ELSEVIER

Solid State Nuclear Magnetic Resonance 15 (1999) 1–19

SOLID STATE
Nuclear Magnetic
Resonance

www.elsevier.nl/locate/solmag

The role of solid state ^{13}C NMR spectroscopy in studies of the nature of native celluloses

R.H. Atalla [a,*], D.L. VanderHart [b]

[a] *USDA Forest Service, Forest Products Laboratory, Chemistry and Pulping Research, 1 Gifford Pinchot Drive, Madison, WI 53705-2398, USA*

[b] *Polymers Division, National Institute for Standards and Technology, Gaithersburg, MD 20899-8544, USA*

Received 8 February 1999; accepted 15 July 1999

Abstract

Published spectroscopic observations pertaining to the crystal structure of native celluloses are reviewed for the purpose of defining our current level of understanding about crystalline polymorphism in these materials. Emphasis is placed on observations from solid state ^{13}C nuclear magnetic resonance (NMR), which first led to the postulate that most native, semicrystalline celluloses are composites of two crystalline allomorphs, labeled I_α and I_β. Historical background is presented, highlighting the structural controversies which mainly arose because different native celluloses were used, each one representing a different mixture of allomorphs. Input from Raman, infrared (IR) and electron diffraction data is included in the discussion of our current understanding of polymorphism in native celluloses. Also noted is the input from more recently studied celluloses (e.g., *Halocynthia*) as well as from newer processes that convert the I_α to the I_β form. On the basis of Raman and IR observations, it is argued that the I_α and I_β allomorphs differ in hydrogen bonding patterns only and that backbone conformations are nearly identical. Also, the point is made that the absence of correlation field splittings in the Raman spectra calls into question (although it does not disprove) whether the normal two-chain-per-unit-cell, monoclinic I_β allomorph really possesses two equivalent chains. Considerable discussion is devoted to the allomorphic composition of cellulose crystallites in higher plants. Published methods of NMR lineshape analysis for the higher plant celluloses are reviewed and critiqued, both from the point of view of lineshape theory and from the point of view of self-consistency of inferences that are based on lineshape analyses for different carbons (particularly C1 and C4). It is concluded that higher plant celluloses most likely possess a minor amount of the I_α allomorph where the I_α/I_β ratio is probably less than 0.25.

Keywords: Cellulose; Native; Carbon-13; NMR; Solid state; Structure; Allomorphs; Raman; Infrared; Electron diffraction; Review

1. Introduction

In the past three decades, our understanding of the diversity of native celluloses has been advanced in large measure due to the possibility of explorations of their structures through the methods of cross-polarization/magic angle spinning (CP/MAS) ^{13}C nuclear magnetic resonance (NMR) spectroscopy. These explorations led to recognition once again of the individuality of native celluloses from different species. Early in the century, the uniqueness of

* Corresponding author. Tel.: +1-608-231-9443; fax: +1-608-231-9262; e-mail: rhatalla@facstaff.wisc.edu

0926-2040/99/$19.00
PII: S0926-2040(99)00042-9

source species had led to serious questions concerning the relevance of crystallography in characterizing native celluloses. Cross and Bevan, among the early pioneers in investigations of the chemistry of cellulose, had written: ''The root idea of crystallography is identical invariability, the root idea of the world of living matter is essential individual variation'' [1]. Yet ever since the introduction of the polymer hypothesis early in the century, studies of the structures of the naturally occurring native celluloses were based on the premise that they all represent variations on a structure that is common to all native forms; it was identified as cellulose I. Though it was recognized that the results of crystallographic studies of the more highly ordered types of native celluloses from different sources were often not consistent with each other, little progress was made towards resolving the inconsistencies. When new spectroscopic methods became available and they were applied to the celluloses, additional questions arose that could not be rationalized on the basis of the then available crystallographic models. The resolution of the dilemma was accomplished in the early 1980s when solid state ^{13}C NMR spectral studies led to the conclusion that all native celluloses are composites of two different forms of cellulose I. In this report, we review the development of the models of cellulose that have arisen from the solid state ^{13}C NMR studies and look ahead to new directions for exploration of native celluloses that are now possible because of these findings.

