these are adequate representations of the organization of the crystalline domains in native celluloses. The majority of crystallographic studies also point to parallel alignment of the cellulose chains in the native celluloses, and this conclusion has been confirmed by electron micrographic observations. Also for cellulose II, the structures derived from X-ray diffraction data suggest a ribbon-like secondary structure approximating twofold helical

organization and, in this instance, antiparallel alignment of the chains. For cellulose II though, the antiparallel proposal has been contradicted by recent electron micrographic observations. The unit cell organization of space group $P2_1$, with two chains per unit cell, has also been suggested for cellulose II, though the degree of confidence is even less than that with respect to the structures of cellulose I.

The early Raman spectroscopic studies clearly could not be reconciled with the premise that both cellulose I and cellulose II possess twofold helical conformations as the crystallographic studies had suggested. The Raman spectra, together with some corresponding infrared spectra, also pointed to the probability that the repeat unit of the structure of crystalline celluloses is anhydrocellobiose, so that alternating nonequivalent glycosidic linkages occur within each chain. To preserve the ribbon-like structural approximation, the different conformations of celluloses I and II were rationalized as incorporating glucosidic linkages that represent alternate left- and right-handed departures from the twofold helical structure, with those in cellulose II representing somewhat larger departures from the twofold helical conformation than those in cellulose I.

The introduction of high-resolution solid-state ^{13}C NMR spectral analyses into the study of celluloses resulted in resolution of one of the fundamental mysteries in the variability of native celluloses by establishing that all native celluloses are composites of two forms. These were identified as the I_α and I_β forms to distinguish them from the I_A and I_B categories that had been introduced more than three decades earlier to distinguish the celluloses produced by algae and bacteria from those produced by higher plants. The correspondence between the two classifications is that those in the I_A category have the I_α form as the dominant component, while those in the I_B category are predominantly of the I_β form. The nature of the difference between the I_α and I_β forms remains the subject of serious inquiry. Recognition of the I_α/I_β duality has facilitated a significant amount of additional research seeking to establish the balance between the two forms in a wide range of higher plant celluloses.

In later studies, the Raman spectra and corresponding infrared spectra indicated that the primary differences between the I_α and I_β forms of native cellulose were in the pattern of hydrogen bonding. Furthermore, the Raman spectra of the two forms raise questions as to whether the structures can possess more than one molecule per unit cell since there is no evidence of any correlation field splittings of any of the bands in the spectra of the two forms.

Electron microscopic studies relying on agents that act selectively either at the reducing or the nonreducing end groups of the cellulose chains have provided convincing evidence that cellulose chains are aligned parallel to one another in native cellulose. More recently, similar evidence has been presented supporting the view that alignment of the chains is also parallel in cellulose II. Other electron microscopic studies using the methods of lattice imaging have been used to demonstrate that the highly ordered microfibrils derived from algal celluloses represent homogeneous lattice structures with respect to the diffraction planes defined by the organization of their heavy atoms.

Electron diffraction studies carried out on algal celluloses after discovery of the I_α/I_β duality have been interpreted to indicate that the two forms may alternate along the length of individual microfibrils. These observations can also be interpreted as manifestations of the slow twisting about the long axis that has been observed in other studies of similar algal celluloses.

The possibility of the coexistence of the I_α and I_β forms within a superlattice structure has been suggested in the context of studies intended to mimic the conditions of biogenesis. These will be examined in greater detail in relation to the discussions of native celluloses and their biogenesis.

This discussion of structure has so far focused on issues of structure at the nanoscale level, identified as corresponding to domains that are of the order of 2 nm in dimension. Organization at the microscale level, defined as the range between 2 and 50 nm, requires consideration of a number of issues that have not been dealt with adequately in the literature on structures of cellulose. These include the well-recognized departures from a linear lattice, which have been generally regarded as measures of disorder when in fact they are more appropriately regarded as indicators of the nonlinear organization in a biological structure. Another issue arising at the microscale level is associated with the occurrence of significant fractions of cellulose molecules at the surface of the microfibrils of most native celluloses, particularly in the case of higher plants. It is a pertinent question as to whether the microfibrillar structure can be viewed as a separate phase in the traditional sense and whether criteria developed for the stability of homogeneous phases in the context of classical thermodynamics can have meaning when applied to native cellulosic structures. These issues arise in relation to discussions of native celluloses and their biogenesis.

Finally, the new developments with respect to structure, which facilitate separate though complementary focus on secondary and tertiary structures, provide a basis for exploring the relationships

between structure and reactivity in new ways. The framework for more detailed characterization of the states of aggregation should permit more systematic exploration of the influence of source and history of cellulosic samples on their entry into chemical and biological processes.

3.16.4 BIOLOGICAL ASPECTS

The biological aspects of cellulose have usually been incorporated briefly in most prior reviews of the chemistry of cellulose. In view of the advances with respect to the characterization of celluloses, particularly in their native forms, and the significant progress in understanding processes of biogenesis and biodegradation, it is particularly appropriate to devote a section to the biological aspects of cellulose in a treatise on natural products chemistry. It is well at the outset to present the perspective that informs our discussion. When viewed in the context of biology, it is increasingly obvious that cellulose can no longer be regarded as another semicrystalline polymer, the phenomenology of which needs to fit within the traditional paradigms of polymer science. Even though investigations of its nature and its derivatives contributed to foundations of the polymer hypothesis and were the basis for developing many of the principles of polymer science, it is now important to recognize its distinctive biological functions and acknowledge that it is a remarkable and unusual biological molecule with characteristics that allow the formation of unique cellulosic structures within particular species and often within different tissues of the same organism. Rather than viewing cellulose as a linear homopolymer capable of crystallization under different sets of conditions, the biological celluloses must be regarded as important constituents of living matter formed through self-assembly of this unusual molecule in a wide variety of forms that are unique to the organisms within which it occurs. In a wider range of biological contexts, cellulose is also capable of co-aggregating with other constituents of the tissues within which it is present. It is therefore important to adopt an approach that is not entirely reductive with respect to the role of cellulose in living matter. It is with this in mind that this section of the chapter begins with a discussion of the processes of biogenesis of cellulose, beginning with the formation of the primary structure through the biosynthetic pathways and continuing with the formation of microstructure which is at the heart of the differences between native celluloses from different biological sources and tissues.

The discussion of biogenesis is followed by a discussion of a few of the native forms of cellulose that have been the subject of extensive investigation. Some of these forms have been studied primarily because they facilitate fundamental understanding of the nature of native celluloses, while others have been characterized exhaustively because of their commercial importance. New among the former group are the celluloses formed by the tunicates, which are marine organisms that have simply been regarded as an interesting curiosity of little interest. The finding that these celluloses represent the only naturally occurring highly ordered forms that are almost pure I_b types has dictated that they receive greater attention than has been the case in the past.

Within the biosphere, the biogenesis of cellulose is balanced by a wide variety of processes that result in its conversion to other forms of biomass. To reflect this, Section 3.16.4.3 discusses the biological processes of disassembly or biodegradation of cellulose as these processes are part of the key to understanding the role of cellulose within the carbon cycle. The biodegradation or bioconversion of cellulose occurs in two major contexts. The first is the natural disassembly of plant matter that has come to the end of its natural life cycle; the disassembly is carried out by microorganisms that are particularly well adapted to this purpose. The second is in the context of the function of cellulose as an important nutrient in the diet of ruminants and herbivores. A different class of microorganisms that are symbiotic with the animal hosts appear to be adapted to this function. But entirely apart from the primary processes of recycling of the celluloses, it is also now increasingly recognized that the disassembly and reassembly of cellulose may be an important part of the morphogenetic processes of plant tissues during their formation and growth, thus adding yet another dimension to studies of the disassembly of cellulose.

Finally, from a systems perspective, it is in order to compare the levels of complexity that are recognized in discussions of the different aspects of the biological processes in which cellulose plays a central role. The questions arise, for example, whether the complexities of the systems that have evolved for the disassembly of cellulose are adequately reflected in our understanding of the processes for its assembly? To the extent that the processes of disassembly reflect diversity in the secondary and tertiary structures of the native celluloses, should consideration of these levels of structure be an integral part of the analyses of structure formation? Are investigations focused on the formation of primary structure adequate to the task of unraveling the mysteries of the formation of native

celluloses, or are the mechanisms of formation of secondary and tertiary structures as central to the answers as the processes of polymerization? These questions are more appropriately addressed at the conclusion of this section, but their statement at the outset will help place the discussions in perspective.

3.16.4.1 Biogenesis

The processes involved in the biogenesis of native celluloses are complex and highly coupled as formation of the primary structure at the nanoscale level is coordinated with the development of organization of secondary and tertiary structures at the microscale level. The primary structure is, by definition, the same for all cellulose-forming organisms, and it may well be that the processes of formation in different groups of organisms have much in common. It is not plausible, however, at the present time to speculate regarding the degree of commonality between the formation of primary structure in the celluloses that occur in the tunics of marine organisms and the processes of the plant kingdom. In contrast with primary structure, secondary and tertiary structures appear distinctive for particular species and, for all but unicellular organisms, can vary with tissue type. In higher plant cell walls, for example, the tertiary structures of the celluloses may differ among the different layers of the wall. In all instances wherein a secondary wall occurs, differences are observed between the primary and secondary walls, and in many plants the secondary wall is layered with significant variation in the organization of the celluloses within the different layers. For these reasons, the study of biogenesis has occurred at two different levels. The first is concerned with biosynthesis or the chemical pathways that lead to formation of the primary structure through the polymerization of glucose. At this level the issues are identification of the precursors and the enzyme systems involved in the processes of polymerization; investigations in this arena are primarily biochemical. At the next level, the focus is on ultrastructure, that is, the aggregated forms of native cellulose and the fibrillar organization that manifests itself at different levels within the morphological features characteristic of the particular cell wall or organism; studies addressing issues at this level rely more heavily on electron microscopy and seek to advance understanding of the secondary and tertiary structures and related phenomena at the microscale level. It is at this level also that many of the methods discussed in the previous section are applied. Though much of the investigation of celluloses at this level is biologically oriented and beyond the scope of this chapter, the self-assembly of cellulose during biogenesis and its modulation by particular chemical reagents are key to understanding the nature of native celluloses. Thus, the ultrastructural level is central to the chemistry of cellulose.

3.16.4.1.1 Biosynthesis

This discussion of biosynthesis will be confined to celluloses of the plant kingdom and related microorganisms that have been regarded as appropriate model systems. Studies of biosynthetic pathways have focused on investigation of model systems, with the expectation that much of what is learned of the biosynthesis of cellulose in the model systems will be common to the broader group of organisms that can produce cellulose. The system studied most often is the cellulose-producing bacterium *Acetobacter xylinum*. This system has two degrees of simplification relative to plant models. The first is that the processes of aggregation of the cellulose are more simple. The fibrils are not part of a cell wall but rather associate to form a pellicle that does not possess a higher level of order beyond the constraints associated with formation parallel to the surface of the culture medium. In most cellulose-producing organisms, in contrast, the cellulose is an integral part of a cell wall and its aggregation into fibrils is subject to additional constraints as it is coordinated with the development of higher levels of organization. The second advantage of *Acetobacter xylinum* over plant models is that it does not produce the b-(1-3)-linked glucan callose, formation of which is part of the wound response process of plants. Since the process of isolation of synthase enzymes obviously requires dramatic disruption of the integrity of the cells, plant systems usually shift to the formation of callose whenever the cells are subjected to enzyme isolation procedures. As a consequence, it has not been possible so far to isolate fully functional cellulose synthase enzyme systems from higher plants. The isolates instead produce glucans that are either the b-(1-3)-linked callose or include mixed linkages, most commonly the b-(1-3)- and b-(1-4)-linked polymer. These polymers also tend to aggregate in insoluble forms and have often been mistaken for cellulose on the basis of their insolubility and hydrolysis to produce glucose.

The biochemistry of cellulose synthesis by *Acetobacter xylinum* has been investigated in considerable detail by Benziman, Delmer, and their associates.[140-145] Important advances have been facilitated by isolation of the cellulose synthase from this organism in a highly active form.[143-145] This has allowed examination of the different factors involved in the synthesis and has established an important point of departure for additional studies exploring cellulose synthesis by other more complex systems. Availability of the synthase has provided access to the genetic encoding for the enzyme system and its constituents, allowing exploration of the degree to which related structures may occur in appropriate protein extracts from other organisms. This, in turn, has led to identification of such genes in a number of important plant systems.