2. Background

The evolution of ideas concerning the crystallography of native celluloses during the different phases of this century has been reviewed more extensively elsewhere [2]. The early 1970s were marked by the reintroduction of unit cell models based on parallel alignment of the cellulose molecular chains [3,4]. These models were not unlike those abandoned by Meyer and Misch [5] in the 1930s on the basis of premises that are now clearly questionable. The new models, however, also incorporated bending of the glycosidic linkage in a manner not envisioned by Meyer and Mark [6]. The bending of the glycosidic linkages was introduced to allow for an intramolecular hydrogen bond as suggested by Hermanns [7]. 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 [8] pointed out, they were also not strongly preferred over an antiparallel structure. In the analysis by French [8], 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 forms of native celluloses was revealed through the high-resolution solid state ^{13}C NMR investigations.

The new crystallographic models remained in question also because the analyses on which they were based assumed a level of symmetry in the unit cell that was inconsistent with some of the diffractometric data. Some of the reflections that are consistently observed in electron diffraction patterns and are disallowed by the selection rules for space group $P2_1$ [9] 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. The spectral information came from two sources, first Raman spectroscopy and later, solid state ^{13}C NMR.

The Raman spectral data were not consistent with the crystallographic results in two ways. First, it was clear that the crystallographic structures for celluloses I and II could not provide a basis for explaining the differences in the Raman spectra in the conformation sensitive region. Both structures were assumed to possess the symmetry of $P2_1$ with the twofold screw axis coincident with the cellulose chain. The Raman spectra, in contrast, indicated a change in the skeletal conformation upon conversion from I to II [10]. In addition, it emerged from analysis of the Raman spectra that the repeat unit in

the structures of cellulose is the dimeric anhydrocellobiose [11]; the two anhydroglucose units within the repeat unit are not symmetrically equivalent as is implicit in the assumption of a structure possessing $P2_1$ symmetry. The Raman spectra, together with conformational energy calculations, pointed to alternating non-equivalent glycosidic linkages along the chain. The alternate linkages envisioned represent small left-handed and right-handed departures from the structure with twofold screw axis symmetry. The overall structure was still regarded as a ribbon-like cellulose molecule with *approximate* twofold screw axis symmetry. If the $P2_1$ symmetry were precise, the nonequivalent residues would require that the twofold axis be located between the chains, thus also requiring that the chains be parallel. But without allowance for the small departures at the linkages, the Raman spectra could not be rationalized. The solid state ^{13}C NMR spectra that are the main concern of this report are consistent with such an interpretation though they do not represent conclusive evidence.

In the mid-1980s, electron microscopic studies based on new staining techniques, specific to the reducing end groups of the polysaccharides, provided evidence supporting the parallel alignment of molecular chains within the microfibrils in native celluloses [12]. 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 clearly were active at only one end of each microfibril [13]. 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.

3. Solid state ^{13}C NMR spectra of the forms of native cellulose

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. The most important characteristic of the spectra acquired using the (CP/MAS) ^{13}C NMR method 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, i.e., if they occur in different environments or if the anhydroglucose rings possess different conformations, within the rings or 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 is to achieve a level of resolution sufficient to distinguish nonequivalences between chemically equivalent carbons, because the magnetic non-equivalence 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 spectra 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 [14,15] 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 [14–20], were based on comparisons with solution spectra of cello-oligosaccharides and of a low DP cellulose [21]. They are indicated in Fig. 1, which shows a spectrum of cotton linters [22]. Beginning at the upfield part of

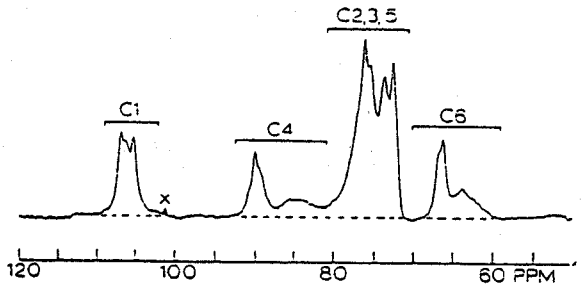

Fig. 1. The 15-MHz ^{13}C CP/MAS spectrum of cotton linters. The horizontal bars indicate the spectral ranges of the corresponding carbon sites in the anhydroglucose monomer unit of cellulose. The "X" marks the position of the small first spinning sideband of linear polyethylene, which was added as a chemical shift reference. The polyethylene center band (not shown) occurs at 33.63 ppm; the zero for reference of chemical shifts is liquid tetramethylsilane. Note the existence of both broad and narrow resonance features. Adapted from Ref. [22].

the spectrum, the region between 60 and 70 ppm is assigned to C6 of the primary alcohol group. The next cluster of resonances, between 70 and 81 ppm, is attributed to C2, C3, and C5, the ring carbons other than those anchoring the glycosidic linkage. The region between 81 and 93 ppm is associated with C4 and that between 102 and 108 ppm with C1, the anomeric carbon.

In one of the first reports on the application of the technique to studies of different celluloses, the splittings of the resonances of C4 and C1 in the spectrum of cellulose II Fig. 2 were regarded as confirmation of the occurrence of nonequivalent glycosidic linkages that had earlier been proposed on the basis of the comparison of the Raman spectra of cellulose II and of cellobiose in the O–H stretching region [14]. 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 [18]. However, such an interpretation leaves open the question as to why the resonances for carbons 2, 3, and 5, do not display splittings of similar magnitude. 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, i.e., C1 and C4, 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 in Fig. 1, the two resonance regions associated with C6 and C4 include sharper resonances overlapping broader upfield wings. After excluding the possibility that the broader wings could arise entirely from molecular mobility [15,16], 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 argued for in a study of different native celluloses undertaken by Earl and VanderHart [16]. 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 C4 and C6 varied approximately as the surface-to-volume ratio. 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 conforma-

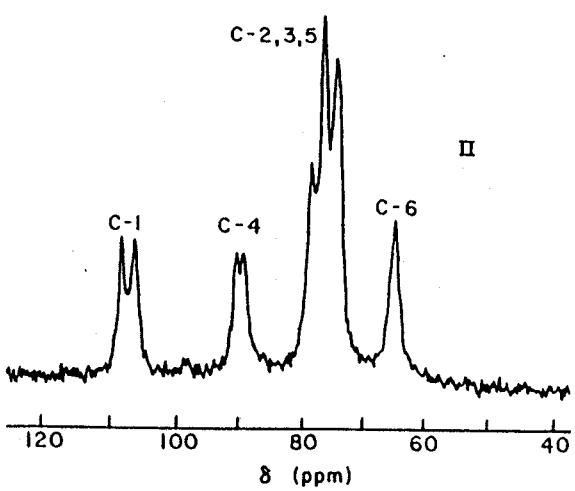

Fig. 2. The 15-MHz ^{13}C CP/MAS spectrum of high crystallinity cellulose II. Chemical shifts are relative to tetramethylsilane. Assignment of the spectra is based on comparisons with pertinent liquid state spectra. Adapted from Ref. [14].

tional 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 C4 and C6 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 C1, 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 C2, C3, and C5 cluster.

The relative contributions of the two categories of environment to the intensity of the upfield wings was assessed [22] in a careful analysis of the C4 wing of a cotton linters spectrum. It was demonstrated that part of the wing could be correlated with the range of the C4 resonance in an amorphous cellulose prepared by ball milling. That part was therefore assigned to cellulose chains occurring in the second type of environment, i.e., 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 C4 is the basis of this allocation of intensities, it can be assumed that the relative contributions are similar for the upfield wing of C6 and for the component that appears to underlie the sharper resonances at C1. It is also expected that these domains contribute to the total intensity of the C2, C3, and C5 cluster between 70 and 81 ppm.

The sharper resonances in the C6 and C4 regions, centered at 66 and 90 ppm, respectively, each appear to have multiplet character even though the resolution is not sufficient to distinguish the multiplet components well. The C6 resonance seems to include at least two, and possibly three, components while the C4 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 C1, also reveals multiplicity and sharp resonance features. However, overlap of the C1 resonances of the crystalline carbons with those of the less-ordered regions complicates interpretation. The lesser influence of the degree of order on the chemical shift of C1 carbons, relative to C4 or C6 carbons, suggests the possibility that, because of the anomeric effect, the internal coordinates surrounding C1 are much less flexible within the range of possible conformational variations than are the internal coordinates around C4 or C6.