It has been recognized for some time that the precursor in cellulose synthesis is a sugar nucleotide, although there had been some uncertainty regarding the specific one involved in cellulose.[146] Following considerable accumulation of evidence in support of this view, it is now generally accepted that, while other nucleotides may be involved in the synthesis of other polysaccharides or oligosaccharides, UDP glucose (uridine 5'-(α-D-glucopyranosyl pyrophosphate)) is the carrier of glucose for the formation of cellulose.[147] Assembly of the polymer from the glucose portion of UDP glucose is accomplished through a sequence of coordinated steps that seem to be orchestrated by the synthase complex. Northcote[148] has suggested that, in addition to the primary linkage formation by a glucosyltransferase enzyme system, the functions of the cellulose synthase include transportation of UDP glucose from the cytoplasm to the outer surface of the membrane and binding of both donor and acceptor molecules in proper orientation for the action of transglucosylase. In addition, he has suggested the occurrence of transmembrane proteins that associate the synthase complex with cytoskeletal elements which facilitate directed movement of the complex and orientation of the microfibrils and other subsidiary proteins that may serve in the assembly of the complex and its location on the membrane. Though Northcote was discussing cellulose synthesis in higher plants, most of the functions he describes are also expected to be present in the synthase of *Acetobacter xylinum.*

After isolation of the synthase in a highly active form, Benziman and associates pursued an understanding of the key factors in conservation of the activity of the synthase upon isolation.[149-152] This led them to discover the occurrence of a unique activator, cyclic diguanylic acid (c-di-GMP), and enzyme systems involved in regulating its level. They identified one enzyme, diguanilate cyclase, responsible for the formation of c-di-GMP and two phosphodiesterases, which, acting in sequence, degrade it. The balance between the two processes regulates the level of c-di-GMP, which in turn regulates the activity of cellulose synthase. Another important observation flowing from investigation of the isolated synthase was the demonstration that no covalently bonded intermediates were required for the process of assembly of the cellulose chains. It was concluded that "catalysis from UDP-glucose occurs via a direct substitution mechanism, in which the phosphoester-activating group at the anomeric carbon of one glucose residue is displaced by the C-4 hydroxyl group of another glucosyl residue, inverting the α configuration to form a *b*-glucosidic bond." This is in contrast to other proposed mechanisms that had implicated lipid-linked oligosaccharide intermediates in the synthesis of cellulose.[146] The relative simplicity of the mechanism proposed by Benziman and associates makes more plausible the possibility of a multienzyme complex wherein a number of synthase systems act in concert to produce cellulose chains that can then enter into the self-assembly process so characteristic of native celluloses.

As noted, another important advance was identification of the genes responsible for encoding the structures of the synthase components. This provided a basis for exploration of the degree to which similar genes may occur in higher plants. While such genes have not been found, ones with considerable similarity with respect to the domains thought to code for the binding of UDP-glucose and for catalysis of the *b*-(1-4)-linkage formation have been identified for cotton *(Gossypium hirsutum)* and rice *(Oryza sative).* [153] Though this represents an important advance, the findings cannot yet be regarded as conclusive since higher plants are likely to contain many other enzyme systems capable of forming *b*-(1-4)-linkages during the biosynthesis of the majority of the hemicelluloses, and it is not clear that the genes identified are not related to this group of enzymes which are thought to function within the Golgi apparatus rather than at the cell membrane.

Similar genes have been associated with cellulose formation in *Arabidopsis thaliana.* [154] In this instance, a mutant defective in a gene that is similar to those identified for cotton was found to be associated with the production of lower levels of cellulose when *Arabidopsis* is grown at 31 °C when compared to controls grown at 23 °C. Instead, under the conditions of elevated temperature, the mutant produced a glucan, identified as a soluble *b*-(1-4)-glucan, in amounts approximately equivalent to the deficiency in cellulose production. Thus, there is little question that the gene encodes a function that is essential to formation of aggregated microfibrillar cellulose, but since a *b*-(1-4)-

glucan is in fact formed, the encoded function is clearly associated with some more subtle transformation that is at an intermediate stage in the aggregation of cellulose in the unique native form. It is generally recognized that unbranched *b*-1,4-glucans are inherently insoluble at the octamer,[155] so the anomalous glucan is likely to include limited amounts of branching at a level sufficient to make it soluble.[156] More detailed characterization of this anomalous glucan must be accomplished before the full role of the gene can be defined. It is clear that the observations represent important advances towards better understanding of the biogenesis of cellulose at the molecular level, but much remains to be explored before the processes of cellulose formation in higher plants are fully understood.

With increased interest in the tunicate celluloses, it will be of considerable interest to compare the pathways of cellulose biosynthesis when there has been sufficient progress in understanding these systems.

3.16.4.1.2 Ultrastructural organization

Ultrastructural studies of native celluloses have focused most often on organization of the microfibrils, which are the basic units of structure at the level next above that of the unit cells in the lateral direction. At this level also, *Acetobacter xylinum* has been the model system of choice in explorations of the processes of biogenesis. Many other native celluloses from algal and plant sources, and, more recently, from the tunicates, have been investigated, but while they are important from a biological perspective, they do not lend themselves as readily to addressing questions of formation at the level of molecular aggregation. Rather, the patterns of formation are complicated by microscale processes associated with cell wall development; these include the influence of other cell wall constituents on the aggregation of cellulose as well as the constraints of cell wall geometry and the geometry of the assembly complexes, which can impose particular patterns of deformation during the process of aggregation. They also include poorly understood relationships between organization of the microtubules of the cytoskeleton and orientation of the microfibrils in plant cell walls. With *A. xylinum*, in contrast, in the absence of the conditions and the constraints of the cell wall environment, the patterns of aggregation reflect more than anything else the self-assembly characteristic inherent in the nature of cellulose, and to an uncertain degree the influence of the assembly complexes at the cell wall membrane that are thought to play a key role in the aggregation of cellulose molecules in the characteristic cellulose I form.

It is useful here to reiterate the issues that are associated with the distinction between the self-assembly of biological macromolecules within the context of native tissues and their crystallization if they are in an isolated form, because the distinctions are subtle and not immediately obvious and can result in arrival at misleading conclusions in some otherwise important contributions. The key point is that a structure does not have to be crystalline in order to result in diffraction patterns, whether with X-rays or electron beams. The prerequisite rather is a periodicity of structure over domains wherein the coherence of order is sufficient in extension to result in diffraction. With modern instruments it is now possible to acquire such patterns from very small domains, and it is usually implicit in most discussions of such observations of diffraction that the substrate is crystalline. The concept of crystallization in the traditional sense, in contrast, though also implying a process of self-assembly, carries with it the notion of phase separation. As noted in an earlier section, the nanoscale and microscale structures encountered in native celluloses rarely if ever meet the criteria for separate phases established on the basis of statistical physics. It is therefore important to avoid ascribing to the highly ordered native celluloses the characteristics that devolve from the definition of a crystalline phase in the classical sense. For this reason, and to avoid confusion, an effort will be made to describe the ordering of cellulose as self-assembly even in reference to prior work wherein the authors cited spoke of crystallization.[161]

Though the processes of formation of cellulose by *A. xylinum* had been investigated quite extensively in earlier studies, [157-160] Brown and his associates[161,162] were the first to focus attention on the relationship of cellulose fibrils to structures in the walls of the bacterial cells in the course of biogenesis. They first established that linear arrays of terminal complexes on the cell surface were associated with the formation of cellulose ribbons produced by the individual cells. In the course of this work they discovered that certain dyes and fluorescent brightening agents resulted in dramatic changes in the patterns of aggregation.[163-167] They also studied in greater detail an effect that had been reported earlier by Ben-Hayim and Ohad[160] that certain cellulose derivatives of limited substitution also modified the processes of aggregation. The investigations by Brown and associates

were directed at elucidating the relationship between the process of polymerization and self-assembly. These results allow clarification of the levels of structure at which the native organization of the cellulose first manifests itself.

Prior to their exploration of the patterns of aggregation of bacterial celluloses. Brown and co-workers[168] had demonstrated a clear association between the points of deposition of cellulose microfibrils and the cellulose assembly complexes at the plasma membrane for a number of organisms. It was generally observed that *A. xylinum* and a variety of cellulose-forming algae possessed linear arrays of terminal complexes,[169] while higher plant organisms were more frequently found to have the terminal complexes organized into rosettes consisting of six globules each.[170] Though the exact relationship between the terminal complexes and the processes of synthesis have not been established, it is generally accepted that the complexes play a role in the ordering or self-assembly of the cellulose deposited. In most other organisms, observation of the complexes by electron microscopy requires freeze fracture of the cell wall in the plane of the membrane, so that the role of the assembly complexes in biogenesis has to be inferred from adjacency of the ends of the cellulose microfibrils. In the case of *A. xylinum*, the association of cellulose synthesis with the assembly complexes is obvious from electron micrographs that can be prepared from systems wherein the bacteria are cultured directly on microscope grids allowing observation without any disruption of the association between the microfibrils and the terminal complexes embedded in the cell wall. In the same manner, it is also possible to explore the effects of modifying agents without the complications of other effects that may be associated with cell wall geometry or other cell wall components.

When grown under normal conditions, that is, in the absence of any agents that can interact with the cellulose at the molecular level, the cellulose appears to be produced through a linear array of complexes aligned parallel to the cell axis. Under such conditions the cellulose appears in the form of a ribbon that is 40-60 nm wide as shown in Figure 20; the ribbon appears to have a twist with a periodicity of 0.6-0.9 μm. The ribbons appear to consist of smaller fibrillar subunits which are thought to be 6-7 nm in lateral dimension; these are smaller than the microfibrils of many of the algae, but larger than the subunits in higher plants.[171] X-ray diffractometry reveals that these cellulosic aggregates possess a higher degree of order coherence than the pure celluloses from higher plants. As noted earlier, the bacterial celluloses in their native form are shown by solid-state ^{13}C NMR to contain both I_α and I_b forms of cellulose with the I_α form being dominant.

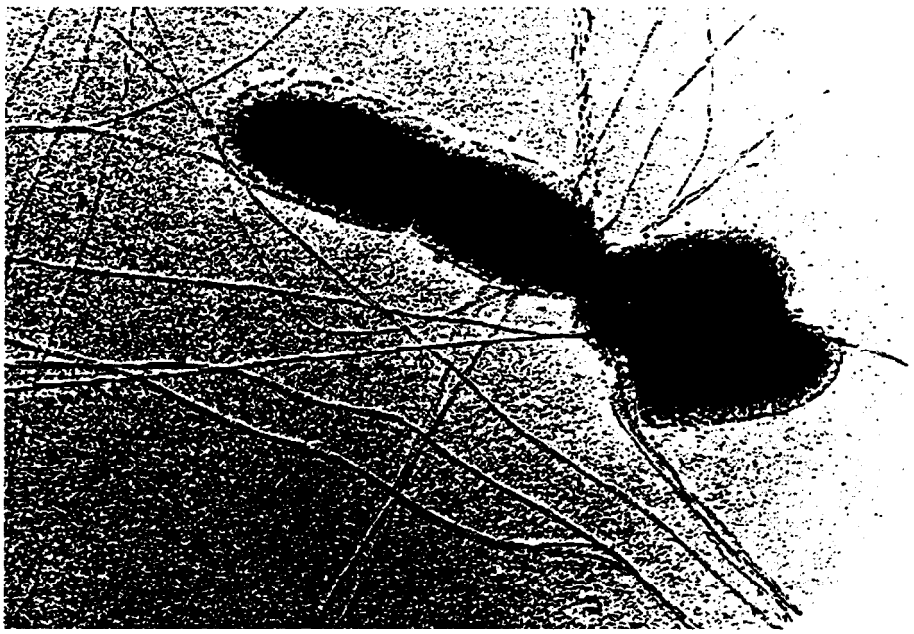


Figure 20 Several normal ribbons of cellulose synthesized by *Acetobacter xylinum*. Twists occur at several points, and in some regions smaller fibrillar subunits are visible (after Haigler[171]).

When the cellulose is formed in the presence of modifying agents, the fibrillar organization is altered to varying degrees. Among the most informative studies are those carried out with fluorescent brightening agents (FBAs) which have conjugated multiple aromatic rings in a planar configuration that facilitates their association with the cellulose molecules in nascent fibrils. The brightening agent studied most extensively is FBA 28, usually identified as Calcofluor White ST or Tinopal LPW. Its

chemical formula is 4,4'-bis[4-anilino-6-bis(2-hydroxyethyl)amino-*S*-triazin-2-ylamino]-2,2'-stilbenedisulfonic acid; CI40622. It has been used in two classes of experiments. When added in small amounts (1-4 μM), the FBA alters the pattern of formation so that the cellulose aggregates as very fine fibrils, approximately 1.5 nm in diameter and they are extruded perpendicular to the cell axis and the plane of the cell wall. It is thought that these very fine fibrils represent aggregation of the cellulose molecules emerging from a single synthase complex system at the cell wall. When the concentration of FBA is raised to 8-250 μM, the fibrillar structure is no longer observed and the cellulose appears to form into many small sheets. Studies of these sheets in the wet state show them to be noncrystalline. One of the most interesting observations from a structural perspective is the transformation that is observed if the noncrystalline sheets are washed in water at a pH of 3.0. The result is that the sheets are transformed into helices of microfibrils about 1.5 nm in diameter as shown in Figure 21.

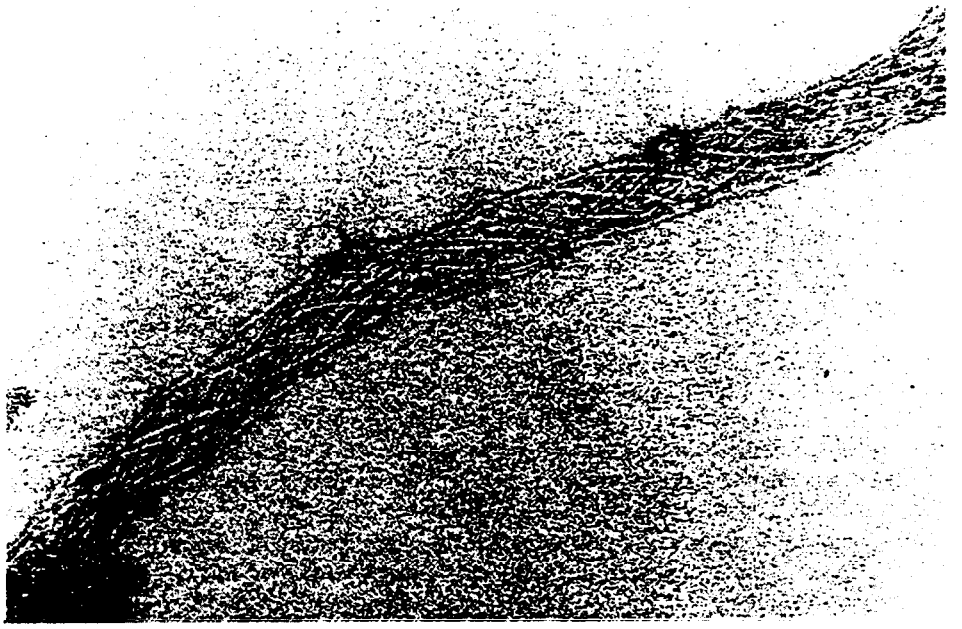


Figure 21 *Acetobacter xylinum* cellulose: a helix of microfibrils formed by gentle washing of the dye from extended sheets formed in its presence. The microfibrils contain cellulose I crystallites with reduced size compared to controls. All of these microfibrils would have been part of one normal ribbon (after Haigler[171]).

The cellulose derivative that was found to have the greatest influence on aggregation of the cellulose from *A. xylinum* is carboxymethyl cellulose (CMC) with a degree of substitution (DS) of 0.7 relative to a maximum possible value of 3.0, and a molecular weight of 90 000 Da. Its effect on the pattern of aggregation was not unlike the lower levels of addition of the FBA in that it resulted in limitations on the aggregation of the ribbon. However, the fibrillar dimensions were somewhat larger than with low levels of FBA and of the order of the lateral dimensions inferred from X-ray diffractometry, in the range of 6-12 nm. The crystallinity of the cellulose formed under these conditions appeared to be similar to that of the controls, suggesting that the action of CMC was at the level of aggregates larger than the 1.5 nm fibrils.

A number of the observations described above have implications with respect to the processes of structure formation at the secondary and tertiary levels. The finding that in the presence of higher levels of FBA the aggregation was not crystalline, but that upon washing, the cellulose I form is recovered, suggests that the cellulose molecules are extruded from the assembly complex in a stable secondary structure typical of the cellulose I form and that it can persist even in the absence of crystallization. Upon washing away the FBA, the cellulose molecules organized in this conformation can then aggregate to form the tertiary structure characteristic of cellulose I. The implication is that the secondary structure is inherently stable or that it is stabilized by association with the FBA molecules that are interleaved between the cellulose chains. At lower concentrations of FBA, where the 1.5 nm fibrils are formed, the implication is that the FBA can be associated with the surfaces of these small fibrils but is not sufficient in amount to penetrate between the individual chains. Both of these observations point to a self-assembly characteristic of cellulose molecules in this conformation. This self-assembly seems to occur at the level of association of the chains into the

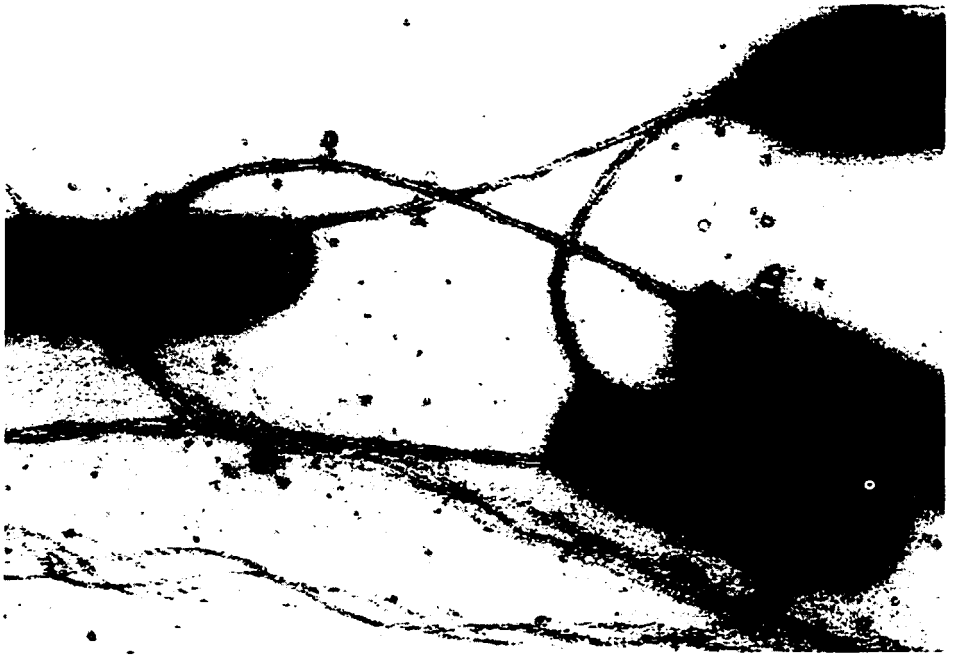


Figure 22 *Acetobacter xylinum* cellulose: the subunits of the normal ribbon remain separate when synthesized in the presence of corboxymethyl cellulose. They may lie closer together, thereby resembling normal ribbons, but close inspection shows separated or sometimes highly splayed subunits (after Haigler[171]).

1.5 nm fibrils when the FBA is not sufficient to associate with all of the chains, but it also seems to function at the next higher level where, in the absence of the FBA, the fibrils associate to form coherently ordered microfibrils at the 7 nm and higher levels, and beyond that in assembly of the fibrils into the ribbons usually produced in control cultures.

Thus, it can be said that the driving force in the process of aggregation is the tendency for self-assembly inherent in the nature of cellulose as it is formed in its native state. A corollary that can be inferred from this observation is that the assembly complexes are the determinants of the secondary structure of the cellulose and that the secondary structure in turn determines the tertiary structure. The process of aggregation of native cellulose can be envisioned as occurring at a number of hierarchic levels, and agents that can modify the processes of aggregation can act at one or more of these levels. In this context, the FBA can be regarded as capable of acting at the most basic level if it is present in sufficient concentration, but at lower concentrations it acts at the next level, that of the 1.5 nm fibrils. CMC can then be viewed as acting at-one level above that of the FBA, resulting in the formation of the 8-l2 nm fibrils.

In a manner not unlike that of the modifying agents, it can be expected that in the cell walls of higher plants, if constituents that associate strongly with cellulose are present, they may well modify the aggregation of cellulose as it is formed and extruded from the assembly complexes at the cell wall membrane. Though such an effect cannot be demonstrated directly in plants, the addition of selected cell wall constituents to cultures of *A. xylinum* does provide a basis for exploring such effects. In a number of studies wherein hemicelluloses have been added to cultures of *A. xylinum*, precisely such an effect is in fact observed. Studies have been carried out for xyloglucans,[172-174] xylans,[172,173] mannans,[172,173] and glucomannans.[175] In all instances the character of the aggregates of bacterial celluloses was shifted towards greater similarity to the aggregates of cellulose found to occur in higher plants. It was also demonstrated that some of these hemicelluloses, particularly the xyloglucan and the glucomannan, co-aggregate with the cellulose so intimately that they cannot be fully extracted even when quite strong caustic solutions are used for the extractions. Finally, it has been observed that the balance between the I_α and the I_b forms, as revealed in the solid-state ^{13}C NMR spectra, is shifted by the hemicelluloses so that the hemicellulose-modified bacteria! celluloses have spectra very similar to the celluloses from higher plants. This has led to the proposal that one of the functions of the hemicelluloses is regulation of the aggregation of celluloses in higher plants.

The view of the processes of association presented here, elaborate as they may be, clearly do not yet represent the full complexity of the system for the assembly of native cellulose. This is perhaps best illustrated by observations by Hirai *et al.* [176] on the effects of temperature. They observed that by allowing *A. xylinum* to grow at 4 °C they were able to induce effects not unlike those that Haigler

et al. [164-167] observed as a result of the addition of FBAs. The effect was equally reversible; that is, when grown at 4 °C the cultures produced noncrystalline bands of cellulose, but when the temperature was increased to 28 °C the production of cellulose reverted to the normal fibrillar form. The full implications of these observations have not yet been developed, but when taken together with observations of the effects of temperature on the biogenesis of cellulose in the *Arabidopsis thaliana* mutant, they clearly suggest that the stability of the synthases or their capacity to generate celluloses in native conformations is more sensitive to temperature than had heretofore been recognized.

The genetic regulation of aggregation at the microscale level is also illustrated in observations by Sugiyama *et a1.* [177,178] Both in studies of algal celluloses[177] and in a systematic examination of approximately 40 tunicates,[178] they observed clear correlations between the I_α/I_b ratios of the celluloses and the species from which they were derived. Among the algal celluloses, there was also an indication that the I_α and I_b forms were spatially organized into nanodomains that varied in the relative distribution of the two forms and that the distributions were species specific. Here again, the full implications of these observations have yet to be developed, but they leave little question that the secondary and tertiary structures of native celluloses are also genetically regulated and that this regulation is at least as sensitive to environmental perturbations as is the formation of the primary structure.

It is worth noting at this point that the formation of the relatively large fibrils of celluloses from the algae and the tunicates are perhaps one of the clearest examples of the tendency of cellulose to self-assemble. When the lateral dimensions are of the order of 20 nm and it is known that the assembly complexes are of smaller dimension, the formation of coherently ordered extended structures points more clearly to self-assembly in a biological sense than to phase separation in the classical sense.

3.16.4.2 Native Celluloses

It was noted earlier that the chemistry of cellulose and its entry into biological processes are as much a function of its state of aggregation as of its primary structure. This reality is manifested as a variability of the response of different celluloses to the action of chemical reagents or cellulolytic enzyme systems. While the origins of this variability have generally been cast in terms of differences in accessibility, this has been primarily because quantitative measures of the organization of the celluloses at the microscale level have been elusive. In search of reproducibility in experimental studies, the vast majority of chemical and biological investigations of cellulose have relied on a handful of readily available substrates that occur in relatively pure form. The substrates most commonly used are cotton, cotton linters, and commercially available microcrystalline celluloses; the cotton linters are to be found in most chemical laboratories in the form of filter paper. Explorations of the effects of aggregation on chemical or biochemical reaction pathways have usually focused on changes that can be induced in the cellulosic substrates through mechanical action to induce varying degrees of decrystallization and disorder or by one or another of the chemical treatments that are known to swell or alter the ordered domains of the native forms. Quantitative measures of the organization of the different aggregated states have usually been cast in terms of accessibility to chemical reagents or in terms of variations in the degree of ordering as quantified by one or another of a number of empirical methods. With the development of some of the new methods for characterizing secondary and tertiary structures described in the preceding section, it is now likely that more studies can be carried out on a wider variety of native forms, and that the relationships between native states of aggregation and the response of the celluloses to the action of chemical and biological agents will be more readily characterized. It is useful, therefore, to develop an overview of what is known of organization in some of the commonly investigated native celluloses.

When considering the structures of native celluloses, it is important to distinguish between two categories of native forms. The first category includes the celluloses that occur in relatively pure form in their native state and which can be isolated using procedures that are relatively mild and that do not perturb or alter the state of aggregation to any significant degree. These are the celluloses that have been the subject of most structural studies. The second category, which includes the vast majority of naturally occurring celluloses, consists of celluloses that are an integral part of the complex architecture of cell walls of higher plants and are intimately blended with the other cell wall constitutents. They are most intimately blended with other cell wall polysaccharides, which are primarily other *b*-(1-4)-linked pentosans or hexosans with varying degrees of limited branching,

collectively identified as the hemicelluloses. The celluloses and other cell wall polysaccharides are together also quite intimately integrated with cell wall lignins. Celluloses in this second category undergo varying degrees of change in their state of aggregation during isolation. However, these changes have not been well recognized or characterized in most instances. One of the major challenges yet to be met in future studies is the definition and characterization of the native state of such celluloses prior to application of isolating procedures that are disruptive to cell wall structure. The usual practice in most prior work on higher plant celluloses has been to assume that the structures of the pure celluloses are adequate models of these more complex forms of cellulose. While for some applications such an approximation may be useful, it has resulted in models of cell wall architecture that ignore the intimacy of the blending of cellulose with the hemicelluloses and lignin. These models will need revision to incorporate findings that the long-accepted two-phase model of plant cell walls is no longer adequate for rationalizing the organization of constituents of plant cell walls. In the following discussion we will first address the issues that complicate characterization of the first category of celluloses. that is. those that naturally occur in a relatively pure state. We will then examine the more complex states of aggregation of celluloses that are so intimately blended with other cell wall constitutents that more severe conditions are necessary for their isolation.