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 Fig. 3. They include ramie fibers (A), cotton linters (B), hydrocellulose prepared from cotton linters by acid hydrolysis (C), a low DP regenerated cellulose I (D), cellulose from *Acetobacter xylinum* (E), and cellulose from the cell wall of *Valonia ventricosa*, an alga (F). While similar observations were reported in a number of studies [14,16–20,23], their implications with respect to structure were more fully developed in the work of VanderHart and Atalla [22] and Atalla and VanderHart [24], which provides the basis for the following discussion.

All of the spectra shown in Fig. 3A–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 C4 and C6. 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

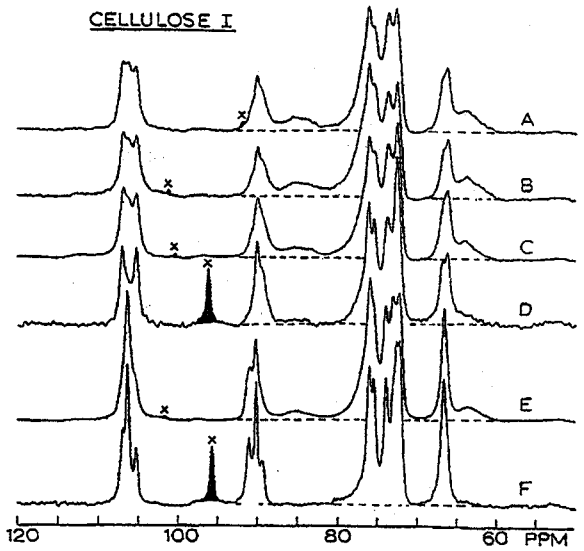

Fig. 3. The 50-MHz ^{13}C CP/MAS spectra of several cellulose I samples: (A) ramie; (B) cotton linters; (C) hydrocellulose from cotton linters; (D) a low DP regenerated cellulose I; (E) *A. xylinum* cellulose; (F) *V. ventricosa* cellulose. Note the varied fine structure particularly at C1 and C4. Signal-to-noise variation due to limited amount of some samples; in those instances, more polyethylene was added so the sideband intensity increased. No line broadening or resolution enhancement techniques were applied in the acquisition of the spectra. Adapted from Ref. [22].

rationalization was not possible because the relative intensities within the multiplets, particularly at C4, 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 of any one species. 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 form.

Further rationalization of the spectra required a careful analysis of the mutliplets at C4 and C6 and the variations of the relative intensities of the lines within each multiplet among the spectra of the different celluloses. The shape of the resonances of C1 was also a factor in these analyses, in spite of the overlap problem alluded to earlier, since the chemical shift spread of multiplet components was comparable to that at C4. 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 contributing a line to the more complex multiplets. Thus, a decomposition of the spectra on the basis of two distinct ordered forms was pursued. The results of the decomposition are shown as spectra (B) and (C) in Fig. 4, and were designated as the I_α and I_β 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 the infrared (IR) spectra of different types of native celluloses [25,26]. 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

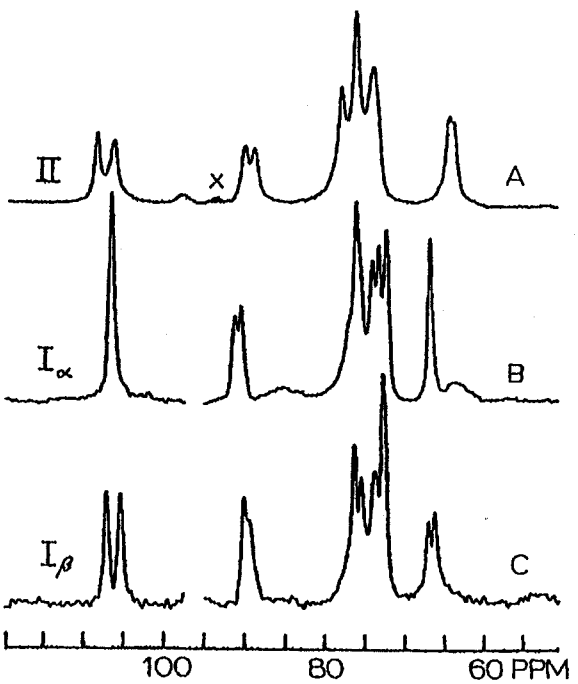

Fig. 4. Comparison of the ^{13}C CP/MAS spectrum of a low DP cellulose II sample and the spectra b and c corresponding, respectively, to the two proposed crystalline forms of cellulose I, namely I_α and I_β. Spectra b and c were obtained by taking linear combinations of the low DP and *A. xylinum* cellulose spectra shown in Fig. 3. 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 C4 and C6 upfield wings. Multiplicities of the C1, C4, and C6 narrower resonances ought to indicate unit cell inequivalences. Adapted from Ref. [22].

cellulose I from the long-known polymorphic variation of the crystallinity of cellulose.