3.16.4.2.1 Pure native celluloses

Much of the discussion in the preceding section has focused on secondary and tertiary structure at the level of the unit cell, with characteristic dimensions of the order of 1 nm, and on the degree of equivalence or nonequivalence of different monomeric or dimeric units within the unit cell. Yet, the unit cell is the foundational concept in the description of crystalline order, which, as noted earlier, is usually predicated on the notion of infinite linear extension of the lattices in all three directions. It is necessary therefore to consider the implications of organization in cellulose at the next level beyond the unit cell; for most native celluloses it is the level at which departures from linear extension of the lattice of the aggregated celluloses is first manifested.

The most common approach, and the one that must be transcended to achieve clearer definition of order in celluloses, is based on categorization of that which is not crystalline as disordered. A number of methods for measurement of the "degree of crystallinity" of celluloses have been established on the basis of this classification. While this approach may allow easy methods of measurements that may be useful for some applications, it is fundamentally misleading because the "nonordered" fraction is usually regarded as disordered and identified as amorphous. The approach is particularly misleading with respect to native celluloses because the methodologies are frequently predicated on the assumption that the nonordered or amorphous fraction is homogeneously disordered, and that is clearly not the case in native celluloses. The reality is that most native celluloses are highly organized, often in a hierarchy of morphological features that are defined at the different scales of observation of the native tissue within which the cellulose occurs.

The difficulty and a measure of the paradoxical situations that arise are best illustrated through comparison of order in algal and cotton cell walls. By all common measures of crystallinity, it is very easy to ascribe to algal celluloses, such as those from *Valonia,* a higher degree of order than one would ascribe to cotton cellulose. Yet at the microscopic level it is clear that the fibrils of cellulose in cotton are more highly organized than those of *Valonia,* and into much more complex morphological features. In cotton, the fibrils are organized into lamella within which the microfibrils are helically wound at 45° to the fiber axis and wherein the direction of the helix is reversed at regular intervals. Furthermore, the lamella occur in concentric diurnal rings with no known correlation between organization of the microfibrils in the different lamella. In *Valonia,* in contrast, the microfibrils are linear and are of much more limited curvature because of the much larger diameters of the *Valonia* cells. The fundamental difference is that to accommodate the much more complex and hierarchic organization of the cellulose in cotton fibers, the microfibrils need to have smaller diameters and are subject to higher degrees of curvature than the microfibrils of *Valonia.* These two factors have a significant influence on the results of measurements used as the basis for most of the common methods for quantification of the crystallinity of cellulose. A number of consequences follow from each of the two factors and it is well to consider their implication for different methods of measurement.

The smaller diameter of cotton microfibrils has been reported as in the range of 3.5–4.5 nm. If, for the purposes of discussion, one were to assume a unit cell of the dimensions generally accepted

for two-chain unit cells, one would find that two-thirds of the unit cells are at the surface of the microfibril. For *Valonia*, in contrast, only one in six would be at the surface. The placement of a unit cell at the surface reduces the packing constraints on the range of values that the internal coordinates defining the secondary structure or conformation can assume. Thus, the freedom of the secondary structure to depart from that in an ideal lattice conformation is far greater for cellulose molecules in a cotton microfibril than for a *Valonia* microfibril. This effect of microfibril diameter is clearly manifested in Figures 9 and 10 where the Raman spectra of ramie and *Valonia* can be contrasted. Ramie has a fibril diameter of about 5-6 nm. which is slightly larger than that of cotton. Furthermore, in the case of ramie the microfibrils are more like *Valonia* with respect to their degree of curvature so that their departure from the ideal lattice is primarily due to the smaller lateral dimensions. Yet it is clear from the broadening of the Raman spectra of ramie fibers relative to those of *Valonia* that the values of the internal coordinates defining the conformation are more broadly distributed than is the case for *Valonia.* This broadening of the distribution of the values of the internal coordinates results in a broadening of the distribution of the frequencies of the normal modes of vibration relative to their values in an ideal lattice.

The effects of microfibril diameter and the presence of significant fractions of cellulose chains near the surface are also manifested in the solid-state ^{13}C NMR spectra. As noted earlier, it was shown by Earl and VanderHart[67] that the areas of the upfield wings of the resonances of C-4 and C-6 increase as the microfibril diameter decreases, indicating that surface chains were subject to different conformational constraints. This variation is seen in Figure 6 where the upfield wing for C-4 is essentially at the level of noise for the spectrum of *Valornia* while its area increases in the celluloses known to have microfibrils of smaller diameters. In the studies by VanderHart and Atalla[73-75] and in those carried out by Newman and Hemmingson,[92] the contribution of the crystalline surfaces to the upfield wings of C-4 was clearly recognized. Newman's interpretation of a doublet resolved in this region as reflecting different crystalline surfaces of the microfibrils suggests that the microfibrils must be viewed as having polyhedral cross-sections; this is in contrast to the premise that the microfibrils may be regarded as cylindrical which is implicit in many discussions of the microfibrils of higher plants. Figure 6 also reveals another manifestation of the effects of microfibril diameter in that the same type of broadening as described above for the Raman spectra in Figures 9 and 10 is obvious in the lineshapes of the NMR resonances. Those that are narrowest and most clearly resolved are those associated with *Valonia.*

In X-ray diffractometry, the method used most often to characterize order in celluloses, the most common measurement method is the acquisition of a powder pattern. A number of different measures of coherence of order are derivable from such patterns. One is based on the width at half-height for the 020 peak at approximately 22.6°, which is inversely related to the lateral dimensions of the crystalline domain by the Debye-Scherer equation. A second is based on the intensity of the 020 peak relative to the intensity at $2q = 18°$, which is taken as representative of the disordered fraction. Yet another, used less often on its own, but sometimes measured to test the consistency of the other two measurements, is the position of the peak of the 020 reflection. It is usually highest for the most highly ordered celluloses, which are expected to have the fewest defects and the highest densities in the crystalline domains.

For all of the above approaches to characterizing the states of aggregation of native cellulose, the curvature of the microfibrils has an effect not unlike that of the smaller diameters. To the extent that the curvature is a measure of departure from the infinitely extended linear lattice, it induces in the internal coordinates departures from their values in an ideal lattice. It also results in a reduction of the degree of coherence in the relationships between unit cells as they are further apart from each other within an individual microfibril. Though curvature clearly contributes to broadening of the lines in the Raman spectra and the solid-state ^{13}C NMR resonances, its most serious misinterpretation occurs in the context of efforts to quantify the crystallinity of cellulose on the basis of X-ray powder patterns, where its contribution to the broader background is taken as an indication of homogeneous disorder.

The effects of microfibril diameter and curvature are noted here, not to discredit the methods that are commonly used to quantify order in different celluloses, but rather to suggest that the interpretation of results derived from such measurements must be approached with caution. It is clear that the types of order that occur in native celluloses at higher levels of organization are not readily described in mathematical terms and therefore it is not easy to relate the experimental measurements to the degrees of departure from the linearly extended infinite lattice. Yet it is also clear that the categorization of that which is not linearly ordered as disordered or amorphous can be quite misleading, particularly in the context of native celluloses. One of the major challenges for the years ahead is the development of a conceptual framework that is adequate for description of the more complex forms of order that occur in native celluloses.

3.16.4.2.2 Complex native celluloses

The term complex native celluloses is used here to indicate the class of native celluloses that occur naturally in intimate combination with other plant cell wall constituents, most often in higher plants. It has been generally assumed that the cell wall structure is adequately represented as a two-phase system, one consisting of the microfibrils of pure cellulose, the other a blend of all of the other constituents. Furthermore, it was implicit in this view that the cellulose phase is not unlike the pure native celluloses and that it is possible to isolate the cellulose from such cell walls by removing the other components, leaving behind the cellulose in a condition that approximates its native state. While these assumptions remain the basis of most discussions of plant cell wall structure, there is now clear evidence that the assumptions must be reassessed. The reassessments need to be at two levels. The first, which is perhaps more readily resolved, addresses the question whether it is possible to remove the other cell wall constituents without disrupting the native organization of the cellulose in its native form. The second raises the more fundamental question as to whether the two-phase model is at al valid in the context of the native state.

The generally accepted model of the cell wall in plants is well represented by Preston's chapter on "General principles of wall architecture",[8] where it is stated that:

> All plant cell walls are multiphase systems which for simplicity we may regard as two-phase. All walls known contain a crystalline polysaccharide of specific composition embedded in a matrix consisting usually of the wide variety of polysaccharides and other compounds already dealt with. As already mentioned, the crystalline polysaccharide of most plants is cellulose.

For primary cell wall structure, a mode! by Carpita and Gibaut[179] has received wide acceptance, but it also implicitly incorporates the two-phase view of the cell wall matrix. In both contexts, the two-phase view is based on the knowledge that cellulose occurs in the cell wall and that some form of insoluble polysaccharide can be isolated, that can be shown to include significant amounts of *b*-(1-4-linked anhydroglucose units, and that in some instances it produces X-ray diffraction patterns typical of cellulose. It is therefore important to consider whether the steps intermediate between the native state and the isolated cellulose alter the state of aggregation, and to examine whether detection of the diffraction pattern of cellulose is an adequate criterion for the definition of cellulose.

The first explorations of this issue were undertaken in the course of assessment of industrial processes for the isolation of cellulose from woody tissue through commercial pulping operations. It had earlier been demonstrated that with mercerized celluloses, the degree of ordering could be enhanced by annealing at temperatures below those used in commercial pulping.[180] The question arose as to whether exposure to the elevated temperatures used in commercial pulping operations (170-190°C) can have a similar influence on the celluloses isolated from woody tissue wherein: the cellulose occurs in intimate association with the cell wall hemicelluloses, particularly in the secondary wall. Samples of wood were subjected to delignification using the acid-chlorite process at 70°C and this was followed by treatment with 4% NaOH to remove the extractable hemicelluloses. A control sample was then subjected to an X-ray diffractometric measurement, while another was sealed in a buffer solution and exposed to a temperature cycle simulating the temperature cycle applied to the wood chips and the pulping liquors during commercial pulping. The X-ray diffractometric pattern reflected a significant increase in the degree of ordering of the cellulose as a result of exposure to the elevated temperature,[181] suggesting that removal of the hemicelluloses had facilitated a transformation in the state of aggregation of the cellulose. While this observation has many implications, the key one in the present context is that the state of aggregation of most commercially available celluloses derived from complex celluloses does not reflect the native state.

Though the observation concerning commercially available celluloses may seem relevant only to the uses of cellulose in commercial applications, it is of considerable importance in cellulose research because it has become quite common for research workers to rely on microcrystalline celluloses that are commercially available from suppliers of chemical reagents. The microcrystalline celluloses are, in the vast majority of instances, prepared from commercial dissolving pulps through acid hydrolysis followed by mechanical disruption and dispersion. The dissolving pulps, in turn, are derived from the complex celluloses that occur in woody species so that in their native state they are intimately blended with hemicelluloses. The majority of hemicelluloses are degraded and extracted in the course of the pulping cycles, but it is rare to find a microcrystalline cellulose that is entirely free of hemicellulose residues.

There are of course more fundamental reasons for exploring the nature of complex celluloses. In a number of studies, the association between hemicelluloses and cellulose was explored in contexts

that simulated the processes of biogenesis. Here again, the model system used was *A. xylinum.* The work represented an extension of the studies by Haigler *et al.* [164-167] concerning the effects of cellulose derivatives on the aggregation of celluloses produced by *A. xylinum*. In these studies the polymers added were not cellulose derivatives, but rather hemicelluloses known to occur in higher plants.[172-174] The clear indication in all of these studies was that the hemicelluloses associated with the cellulose during its biogenesis. It was clear that the association is at the most elementary level in that it resulted in co-aggregation of the hemicelluloses with cellulose resulting in what has been defined as complex celluloses. Furthermore, it appeared that each of the hemicelluloses modified the process of aggregation in a different manner so that the resulting complex celluloses were different from each other when characterized by methods that are sensitive to variations in secondary and tertiary structure.

With respect to understanding the nature of complex celluloses, two features which emerged in the studies of these model systems are important and suggest the possibility of errors in interpretations of measurements in studies of plant celluloses. The first is related to the degree to which the crystallization habit of cellulose is dominant even when significant amounts of hemicelluloses are co-aggregated with it. For example, it was observed that X-ray powder patterns recorded for the complex celluloses appeared very similar to those usually recorded for less ordered celluloses.[172-174] This phenomenon is not unusual in the context of the X-ray diffractometry of polymers where it is often observed that in polymer blends or in block copolymer systems, the diffraction pattern is often that of the dominant component though it may appear somewhat modified.[11] This effect is also noted in the solid-state ^{13}C NMR spectra of the complex celluloses; the major features of the spectra are very similar to those of the pure celluloses when the latter are less ordered.

The second feature explored in the studies of complex celluloses generated with the *A. xylinum* system is that of the degree to which the hemicelluloses are integrated into the structure of cellulose. One of the observations in these studies was that although the hemicelluloses were added at the level of 0.5% in the culture medium, uptake into the structure of the cellulose was at the level of 30% for the xyloglucan and 24% in the case of the glucomannan. It was also observed that extraction with 17.5% NaOH resulted in removal of less than half of the amount of hemicellulose incorporated into the complex cellulose. In order to carry out comparisons with the traditional α–cellulose content test usually applied to commercial celluloses derived from wood pulps, the complex celluloses prepared from *A. xylinum* would have had to be subjected to an alkaline pulping process first; this was not done. The difficulty of extracting the hemicelluloses is clearly indicative of incorporation of the hemicelluloses at an early stage of aggregation. It would appear that the hemicelluloses, both with *b*-(l-4)-linked backbones, are able to enter into the self-assembly process of the cellulose, quite probably at the earliest stage of aggregation intermediate between the 1.5 nm fibrils reported by Haigler [171] and the level above. The intimacy of incorporation of the hemicelluloses and the difficulty of extracting them may well have been the basis for the opinion expressed by many senior researchers in the field of cellulose science late into the first half of this century,[182] which held that some of the hemicellulose derived sugars are covalently bonded and an integral part of the cellulose backbone.

The effects of polymeric additives on cultures of *Acetobacter xylinum* were also studied by Yamamoto and Horii[86] who also detected shifts in the balance between the I_α and I_b forms that were consistent with earlier reports based on other observational methods.[171-174] as well as on the basis of ^{13}C NMR.[175]The earlier studies had shown that polymers with a *b*-1,4 glucan backbone tend to shift the structures to ones more similar to higher plant celluloses. Yamamoto and Hori found that addition of xyloglucan and the sodium salt of carboxymethylcelulose at 2% and 2.5%, respectively, reduced the I_α content of the bacterial cellulose to a significant degree, while the addition of polyvinyl alcohol or polyethylene glycol had no effect. As noted in an earlier section (3.16.3.2.2) they also observed an effect of temperature that was quite interesting in that lower temperatures favored a higher level of the I_α form.

In another investigation of celluloses in plants, Hackney, Atalla, and VanderHart examined a wide range of cell wall constituents.[183,184] The general pattern which emerged was that in many primitive algae, where the other cell wall polysaccharides tended to be primarily acidic or pectic, there was very little evidence of co-aggregation of the polysaccharides with cellulose, and the celluloses isolated were sufficiently free of other matter that they could be classified among the pure celluloses. At higher levels of evolution, there emerged a pattern of occurrence of neutral polysaccharides with *b*-(l-4)-linked backbones, together with the cellulose. In some instances the other polysaccharides were the dominant components. However, in the vast majority of instances, the cellulose was the dominant component. It appeared that at a threshold level of 65% cellulose content, the aggregation habit of cellulose became the one that was detected by both X-ray powder

patterns and solid-state ^{13}C NMR. In all instances, it was clear that while the X-ray patterns and the NMR spectra were the distinctive ones characteristic of native cellulose, they reflected a much lower level of ordering than would be observed from similar observations of pure celluloses.

The observations on complex celluloses are particularly important in relation to the common practice of using diffraction patterns to distinguish celluloses isolated from plant cell walls using mild extraction procedures. It needs to be noted in future studies that although patterns characteristic of cellulose are observed, the isolated insoluble substance may contain significant amounts of hemicelluloses as minor components and yet retain the characteristic aggregative properties of cellulose as characterized by X-ray diffraction, solid-state ^{13}C NMR spectroscopy, or on the basis of insolubility in acid media, the latter of course in the absence of strong hydrolytic conditions.

In relation to the discussion of complex celluloses, it is also helpful to revisit some of the relatively recent solid-state ^{13}C NMR studies of these celluloses. Newman *et al.* [185] applied the procedures they developed to investigations of a number of native celluloses of particular interest from a biological perspective, all of which are in the class of complex celluloses as they have been defined here. These include primary wall celluloses from apple cell walls[185] and from the leaves of *Arabidopsis thaliana*[186] and, in 1997, tissue from the silver tree fern *Cyathea dealbata.* [187] In the studies of primary wall celluloses from both apple cells and *A. thaliana* leaves, it was possible to resolve two resonances associated with the surfaces of fibrils and to suggest that the I_α and I_b forms occur in similar proportions. The intensities of the resonances associated with the surfaces of fibrils, when compared with the intensities of resonances associated with crystalline interiors, suggested lateral dimensions of the fibrils of 2.5-3.0 nm. These were then interpreted in terms of fibrils consisting of 23 chains of cellulose molecules, 14 of which were at the surface and nine in the interior; of the interior molecules, eight were only one layer removed from the surface. While the observations and their interpretation are not questioned here within the context of the accepted models of order in native celluloses, they clearly point to the need to ask again whether the traditional concepts associated with crystalline order are meaningful when so much of the substance under examination is at the surface. It would appear that some of the questions arising in the earlier periods of research on cellulose have yet to be answered.

Newman's study of the cellulose from fibrous material from the silver tree fern *C. dealbata* did not address the dimensions of the fibrils, but suggested the occurrence of similar proportions of the I_α and I_b forms. This, taken together with the earlier observations on primary wall celluloses, may raise questions concerning the earlier generalizations that higher plant celluloses have the I_b form as the dominant component. The question, however, must remain open until the possibility of quantitation on the basis of resolution-enhanced spectra is more firmly established. The question concerning quantitation of the relative amounts of the I_α and I_b forms on the basis of resolution-enhanced spectra also arises with respect to the comparison of celluloses from softwoods and hardwoods carried out by Newrnan.[188]

It is also useful to review the results of Larsson *et al.* concerning the solid-state ^{13}C NMR spectra of the celluloses they investigated within the present perspective.[100] It emerges that the wood celluloses, which are the only complex celluloses examined by them, are those which appear to have as their largest fractions the category characterized as paracrystalline and the category of hemicelluloses and cellulose oligomers. The I_α and I_b forms together represent only 9.1% of the total. Cotton cellulose, which is a pure cellulose but possesses a very complex morphology, has approximately 32% of the I_α and I_b forms, with the remaining large fractions belonging almost equally to the categories of paracrystalline and amorphous and surface fractions. The *Holocynthis, Cladophera,* and *Valonisa* celluloses are clearly among the pure celluloses, but the fact that the microfibrils need not enter into a more complex morphological architecture allows the largest fraction of the cellulose to retain the coherence of order of the molecular chains within the fibrils.

With the development of new methods to probe the states of aggregation, it is quite likely that new approaches will be developed to establish the relationships between microscale order and the properties of the celluloses as substrates. The key will be movement from characterization of the level of order in terms of a partitioning between fully ordered and disordered towards a more detailed account of the distribution among the different categories of order that prevail within the context of a particular state of aggregation.

New approaches to characterization of celluloses may well be based on the variability in secondary and tertiary structures. Variations in the secondary structure are most likely to manifest themselves as steric effects, which are issues associated with reaction pathways at the molecular level. Differences in tertiary structure, in turn, are expected to have their primary influence on accessibility to reagents, which is an issue arising from heterogeneity at the microscale level. Because of the difficulty in

separating these effects, efforts to examine their separate influences have been quite limited. New methods for categorizing order in celluloses are likely to be developed on the basis of spectroscopic measurements.

3.16.4.3 Biodegradation

Next to biogenesis, the biodegradation of cellulose is the most significant life process connecting cellulose with the biosphere, for it is an essential component in the pathway to the re-entry of dead plant and algal biomass into the processes of life. The biodegradation is, in most instances, the result of action of microorganisms adapted to the degradation of algal and plant biomass. Through the action of hydrolytic enzymes, the cellulose and other cell wall polysaccharides are converted to glucose and other pyranoses or oligosaccharides which then become available as nutrients for the host organisms producing the enzymes. The biodegradation occurs in two broad categories of environments. The first is in the digestive systems of a wide range of ruminant and herbivorous organisms where the by-products of hydrolysis become available to both the microorganisms producing the cellulolytic enzymes and the host organisms. The second and broader category is the wide range of environments in which biomass is degraded in open fields and aquatic environments as part of the cyclic death and renewal of annual or perennial plants, and of shrubs, dead trees, and foliage on the forest floor. This second broader category also includes many of the plant pathogenic microorganisms that secrete hydrolytic enzymes as an integral part of their process of penetrating and metabolizing the constituents of the plants that they colonize.

The processes of biodegradation of cellulose and related cell wall polysaccharides have been investigated quite extensively both from a fundamental perspective in search of deeper understanding of the mechanisms of their action and within the context of efforts toward practical application of this understanding to protect celluloses from decay or, when specific enzymes of known and controllable action can be isolated, to use them to modify cellulosic fibers in a number of commercial applications. For example, cellulases have been used to modify the properties of cellulosic fibers in textile applications, and they have been used to facilitate removal of inks from recycled cellulosic pulp fibers in the manufacture of paper. The xylanases from some of the same organisms have been used to enhance the bleachability of virgin kraft pulp fibers in the manufacture of paper.

Here the author discusses briefly some of the fundamental aspects of cellulolytic action by organisms that have developed the capacity to disassemble cellulose through evolutionary adaptation to environments wherein lignocellulosic biomass represents the primary source of nutrients for the organisms. The emphasis will be on the functional aspects of cellulolytic systems. Much more comprehensive discussions of the structures and protein chemistry of the enzyme systems are available to the interested reader.[189-193] The enzyme systems that complement the cellulolytic systems by disassembling other constituents of plant biomass are beyond the scope of this chapter.

While the occurrence of cellulolytic systems was recognized late in the nineteenth century, mechanistic explorations of the processes began in the middle of the twentieth century. The pioneering work of Rees and Mandels[194] established that in the instance of native celluloses used in textile applications such as cotton or linen, which have highly ordered crystalline domains as a major fraction, a complex enzyme system was necessary for breaking down the cellulose. They were able to separate two active fractions that appeared to act in synergy with each other to disassemble the celluloses. One fraction, identified as C_x was found to be capable of hydrolyzing celluloses that had been modified by the application of swelling treatments such as, for example, Walseth cellulose,[195] which is prepared by swelling with phosphoric acid. But acting alone, it was not effective for hydrolyzing cotton. When the other fraction identified as C_l was added, it appeared to act in synergy with the C_x fraction to bring about hydrolysis. It was proposed that the C_l fraction activates the crystalline cellulose so that it becomes susceptible to attack by the C_x fraction.

While the C_l-C_x hypothesis represented an important advance in recognizing that the action of cellulolytic organisms proceeded through the action of multienzyme systems, it was soon recognized that most cellulolytic organisms produce a wider array of enzymes that act in synergy to disassemble the cellulose. Eriksson and co-workers[196,197] suggested that the C_1 enzyme is in reality an “endwise acting” exoglucanase enzyme that can detach cellobiose units from the nonreducing ends of the cellulose chain. The C_x enzyme was viewed as an endoglucanase that can attack and hydrolyze the *b*-(1-4)-linkages at any point in an exposed chain. Thus, the new nonreducing chain ends would become available for the action of the exoglucanase and therein lay the basis for the synergy. In addition it was recognized that a (1-4)-*b*-glucosidase was also present in the cellulolytic system to

convert the cellobiose and other soluble cellodextrins to glucose. Soon thereafter exoglucanases were isolated from a number of cellulose-degrading fungi.

With the work of this period it became clear that the enzyme systems of most cellulolytic organisms included a wider variety of enzymes with each optimized for a different aspect of the disassembly of cellulose molecules and their ordered aggregates. In addition to the participation of a number of different hydrolases, each with a specific function in relation to the overall process, some oxidative enzymes were also identified although their function is not as well understood. The review by Eriksson and Wood provides a good overview of cellulolytic enzyme systems as they were understood in terms of function, as well as the synergies between them.[198] Since that time advances in protein chemistry have added a different dimension to studies of cellulolytic systems, for in addition to defining them by function it has become possible to classify them by structure. The additional structural information has made possible explorations of relationships between structure and function. One of the most comprehensive reviews is that by Tomme *et a1.;* [191] the perspective presented therein forms the basis of the following brief overview and additional discussions.