Spectra B and C in Fig. 4 were, in fact, derived from appropriate linear combinations of the spectra of the low DP cellulose I (D) and of the *A. xylinum* cellulose (E) in Fig. 3. 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 reject the resonances arising from the two regions of less-ordered cellulose. Spectrum 4b does have some intensity in the upfield wings of C4 and C6, but spectrum 4c has very little evidence of such wings. There is very little question, however, that the sharp components of spectra 4b and c include the key features in the spectra of the I_α and I_β forms. It is of interest to note here that among the distinct resonances of the I_α form at C1, C4, and C6, only the one at C4 appears to be split, while for the I_β form, all three resonances associated with these carbons show splitting, with the one at C1 being the most pronounced. Nevertheless, the splittings observed indicate that both the I_α and I_β forms have a minimum of two magnetically inequivalent anhydroglucose residues per unit cell.

In an effort to further validate the proposal that the I_α and I_β forms were the primary constituents of native celluloses, VanderHart and Atalla [27] undertook another extensive study to verify that the sharper resonance features, particularly the multiplets at C1 and C4, were true indicators of the inequivalences within the unit cells. To accomplish this, they again considered the possibility that these multiplets might include contributions from ordered chains at crystallite surfaces. They also explored the possibility that the small shifts within a multiplet might have their origin in a particular packing geometry of the crystallites with the necessary added feature that these crystallites could exhibit a significantly anisotropic bulk magnetic susceptibility [28]. They concluded from several experimental indicators that the multiplet character of C4 in the 88–92 ppm range gave the most faithful indication of the inequivalences within the unit cell(s); i.e., susceptibility effects did not dictate multiplet structure and, furthermore, C4 carbons on residues at the crystal surface do not contribute intensity in the 88–92 ppm range. They also surveyed several spectra of higher plant celluloses with different native morphologies and concluded that the C4 multiplet shapes were quite similar, even though the definition of the relative multiplet intensities was limited, owing to low spectral resolution. In addition, it was observed that in the higher plant celluloses (in Fig. 1), the downfield shoulder of C4, a feature associated exclusively with the I_α component in Fig. 4, always appears to be about half as intense as the upfield shoulder of C4. According to the proposed I_α and I_β lineshapes, this ratio predicts that the corresponding three multiplet components at C1 ought to have about equal intensities; however, experimentally, they do not; i.e., in the crystalline contribution to the C1 lineshape, the central C1 multiplet (I_α) is generally much weaker than the other two C1 multiplet components in the spectra of the higher plant celluloses. Hence, the question was raised as to whether I_α occurs at all in these higher plant celluloses. Implicit in this question are the possibilities that: (a) the downfield shoulder of the C4 multiplet could be explained by some slight packing irregularities near, but not at, the surface of these I_β crystallites that have lateral dimensions of the order of 3–5 nm, or (b) if the C4 multiplet shape in the higher plant celluloses indicated the true unit cell for the I_β form, one would need more than four anhydroglucose residues per unit cell. Regardless of which is the true reason, the implication is clear that, from X-ray or electron-diffractometric investigations of a higher-plant cellulose sample, one cannot expect to extract a single, uniform structure. Certainly, subsequent publication [29] of the CP/MAS spectrum of *Halocynthia* cellulose, a tunicate cellulose representing a much more perfect form of the I_β allomorph, has eliminated the second of the foregoing possibilities as the central C1 multiplet does not occur in that spectrum; other implications of the tunicate spectrum will be discussed in a later section. Thus, the question remains open as to whether the downfield shoulder of C4 in the higher plant celluloses could be a "near-surface" contribution associated with imperfect I_β crystallites or whether it indicates a true mixture of the I_α and I_β forms.