Beyond the realization that most cellulolytic microorganisms rely on a multiplicity of enzymes that act in synergy, one of the central results of structural studies is the recognition that the enzymes are modular in organization. While there are some exceptions, the majority consist of three key domains. The first is a cellulose binding domain (CBD), the primary function of which is to adsorb the enzyme on the substrate. In addition, in some instances the CBD appears to condition the surface to enhance its susceptibility to attack by the catalytic domain which is responsible for the hydrolysis. These two modules are connected by a third domain referred to as the linker region. The structural information has also allowed the classification of cellulases from different organisms into a number of structurally related families on the basis of amino acid sequence similarities. The tabulation by Tomme *et al.* [191] includes 17 such families.

Another important finding is related to the manner in which the cellulases of a particular cellulolytic system function to complement each other. Among those described by Tomme *et al.* as "Noncomplexed systems," the organism secretes into the environment a mix of endoglucanases and exoglucases or cellobiohydrolases which act in synergy upon the substrate. Such systems are generally characteristic of aerobic fungi and bacteria. The most extensively studied system is *Trichoderma reesei,* which secretes two major cellobiohydrolases, CBHI and CBHII, two major endoglucanases, EGI and EGII, and at least two low molecular weight endoglucanases EGIII and EGV. In addition, some *b*-glucosidases are secreted to relieve product inhibition by hydrolyzing cellobiose and other soluble celloligosaccharides. Another extensively studied system is that used by the fungus *Phanaerochete chrysosporium* which is of particular interest because it possesses multiple CBHI genes.[199] The expression of the different genes in response to differences in the cellulosic substrates [200,201] reflects a higher order in the mechanism of induction than had heretofore been recognized.

The second major group of cellulolytic systems are the complexed systems wherein the cellulases are organized in relation to each other so that their action can be collective. They are typically produced by anaerobic organisms including bacteria and fungi that colonize anaerobic environments such as the rumen and hind-gut of herbivores, composting biomass, and sewage. Among the most interesting of the anaerobic bacteria are those such as the *Clostridium* spp. where the complexed enzymes occur in distinct high molecular weight protein complexes called cellulosomes. The cellulolytic systems of *Clostridium thermocellum* have been studied extensively by Lamed, Bayer and co-workers.[202-205] In these systems the cellullases appear spatially organized through their association with a special protein referred to as a scafoldin. Its function appears to be to localize the different cellulases so that they act in concert. Furthermore, it has been demonstrated that this protein includes a CBD so that it can position the associated cellulases in relation to the cellulosic substrate. The cellulosomes appear to be exocellular organelles protruding from the surfaces of the cells, but they also seem to detach and attach themselves to the cellulosic substrates. It is of particular interest that the specific activity on crystalline celluloses appears to be 50-fold higher than the extracellular system produced by *Trichoderma reesei.* There has been significant progress in understanding the structure and functional organization of cellulosomes and these have been discussed in several recent reviews.[205-207]

Two other arenas of research on the natural disassembly of celluloses are worthy of note here. Recent studies of the action of selected cellulases on relatively pure celluloses have demonstrated a greater susceptibility of the I_α form to attack although the I_b form was also susceptible to attack but in different patterns.[208,209] There is little question that further pursuit of these patterns will add to understanding of both the native structures and the nature of cellulolytic action.

The second area that is of equal importance to the processes of biodegradation discussed above

but is beyond the scope of this chapter is the role of endoglucanases in life processes of plants. Cellulase activity has been detected in both growing and senescing plants.[210] To the extent that native celluloses are the primary structural components of plant tissue. it is evident that changes in form cannot take place without both assembly and disassembly on a continuing basis.

3.16.5 FUTURE DIRECTIONS

In Section 3.16.2, it was noted that at the beginning of the twentieth century studies of cellulose played an important role in formulating the polymer hypothesis. In the main, cellulose has been investigated as one of the natural homopolymers within the framework of polymer science; much of the work has been based on the premise that a single ideal structure was the key to understanding cellulose. The discovery that all native celluloses are composites of two native forms and that the mixture of these two forms is species specific will open the possibility for future focus on the individuality of native celluloses. The findings that microorganisms can detect subtle differences between celluloses resulting in the expression of different isozymes for their disassembly[200,201] suggests that the interfaces between cellulose science and biology and microbiology will be fruitful ones for research in the twenty-first century.

It is anticipated that the physics and chemistry of nanodomains will advance considerably during the early part of the twenty-first century, in part to accommodate the drive towards miniaturization and in part due to the development of instrumental methods capable of greater resolution of structure than is possible at the present time. It is likely that these advances in scientific methods in general will have parallels in the arena of cellulose science. It is likely that as a consequence there will be greater recognition of the individuality of native celluloses as unique composites at the nanoscale level, and that this in large measure is a reflection of the capacity of the cellulose molecule for self-assembly in different patterns depending on the microenvironment within which the assembly takes place.

Finally, from a systems perspective, it will be in order to compare the levels of complexity that are recognized in the discussions of the different aspects of the biological processes in which cellulose plays a central role. The question arises, for example, whether the complexities of the systems that have evolved for the disassembly of cellulose, as in the case of cellulosomes, need be more complex than the systems that assemble cellulose? Or conversely, is the fact that systems as complex as the cellulosomes have evolved for disassembly a measure of the complexity of assembly? To the extent that the processes of disassembly frequently reflect diversity in the secondary and tertiary structures of the native celluloses, it would appear that a complete picture of the biogenesis will require a better understanding of the processes of formation of secondary and tertiary structure, and that investigations focused on the formation of primary structure are not adequate for the task of unraveling the mysteries of formation of native celluloses as they occur in higher plants. It is inviting to speculate that the assembly complexes usually visualized through electron microscopy are organelles at least as complex as the cellulosomes. Though Northcote did not explicitly suggest this,[148] his description of the many functions that must be coordinated to produce the primary structure suggests the orchestrated action of multiple enzyme systems, and his focus was on primary structure formation. If one then considers the need for regulation of secondary and tertiary structure formation, it seems plausible that the assembly complexes are more than synthases that are simply of a higher order of complexity than those observed in *Acetobacter xylinum.* Rather, it would appear that they are more highly regulated, complex assemblies of subsystems analogous to the bacterial synthase. The fact that they appear to be stable only in the context of living cells poses serious challenges in the search for deeper understanding of the biogenesis of celluloses of higher organisms. It may well be that investigations of model systems in tissue culture will prove to be the key to understanding these systems.

ACKNOWLEDGMENTS

The author's work in this area has been supported over the years by the USDA Forest Service, the Division of Energy Biosciences of the Office of Basic Energy Sciences of the US Department of Energy and by the Institute of Paper Chemistry, Appleton, Wisconsin (now the Institute of Paper Science and Technology, Atlanta, Georgia); all are gratefully acknowledged. The author's many collaborators in the area of Raman spectroscopy, whose work has been cited, have contributed

immeasurably during many discussions of the different aspects of the work. The long standing collaboration with Dr. David VanderHart, of the National Institute of Standards and Technology, has been particularly important to much of the progress in the arena of solid state ^{13}C NM R studies of cellulose. Dr. VanderHart's expertise in this arena and his wisdom in addressing complex issues have been of great aid to the author. Discussions with colleagues over the years have also contributed to the author's perspective. Of particular value have been discussions with Dr. A. Isogai, Dr. C. Haigler, Dr. H. Chanzy, Dr. J. Sugiyama, Dr. R. A. J. Warren, Dr. T. Teeri, and Dr. L. Viikari; all are gratefully acknowledged.

3.16.6 REFERENCES

1. C. B. Purves, in "Cellulose and Cellulose Derivatives, Pt. I," eds. E. Ott, H. M. Spurlin, and M. W. Graffline, Wiley, New York, 1954, p. 29.
2. P. H. Hermans, "Physics and Chemistry of Cellulose Fibers," Elsevier, New York, 1949.
3. P. J. Flory, "Principles of Polymer Chemistry," Cornell University Press, Ithaca, NY, 1953.
4. C. F. Cross and E. J. Bevan, Researches on Cellulose (III), Longmans, Green & Co., London, 1912.
5. E. Heuser, The Chemistry of Cellulose, Wiley, New York, 1944.
6. D. W. Jones, in "Cellulose and Cellulose Derivatives, Pt. IV," eds. N. M. Bikales and L. Segal, Wiley, New York, 1971, p. 117.
7. B. A. Tonessen and O. Ellefsen, in "Cellulose and Cellulose Derivatives, Pt. IV," eds. N. M. Bikales and L. Segal, Wiley, New York, 1971, p. 265.
8. R. D. Preston, "The Physical Biology of Plant Cell Walls," Chapman and Hall, London, 1974.
9. A. Frey-Wyssling, "The Plant Cell Wall," Gebruder Borntrager, Berlin, 1976.
10. R. H. Atalla, in "The Structures of Celluloses," ed. R. H. Atalla, ACS Symposium Series No. 340, American Chemical Society, Washington, DC, 1987 p. 1.
11. M. Kakudo and N. Kasai, "X-Ray Diffraction by Polymers," Elsevier, New York, 1972, p. 285.
12. S. Arnott, in "Fiber Diffraction Methods," ACS Symposium Series No. 141, American Chemical Society, Washington, DC, 1980, p. 1.
13. E. D. T. Atkins, in "Fiber Diffraction Methods," ACS Symposium Series No. 141, American Chemical Society, Washington, DC, 1980, p. 31.
14. H. Tadokoro, in "Fiber Diffraction Methods," ACS Symposium Series No. 141, American Chemical Society, Washington, DC, 1980, p. 43.
15. H. Tadokoro, "Structure of Crystalline Polymers," Wiley, New York, 1979, p. 6.
16. J. A. Howsmon and W. A. Sisson, in "Cellulose and Cellulose Derivatives, Pt. I," eds E. Ott, H. M. Spurlin, and M. W. Graffline, Wiley, New York, 1954, p. 231.
17. T. Petipas, M. Oberlin, and J. Mering, *J. Polym. Sci. C*, 1963, **2**, 423.
18. M. Norman, *Text. Res. J.*, 1963, **33**, 711.
19. K. H. Gardner, and J. Blackwell, *Biopolymers*, 1974, **13**, 1975.
20. J. J. Hebert and L. L. Muller, *J. Appl. Polym. Sci.*, 1974, **18**, 3373.
21. D. W. Jones, in "Cellulose and Cellulose Derivatives, Pt. IV," eds. N. M. Bikales and L. Segal, Wiley, New York, 1971, p. 138.
22. J. A. Howsmon and W. A. Sisson, in "Cellulose and Cellulose Derivatives, Pt. I," eds. E. Ott, H. M. Spurlin, and M. W. Graffline, Wiley, New York, 1954, p. 237.
23. K. H. Meyer and H. Z. Mark, *Physik. Chem.*, 1929, **B2**, 115.
24. K. H. Meyer and L. Misch, *Helv. Chim. Acta*, 1937, **20**, 232.
25. F. T. Pierce, *Trans. Faraday Soc.*, 1946, **42**, 545.
26. A. Sarko and R. Muggli, *Macromolecules*, 1974, **7**, 486.
27. A. D. French, W. A. Roughead, and D. P. Miller, in "The Structures of Celluloses," ed. R. H. Atalla, ACS Symposium Series No. 340, American Chemical Society, Washington, DC, 1987, p. 15.
28. A. D. French, *Carbohydrate Res.*, 1978, **61**, 67.
29. R. H. Marchessault and P. R. Sundararajan, in "The Polysaccharides," ed. G. O. Aspinall, Academic Press, New York, 1983, vol. 2, p. 11.
30. O. Ellefsen and B. A. Tonessen, in "Cellulose and Cellulose Derivatives, Pt. IV," eds. N. M. Bikales and L. Segal, Wiley, New York, 1971, p. 151.
31. J. Blackwell and R. H. Marchessault, in "Cellulose and Cellulose Derivatives, Pt. IV," eds. N. M. Bikales and L. Segal, Wiley, New York, 1971, p. 1.
32. H. J. Marinan and J. Mann, *J. Polym. Sci.*, 1956, **21**, 301.
33. J. Mann and H. J. Marinan, *J. Polym. Sci.*, 1958, **32**, 357.
34. M. L. Nelson and R. T. O'Connor, *J. Appl. Polym. Sci.*, 1964, **8**, 1311.
35. M. L. Nelson and R. T. O'Connor, *J. Appl. Poly. Sci.*, 1964, **8**, 1325.
36. L. Pauling, "The Nature of the Chemical Bond," 3rd edn., Cornell University Press, Ithaca, NY, 1960, p. 65.
37. L. A. Woodward, "Introduction to the Theory of Molecular Vibrations and Vibrational Spectroscopy," University Press, Oxford, 1972, p. 344.
38. R. H. Atalla, *Appl. Polym. Symp.*, 1976, **28**, 659.
39. J. Blackwell, P. D. Vasko, and J. L. Konig, *J. Appl. Phys.*, 1970, **41**, 4375.
40. L. J. Pitzner, Ph.D. Dissertation, The Institute of Paper Chemistry, Appleton, WI, 1973.
41. L. J. Pitzner and R. H. Atalla, *Spectrochim. Acta Part A*, 1975, **31A**, 911.
42. G. M. Watson, Ph.D. Dissertation, The Institute of Paper Chemistry, Appleton, WI, 1974.
43. S. L. Edwards, Ph.D. Dissertation, The Institute of Paper Chemistry, Appleton, WI, 1976.