Attention was then directed to analysis of the spectra of algal celluloses where the I_α component is, in general, dominant. Relaxation experiments confirmed that the essential spectral features identi-

fied with the two forms of cellulose were characteristic of the core crystalline domains. When measurements were conducted such that magnetization of the less-ordered regions, including surface domains, was first allowed to undergo relaxation, very little change in the relative intensities of the remaining sharper spectral features was observed. The relaxation experiments indicate that both the I_α and I_β crystalline domains have comparable proximity to crystal surfaces because they have similar, although very slow, longitudinal ^{13}C relaxation rates.

Two groups of modifying experiments were carried out with the *Cladophera* cellulose, an algal cellulose whose lateral fibrillar dimension is about 20 nm. This algal cellulose was first subjected to severe mechanical actions in a Waring blender. In the second experiment, this algal cellulose was subjected to harsh 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, the acid hydrolysis resulted in a dramatic reduction, sufficient indeed to make the spectrum resemble those of the higher plants, except that the resolution of the spectral lines remained excellent and at a level much higher than 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 very similar to those of the original samples. These observations were interpreted to imply that the I_α form is more susceptible to hydrolysis than the I_β form. Moreover, it seemed that the mass loss was not due to any thinning of the crystallites but rather due to the degradation of complete crystalline domains. This suggests that any variation of the I_α and I_β forms across the cross-section of a typical, 20-nm-wide crystallite, is minimal. An earlier study of the effect of hydrolysis, under similar conditions but for only 4 h, had been carried out with cellulose from *Rhizoclonium heiroglyphicum* with no discernible effect on the spectra [30]. The difference in duration of the hydrolysis may 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_β form [31].

VanderHart and Atalla [27] 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_β forms with significantly reduced contributions from regions of lower order. These are shown in Fig. 5 where it is clear that even in the spectrum representative of the I_α form, the upfield wings of the C4 and C6 resonances are reduced to a minimum.

Finally, VanderHart and Atalla [27] used the phenomenon of slow polarization exchange between two isolated natural-abundance ^{13}C nuclei within a radius of about 0.7–1.0 nm [32] in order to generate what was called "near-neighbor" spectra. The idea was straightforward, namely, that one initially prepare a state where the carbon polarization of one of the multiplet lines at C1 or C4 was very different from the average polarizations of all the other lines in the spectrum. Then the carbons at the unique site were allowed to undergo polarization exchange, for some long specified time, with whatever other kinds of carbons were within this defined radius. Finally, the spectrum of the carbons that had undergone exchange was isolated in order to seek support for the hypothesis that only carbons belonging to the I_α (or I_β) form would appear in the near-neighbor spectrum when the initially perturbed polarization resided in an exclusively I_α (or I_β) multiplet. These experiments strongly supported the crystal-composite model for the algal celluloses, even though the initial polarization states could not be prepared as cleanly as one

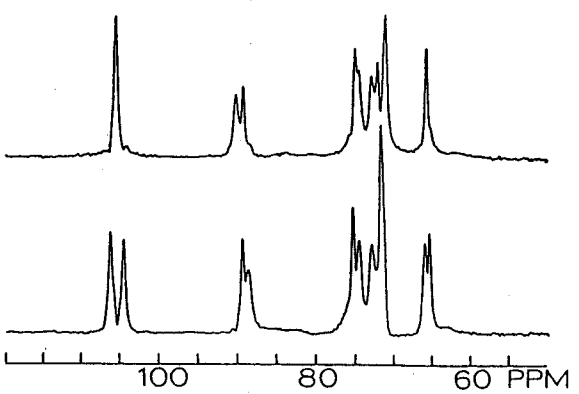

Fig. 5. Alternative candidates for the spectra of celluloses I_α (top) and I_β (bottom) derived from linear combinations of the spectra of I_α rich *Cladophera glomerata*, before and after acid hydrolysis, which resulted in a I_β-rich cellulose. Adapted from Ref. [27].

would have liked. These exchange experiments were also applied to an acid-hydrolyzed sample of cotton linters. The inferior spectral resolution and the shorter relaxation times associated with the higher plant celluloses, relative to many algal celluloses, made this experiment much more difficult as the near-neighbor spectra had poorer sensitivity and resolution. What results there were pointed to a crystal structure for the higher plant celluloses with a dominance of the I_β form and a questioning if there was, in fact, any I_α component.

With the completion of this study by VanderHart and Atalla, the notion that algal celluloses were semicrystalline materials possessing two crystalline components was firmly established. On the other hand, the proposal that higher plant celluloses contained similar heterogeneity of crystalline form, rather than not-so-perfectly formed I_β crystallites or a more complex unit cell, was less well-established. The published spectrum of *Halocynthia* (tunicate) cellulose [29] demonstrated that the downfield shoulder of the C4 resonance, typical of higher plant celluloses, is nearly absent; hence, a perfectly formed I_β crystallite is not expected to exhibit such a feature. This is illustrated in the middle spectrum of Fig. 6, which is a 25-MHz ^{13}C spectrum we have taken of the more crystalline regions of a moist tunicate cellulose (the other spectra displayed in Fig. 6 will be discussed later). This spectrum represents the more rigid carbons, i.e., those carbons which have survived a 75-s interval of longitudinal carbon relaxation, in the pulse sequence of Torchia [33]. Note the absence of any significant downfield shoulder at about 90.3 ppm. Note also the near-baseline resolution of the doublet at C1. There is no hint here of any contribution from the I_α form. Therefore, either the higher plant celluloses are composites containing I_α and I_β crystallites, or the crystallite packing of the I_β structure in the higher plants is imperfect, giving rise to an anomalous downfield shoulder at C4.

With the above clarifications concerning the nature of native celluloses in mind, it was possible to classify particularly the algal and bacterial celluloses with respect to the relative amounts of the I_α and I_β forms occurring in the celluloses produced by particular species. It emerged in these early studies that the celluloses from more primitive organisms such as *V. ventricosa* and *A. xylinum* are predominantly of the

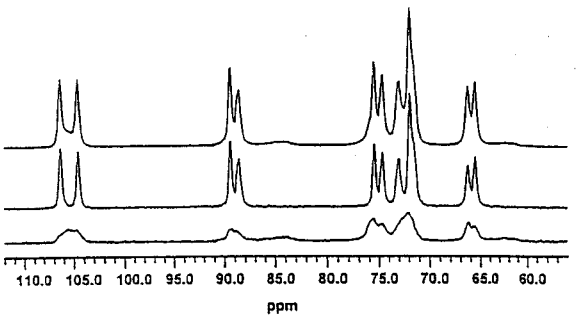

Fig. 6. The 24-MHz ^{13}C CP/MAS spectra of moist *Halocynthia* cellulose. Top: normal CP/MAS spectrum of the whole material; (vertical factor = 0.7). Middle: after 75 s of longitudinal carbon relaxation where signals from the slower relaxing, least mobile, carbons in the interior of the crystallite survive best; (vertical factor = 1.0). Bottom: difference spectrum (top–middle). Scaling factors are chosen so that the lineshape influence of the sharpest features is minimized in the difference spectrum. Note the lack of I_α contributions in the middle spectrum. Also, the fact that in the 88–91 ppm region, about 25% of the upper intensity appears in the difference spectrum means that resonances from the crystalline region have varied linewidths where the components with the broader features relax more quickly. Finally, note that the components of the doublet at C4 in the middle spectrum have unequal linewidths but the same integrated intensities.

I_α form. We can also say with confidence that the celluloses from higher plants such as cotton and ramie are predominantly of the I_β form although the perfection of the lattice is still an open question when the diameter of the crystallites is only 3–5 nm. 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 IR spectra in the OH stretching region. In relation to that classification, the NMR spectra 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_β form.

With the recognition that the ordered regions of all plant celluloses are either imperfect I_β crystallites or composites of the I_α and I_β allomorphs, 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 many of the celluloses produced by different organisms were composites of two crystalline forms that were blended

in different proportions. As noted earlier, the finding that celluloses from different biological sources represent unique blends of the two different forms also resolved the paradox posed early in this century by Cross and Bevan [1] when they contrasted the observation of individuality in biological systems with the invariance of crystalline forms.

4. Further studies of the two forms: I_α and I_β

With the wide acceptance of the proposal of the two forms (I_α and I_β) came the challenge of understanding the differences between them and their relationship with 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 in order to place the remaining questions in perspective.