44. R. M. Williams, Ph.D. Dissertation, The Institute of Paper Chemistry, Appleton, WI, 1977.
45. R. M. Williams and R. H. Atalla, *J. Phys. Chem.*, 1984, **88**, 508.
46. H. A. Wells, Ph.D. Dissertation, The Institute of Paper Chemistry, Appleton, WI, 1977.
47. H. A. Wells and R. H. Atalla, *J. Mol Struct.*, 1990, **224**, 385.
48. K. P. Carlson, Ph.D. Dissertation, The Institute of Paper Chemistry, Appleton, WI, 1978.
49. H. Wiley and R. H. Atalla, *Carbohyd. Res.*, 1987, **160**, 113.
50. D. A. Rees and R. J. Skerret, *Carbohyd. Res.*, 1968, **7**, 334.
51. A. Sarko and R. Muggli, *Macromolecules* 1974, **7**, 486.
52. S. S. C. Chu and G. A. Jeffrey, *Acta Cryst.*, 1968, **B24**, 830.
53. J. T. Ham and D. G. Williams, *Acta Cryst.*, 1970, **B29**, 1373.
54. E. B. Wilson, Jr., J. C. Decius, and P. C. Cross, "Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra," McGraw-Hill, New York, 1955, p. 188.
55. J. J. Cael, J. L. Koenig, and J. Blackwell, *Carbohydr. Res.*, 1973, **29**, 123.
56. J. C. Decius, *J. Chem. Phys.*, 1948, **16**, 1025.
57. S. Melberg and K. Rasmussen, *Carbohydr. Res.*, 1979, **71**, 25.
58. B. Henrissat, S. Perez, I. Tvaroska, and W. T. Winters, in "The Structures of Celluloses," ed. H. Atalla, ACS Symposium Series No. 340, American Chemical Society, Washington, DC, 1987 p. 38.
59. R. H. Atalla, *Adv. Chem. Ser.*, 1979, **181**, 55.
60. R. H. Atalla, *Appl. Polym. Symp.*, 1976, **28**, 659.
61. R. H. Atalla, J. C. Gast, D. W. Sindorf, V. J. Bartuska, and G. E. Maciel, *J. Am. Chem. Soc.*, 1980, **102**, 3249.
62. D. Page, personal communication.
63. R. H. Atalla, in "Proceedings of the International Symposium on Wood and Pulping Chemistry," SPCI Rept. No. 38, Stockholm, 1981, vol. 1, p. 57.
64. R. H. Atalla, in "Structure, Function and Biosynthesis of Plant Cell Walls," eds. W. M. Dugger and S. Bartinicki-Garcia, American Society of Plant Physiologists, Rockville, MD, 1984, p. 381.
65. R. H. Atalla, *J. Appl. Pol. Symp.*, 1983, **37**, 295.
66. W. L. Earl and D. L. VanderHart, *J. Am. Chem. Soc.*, 1980, **102**, 3251.
67. W. L. Earl and D. L. VanderHart, *Macromolecules* 1981, **14**, 570.
68. D. L. VanderHart, Deductions about the morphology of wet and wet beaten cellulose from solid state ^{13}C NMR, NBSIR 82-2534, National Bureau of Standards, 1982.
69. R. L. Dudley, C. A. Fyfe, P. J. Stephenson, Y. Deslandes, G. K. Hamer, and R. H. Marchessault, *J. Am. Chem. Soc.*, 1983, **105**, 2469.
70. F. Horii, A. Hirai, and R. Kitamaru, *Polym. Bull.*, 1982, **8**, 163.
71. G. E. Maciel, W. L. Kolodziejski, M. S. Bertran, and B. E. Dale, *Macromolecules*, 1982, **15**, 686.
72. J. C. Gast, R. H. Atalla, and R. D. McKelvey, *Carbohydr. Res.*, 1980, **84**, 137.
73. D. L. VanderHart and R. H. Atalla, *Macromolecules*, 1984, **17**, 1465.
74. R. H. Atalla and D. L. VanderHart, *Science* 1984, **223**, 283.
75. D. L. VanderHart and R. H. Atalla, in "The Structures of Celluloses," ed. R. H. Atalla, ACS Symposium Series No. 340, American Chemical Society, Washington, DC, 1987 p. 88.
76. R. H. Atalla, R. E. Whitmore, and D. L. VanderHart, *Biopolymers*, 1985, **24**, 421.
77. F. Horii, H. Yamamoto, R. Kitamaru, M. Tanahashi, and T. Higuchi, *Macromolecules*, 1987, **20**, 2946.
78. J. H. Wiley and R. H. Atalla, in "The Structures of Celluloses," ed. R. H. Atalla, ACS Symposium Series No. 340, American Chemical Society, Washington, DC, 1987 p. 151.
79. J. Sugiyama, R. Yuong, and H. Chanzy, *Macromolecules*, 1991, **24**, 4168.
80. H. Yamamoto, F. Horii, and H. Odani, 1989, **22**, 4130.
81. J. Sugiyama, T. Okano, H. Yamamoto, and F. Horii, *Macromolecules*, 1990, **23**, 3196.
82. R. H. Atalla, J. Hackney, U. P. Agarwal, and A. Isogai, to be published.
83. D. I. Bower and W. F. Maddams, "The Vibrational Spectroscopy of Polymers," Cambridge University Press, Cambridge, 1989.
84. F. Horii, A. Hirai, and R. Kitamaru, in "The Structures of Celluloses," ed. R. H. Atalla, ACS Symposium Series No. 340, American Chemical Society, Washington, DC, 1987, p. 119.
85. D. L. VanderHart and G. C. Campbell, *J. Magn. Reson.*, 1998, **134**, 88.
86. H. Yamamoto and F. Horii, *Macromolecules*, 1993, **26**, 1313.
87. H. Yamamoto, F. Horii, and H. Odani, *Macromolecules*, 1989, **22**, 4130.
88. H. Yamamoto and F. Horii, *Macromolecules*, 1993, **26**, 1313.
89. P. S. Belton, S. F. Tanner, N. Cartier, and H. Chanzy, *Macromolecules*, 1989, **22**, 1615.
90. H. Yamamoto and F. Horii, *Cellulose*, 1994, **1**, 57.
91. R. H. Newman and J. A. Hemmingson, *Holzforschung*, 1990, **44**, 351.
92. R. H. Newman and J. A. Hemmingson, *Cellulose*, 1995, **2**, 95.
93. R. H. Newman, *J. Wood Chem. Tech.*, 1994, **14**, 451.
94. R. H. Newman, M. A. Ha, and L. D. Melton, *J. Agric. Food. Chem.*, 1994, **42**, 1402
95. R. H. Newman, L. M. Davies, and P. J. Harris, *Plant Physiol.*, 1996, **111**, 475.
96. R. H. Newman, *Cellulose*, 1997, **4**, 269.
97. H. Lenholm, T. Larsson, and T. Iversen, *Carbohydr. Res.*, 1994, **261**, 119.
98. H. Lennholm, T. Larsson, and T. Iversen, *Carbohydr. Res.*, 1994, **261**, 119.
99. T. Larsson, U. Westermark, and T. Iversen, *Carbohydr. Res.*, 1995, **278**, 339.
100. T. Larsson, K. Wickholm, and T. Iversen, *Carbohydr. Res.*, 1997, **302**, 19.
101. C.W. Hock, in "Cellulose and Cellulose Derivatives, Pt. 1," eds. E. Ott, H. M. Spurlin, and M. W. Grafflin, Wiley, New York, 1954, p. 347.
102. F. F. Morehead, in "Cellulose and Cellulose Derivatives, Pt. IV," eds. N. M. Bikales and L. Segal, Wiley, New York, 1971, 213.
103. K. Hieta, S. Kuga, and M. Usuda, *Biopolymers*, 1984, **23**, 1807.
104. H. Chanzy and B. Henrissat, *FEBS Lett.*, 1985, **184**, 285.

105. A. Maurer and D. Fengel, *Holz als Roh- und Werkstoff*, 1992, **50**, 493.
106. J. Sugiyama, H. Harada, Y. Fujiyoshi, and N. Uyeda, *Mokuzai Gakkaishi*, 1984, **30**, 98.
107. J. Sugiyama, H. Harada, Y. Fujiyoshi, and N. Uyeda, *Mokuzai Gakkaishi*, 1985, **31**, 61.
108. J. Sugiyama, H. Harada, Y. Fujiyoshi, and N. Uyeda, *Planta*, 1985, **166**, 161.
109. J. Sugiyama, T. Okano, H. Yamamoto, and F. Horii, *Macromolecules*, 1990, **23**, 3196.
110. G. Honjo and M. Watanabe, *Nature*, 1958, **181**, 326.
111. J. Sugiyma, R. Vuong, and H. Chanzy, *Macromolecules*, 1991, **24**, 4168.
112. E. Roche and H. Chanzy, *J. Biol. Macromol.*, 1981, **3**, 201.
113. B. J. Hardy and A. Sarko, *Polym. Prepr.*, 1995, **36**, 640.
114. L. M. J. Kroon-Batenburg, B. Bouma, and J. Kroon, *Macromolecules*, 1996, **29**, 5695.
115. L. M. J. Kroon-Batenburg and J. Kroon, *Glycoconjugate J.*, 1997, **14**, 677.
116. A. P. Heiner, J. Sugiyama, and O. Teleman, *Carbohyd. Res.*, 1995, **273**, 207.
117. A. P. Heiner and O. Teleman, *Langmuir*, 1997, **13**, 511.
118. M. S. Baird, A. C. O'Sullivan, and W. B. Banks, *Cellulose*, 1998, **5**, 89.
119. H. Chanzy, B. Henrissat, M. Vincendon, S. Tanner, and P. S. Belton, *Carbohydr. Res.*, 1987, **160**, 1.
120. R. H. Atalla, B. E. Dimick, and S. C. Nagel, in "Cellulose Chemistry and Technology," ed. J. C. Arthur, Jr., ACS Symp. Series No. 40, 1977, p. 30.
121. R. H. Atalla, J. D. Ellis, and L. R. Schroeder, *J. Wood Chem. Technol.*, 1984, **4**, 465.
122. H. Chanzy, K. Imada, A. Mollard, R. Vuong, and F. Barnoud, *Protoplasma*, 1979, **100**, 303.
123. A. Sakthivel, A. D. French, B. Eckhardt, and R. A. Young, in "The Structures of Celluloses," ed. R. H. Atalla, ACS Symposium Series No. 340, American Chemical Society, Washington, DC, 1987, p. 68.
124. L. D. Landau and E. M. Lifshitz, "Statistical Physics", Addison-Wesley, Reading, MA, 1958.
125. R. H. Atalla and S. C. Nagel, *Science*, 1974, **185**, 522.
126. R. E. Whitmore and R. H. Atalla, *Int. J. Biol. Macromol.*, 1985, **7**, 182.
127. N. R. Bertoniere and S. H. Zeronian, in "The Structures of Celluloses," ed. R. H. Atalla, ACS Symposium Series, No. 340, American Chemical Society, Washington, DC, 1987, p. 255.
128. S. P. Rowland and E. J. Robserts, *J. Polym. Sci., A-1*, 1972, **10**, 2447.
129. R. F. Nickerson, *Text. Res. J.*, 1951, **21**, 195.
130. S. P. Rowland, E. J. Roberts, and C. P. Wade, *Text. Res. J.*, 1969, **39**, 530.
131. S. P. Rowland, in "Modified Cellulosics," eds. R. M. Rowell and R. A. Young, Academic Press, New York, 1978, p. 147.
132. S. P. Rowland, E. J. Roberts, and J. L. Bose, *J. Polym. Sci., A-1*, 1971, **9**, 1431.
133. S. P. Rowland, E. J. Roberts, J. L. Bose, and C. P. Wade, *J. Polym. Sci., A-1*, 1971, **9**, 1623.
134. S. H. Zeronian, M. L. Coole, K. W. Alger, and J. M. Chandler, *J. Appl. Polym. Sci., Appl. Pol. Symp.*, 1983, **37**, 1053.
135. J. E. Stone and A. M. Scallan, *Pulp. Pap. Mag. Can.*, 1968, **69**, 69
136. J. E. Stone, E. Treiber, and E. Abrahamson, *TAPPI*, 1969, **52**, 108.
137. R. P. Bell, "The Proton in Chemistry," Chapman & Hall, London, 2nd edn., 1973.
138. D. C. Johnson, M. D. Nicholson, and F. C. Haigh, *Appl. Polym. Symp.*, 1976, **28**, 931
139. D. C. Johnson and M. D. Nicholoson, *Cellul. Chem. Technol.*, 1977, **11**, 349.
140. D. P. Delmer, M. Benziman, and E. Padan, *Proc. Natl. Acad. Sci. USA*, 1982, **79**, 5282.
141. M. Benziman, C. H. Haigler, R. M. Brown, Jr., A. R. White, and K. M. Cooper, *Proc. Natl. Acad. Sci. USA*, 1980, **77**, 6678.
142. Y. Aloni and M. Benziman, in "Cellulose and Other Natural Polymer Systems," ed. R. M. Brown, Plenum, New York, 1982, p. 341.
143. Y. Aloni, D. P. Delmer, and M. Benziman, *Proc. Natl. Acad. Sci. USA*, 1982, **79**, 6448.
144. M. Benziman, Y. Aloni, and D. P. Delmer, *J. Appl. Polym. Sci., Appl. Pol. Symp.*, 1983, **37**, 131
145. Y. Aloni, R. Cohen, M. Benziman, and D. P. Delmer, *J. Biol. Chem.*, 1983, **258**, 4419.
146. H. Hori and A. D. Elbein, in "Biosynthesis and Biodegradation of Wood Components," ed. T. Higuchi, Academic Press, New York, 1985, p. 109.
147. Y. Kawagoe and D. P. Delmer, in "Genetic Engineering," ed. J. K. Setlow, Plenum, New York, 1997, vol. 19, p. 63.
148. D. H. Northcote, in "Biosynthesis and Biodegradation of Cellulose," eds. C. H. Haigler and P. J. Weimer, Marcel Dekker, New York, 1991, p. 165.
149. P. Ross, Y. Aloni, C. Weinhouse, D. Michaeli, P. Weinberger-Ohana, R. Mayer, and M. Benziman, *FEBS Lett.*, 1985, **186**, 191.
150. P. Ross, Y. Aloni, C. Weinhouse, D. Michaeli, P. Weinberger-Ohana, R. Mayer, and M. Benziman, *Carbohyd. Res.*, 1986, **149**, 101.
151. P. Ross, C. Weinhouse, Y. Aloni, D. Michaeli, P. Weinberger-Ohana, R. Mayer, S. Braun, E. de Vroom, G. A. van der Marel, J. H. van Boom, and M. Benziman, *Nature*, 1987, **325**, 279.
152. P. Ross, R. Mayer, and M. Benzima, in "Biosynthesis and Biodegradation of Cellulose," eds. C. H. Haigler and P. J. Weimer, Marcel Dekker, New York, 1991, p. 219.
153. J. R. Pear, Y. Kawagoe, W. E. Schreckengost, D. P. Delmer, and D. M. Stalker, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 12637.
154. T. Arioli, L. Peng, A. S. Betzner, J. Burn, W. Wittke, W. Herth, C. Camilleri, H. Hofte, J. Plazinski, R. Birch, A. Cork, J. Glover, J. Redmond, and R. E. Williamson, *Science*, 1998, **279**, 717.
155. B. A. Tonnesen and O. Ellefsen, in "Cellulose and Cellulose Derivatives," Pt. IV, eds. N.M. Bikales and L. Segal, Wiley, New York, 1971, vol. 5, p. 265.
156. R. H. Atalla, *Science*, 1998, **282**, 591.
157. J. R. Colvin, in "The Fromation of Wood in Forest Trees," ed. M. H. Zimmerman, Academic Press, New York, 1964, p. 189.
158. J. R. Colvin, in "Cellulose and Cellulose Derivatives," Pt. IV, eds. N.M. Bikales and L. Segal, Wiley, New York, 1971, p. 695.
159. I. Ohad, D. Danon, and S. Hestrin, *J. Cell Biol.*, 1962, **12**, 31.
160. G. Ben-Hayim and I. Ohad, *J. Cell Biol.*, 1965, **25**, 191.

161. R. M. Brown, Jr., J. H. M. Willison, and C. L. Richardson, *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 4565.
162. T. E. Bureau and R. M. Brown, Jr., *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 6985.
163. C. H. Haigler and R. M. Brown, Jr., *J. Cell Biol.*, 1979, **83**, 70a.
164. C. H. Haigler, R. M. Brown, Jr., and M. Benziman, *Science*, 1980, **210**, 903.
165. M. Benziman, C. H. Haigler, R. M. Brown, Jr., A. R. White, and K. M. Cooper, *Proc. Natl. Acad. Sci. USA*, 1980, **77**, 6678.
166. R. M. Brown, Jr., C. H. Haigler, and K. Cooper, *Science*, 1982, **218**, 1141.
167. R. M. Brown, Jr., C. H. Haigler, J. Suttie, A. R. White, E. M. Roberts, C. A. Smith, T. Ito, and K. H. Cooper, *J. Appl. Polym. Sci., Appl. Polym. Symp.*, 1983, **37**, 33.
168. R. M. Brown and D. Montezinos, *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 143.
169. H. Quader, in "Biosynthesis and Biodegradation of Cellulose," eds. C. H. Haigler and P. J. Weimer, Marcel Dekker, New York, 1991, p. 51.
170. A. M. C. Emons, in "Biosynthesis and Biodegradation of Cellulose," eds. C. H. Haigler and P. J. Weimer, Marcel Dekker, New York, 1991, p. 71.
171. C. H. Haigler, in "Biosynthesis and Biodegradation of Cellulose," eds. C. H. Haigler and P. J. Weimer, Marcel Dekker, New York, 1991, p. 99.
172. K. I. Uhlin, Doctoral Dissertation, The Institute of Paper Science and Technology, Atlanta, GA, 1990.
173. K. I. Uhlin, R. H. Atalla, and N. S. Thompson, *Cellulose*, 1995, **2**, 129.
174. R. H. Atalla, J. M. Hackney, I. Uhlin, and N. S. Thompson, *Int. J. Biol. Macromol.*, 1993, **15**, 109.
175. J. M. Hackney, R. H. Atalla and D. L. VanderHart, *Int. J. Biol. Macromol.*, 1994, **16**, 215
176. A. Hirai, M. Tsuji, and F. Horii, *Cellulose*, 1997, **4**, 239.
177. T. Imai and J. Sugiyama, *Macromolecules*, 1998, **31**, 6275.
178. T. Okamoto, J. Sugiyama, and T. Itoh, *Wood Res.*, 1996, **83**, p. 27.
179. N. C. Carpita and D. M. Gibeaut, *Plant J.*, 1993, **3**, 1.
180. R. H. Atalla and S. C. Nagel, *J. Polym. Sci., Polym. Lett.*, 1974, **12**, 565.
181. R. H. Atalla and R. Whitmore, *J. Polym. Sci., Polym. Lett.*, 1978, **16**, 601.
182. E. Heuser, in "Nature of the Chemical Components of Wood," ed. C. J. West, TAPPI Monograph Series No. 6, Technical Association of the Pulp and Paper Industry, New York, 1948, p. 8.
183. R. H. Atalla and J. M. Hackney, in "Hierarchically Structured Materials," eds. I. A. Aksay, E. Baer, M. Sarikaya, and D. A. Tirrell, Materials Research Society Symposium Proceedings, 1992, vol. 255, p. 387.
184. R. H. Atalla, J. M. Hackney, and D. L. VanderHart, unpublished.
185. R. H. Newman, M. A. Ha, and L. D. Melton, *J. Agric. Food Chem.*, 1994, **42**, 1402.
186. R. H. Newman, L. M. Davies, and P. J. Harris, *Plant Physiol.*, 1996, **111**, 475.
187. R. H. Newman, *Cellulose*, 1997, **4**, 269.
188. R. H. Newman, *J. Wood Chem. Technol.*, 1994, **14**, 451.
189. T. M. Wood, in "Biosynthesis and Biodegradation of Cellulose," eds. C. H. Haigler and P. J. Weimer, Marcel Dekker, New York, 1991, p. 491.
190. P. Rapp and A. Beermann, in "Biosynthesis and Biodegradation of Cellulose," eds. C. H. Haigler and P. J. Weimer, Marcel Dekker, New York, 1991, p. 535.
191. P. Tomme, R. A. J. Warren, and N. R. Gilkes, *Adv. Microb. Physiol.*, 1995, **37**, 1.
192. T. T. Teeri, *Trends Biotechnol.*, 1997, **15**, 160.
193. E. A. Bayer, E. Morag, R. Lamed, S. Yaron, and Y. Shoham, in "Carbohydrases from *Trichoderma reesei* and other Organisms," eds. M. Clayssens, W. Nerinckx, and K. Piens, The Royal Society of Chemistry, London, 1998, p. 39.
194. E. T. Reese and M. Mandels, in "Cellulose and Cellulose Derivatives," Pt. V, eds. N.M. Bikales and L. Segal, Wiley, New York, 1971, p. 1079.
195. C. S. Walseth, *TAPPI*, 1952, **35**, 228.
196. K.-E. Eriksson, *Adv. Chem. Series*, 1969, **95**, 83.
197. K.-E. Eriksson and B. Petterssen, *Eur. J. Biochem.*, 1975, **51**, 213.
198. K.-E. Eriksson and T. M. Wood, in "Biosynthesis and Biodegradation of Wood Components," ed. T. Higuchi, Academic Press, New York, 1985, p. 469.
199. P. F. G. Sims, M. S. Soares-Felipe, Q. Wang, M. E. Gent, C. Tempelaars, and P. Broda, *Mol. Microbiol.*, 1994, **12**, 209.
200. M. A. Vallim, B. J. H. Janse, J. Gaskell, A. A. Pizzirani-Kleiner, and D. Cullen, *Appl. Environ. Microbiol.*, 1998, **64**, 1924.
201. D. Cullen and R. H. Atalla, unpublished.
202. R. Lamed, J. Naimark, E. Morgenstern, and E. A. Bayer, *J. Bacteriol.*, 1987, **169**, 3792.
203. R. Lamed and E. A. Bayer, in "Biochemistry and Genetics of Cellulose Degradation," eds. J.-P. Aubert, P. Beguin, and J. Millet, Academic Press, London, 1988, p. 101
204. R. Lamed and E. A. Bayer, in "Biosynthesis and Biodegradation of Cellulose," eds. C. H. Haigler and P. J. Weimer, Marcel Dekker, New York, 1991, p. 377.
205. E. A. Bayer, E. Morag, and R. Lamed, *Trends Biotechnol.*, 1994, **12**, 379.
206. P. Beguin and J.-P. Aubert, *FEMS Microbiol. Rev.*, 1994, **13**, 25.
207. C. R. Felix and L. G. Ljungdahl, *Ann. Rev. Microbiol.*, 1993, **47**, 791.
208. N. Hayashi, J. Sugiyama, T. Okano, and M. Ishihara, *Carbohydr. Res.*, 1998, **305**, 109.
209. N. Hayashi, J. Sugiyama, T. Okano, and M. Ishihara, *Carbohydr. Res.*, 1998, **305**, 261.
210. G. Maclachlan and S. Carrington, in "Biosynthesis and Biodegradation of Cellulose," eds. C. H. Haigler and P. J. Weimer, Marcel Dekker, New York, 1991, p. 599.

Comprehensive Natural Products Chemistry

Editors-in-Chief
Sir Derek Barton†
Texas A&M University, USA
Koji Nakanishi
Columbia University, USA

Executive Editor
Otto Meth-Cohn
University of Sunderland, UK

Volume 3

CARBOHYDRATES AND THEIR DERIVATIVES INCLUDING TANNINS, CELLULOSE, AND RELATED LIGNINS

Volume Editor
B. Mario Pinto
Simon Fraser University, Canada

1999
ELSEVIER

AMSTERDAM - LAUSANNE - NEW YORK -OXFORD- SHANNON -SINGAPORE -TOKYO

Elsevier Science Ltd., The Boulevard, Langford Lane, Kidlington, Oxford, OX5 IGB, UK

First edition 1999

Library of Congress Cataloging-in-Publication Data
Comprehensive natural products chemistry /editors-in-chief. Sir Derek Barton, Koji Nakanishi : executive editor, Otto Meth-Cohn. -- 1st ed.
p. cm.
Includes index.
Contents: v. 3. Carbohydrates and their derivatives including tannins, cellulose. and related lignins /volume editor B. Mario Pinto
1. Natural products. I. Barton, Derek, Sir. 1918-1998. II. Nakanishi, Koji. 1925- . III. Meth-Cohn, Otto.
QD415.C63 1999
547.7--dc21 98-15249

British Library Cataloguing in Publication Data
Comprehensive natural products chemistry
1. Organic compounds
I. Barton, Sir Derek, 1918-1998 II. Nakanishi Koji III. Meth-Cohn Otto
572.5

ISBN 0-08-042709-X (set : alk. paper)
ISBN 0-08-043155-0 (Volume 3 : alk. Paper)

♾™ 'The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences–Permanence of Paper for Printed Library Materials, ANSIZ39.48-1984

Typeset by BPC Digital Data Ltd., Glasgow, UK.
Printed and bound in Great Britain by BPC Wheatons Ltd., Exeter, UK.