4.1. Raman and IR spectra

The categorization of native celluloses into the I_A and I_B group by Howsmon and Sisson [25] and Blackwell and Marchessault [26] on the basis of the appearance of the OH stretching region of their IR 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 [34] and in a comprehensive study of the IR spectra of a number of celluloses of the two forms [35]. 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 [34]. The key observations were in comparisons of the spectra of ramie fibers with those of cellulose from the walls of *V. ventricosa*, an alga; the ramie cellulose is predominantly of the I_β form, while the *Valonia* is predominantly of the I_α forms. In most regions of the spectra, those of ramie and *Valonia* were the same except for a limited broadening in the bands for ramie because of the smaller fibril diameters. The similarity of the spectra, particularly in the conformation-sensitive region, indicates that the two forms of cellulose have the same skeletal conformation. In the OH stretching region in contrast, the differences were significant, suggesting that the two forms have different hydrogen bonding patterns. These conclusions were further confirmed by more recent observations in which Raman spectra of celluloses from *Valonia* and from *Halocynthia* were compared. *Halocynthia*, a tunicate cellulose, appears to be predominantly of the I_β form. This comparison was sought because *Valonia* and *Halocynthia* have fibrils of approximately equal (20 nm) lateral dimensions so that the Raman spectra are of equal resolution. Here again, the primary difference between the two spectra was in the OH stretching region.

The IR spectral studies of the I_α and I_β forms were carried out by Sugiyama et al. [35] on a number of different native celluloses of both forms. Furthermore, they included examination of a number of I_α-rich celluloses that were converted to the I_β form through the annealing process first reported by Yamamoto et al. [36]. In order to complement the IR spectra, Sugiyama et al. [37] recorded electron diffraction patterns for the samples, which allowed classification of the celluloses, through comparison with the diffraction patterns acquired in an earlier electron diffractometric study.

The key finding emerging from the examination of the IR 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_β form through annealing. The authors concurred with the interpretation of the differences between the two forms suggested earlier on the basis of the Raman spectral studies, and concluded that the I_α-to-I_β 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 IR spectral studies by Atalla et al. [38] are particularly interesting because they included the spectra of both *Valonia* and *Halocynthia*, the Raman spectra of which have also been investigated at high resolution. In summary then, the Raman and IR spectral studies undertaken after discovery of the composite nature of native celluloses point to the conclusion that the primary difference between the two forms is in the pattern of hydrogen bonding the difference between chain conformations is neglegible.

4.2. Electron microscopy and electron diffractometry

Electron microscopy and electron diffractometry made two important contributions to advancing understanding of the structures of native celluloses. The first was the acquisition of lattice images of some highly crystalline celluloses, the second was in developing a crystallographic perspective on the nature of the difference between the two forms of cellulose I. In an important set of investigations by Sugiyama et al. [39–41], 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 [39,40]. Later, the techniques were refined to allow the acquisition of lattice images of cross-sections of microfibrils [41]. The significance of these observations was that it was 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 that 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 [42,43]; the findings of Sugiyama et al. were the first direct evidence that the approximately 20 cm × 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 the organization of the heavy atoms in the molecular chains and would be insensitive to any non-uniformity in the hydrogen bonding patterns within the interior of the 20 nm × 20 nm fibrils.

The electron diffractometric studies, on the other hand, have led to conclusions that the two forms, I_α and I_β, represent two crystalline phases with different crystal habits [37]. The key conclusion drawn from the electron diffractometric data was that the I_α form represents 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 $P2_1$.

The conclusion, that the I_β form has two chains per unit cell, may well be in conflict with the observations in the Raman spectra of *Halocynthia*. The key issue is that when crystal structures possess more than one molecule per unit cell and all molecules have the same vibrational frequencies, then the vibrational modes of the unit cell become degenerate. Under these circumstances, weak additional couplings between equivalent modes in the different molecules will result in splittings of the bands associated with these key vibrational modes. The type of splitting that is relevant in the case of cellulose is that described as correlation field splitting [44]. This effect arises because, as a result of the additional weak 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. For example, such correlation field effects result in doublets with a splitting of 10–15 cm^{-1} in some modes of crystalline orthorhombic polyethylene which has two chains per unit cell. Since no evidence of such splittings occurs in the Raman spectrum of *Halocynthia*, it is necessary to question whether the I_β form truly has more than one molecule per unit cell. It is recognized that, while the presence of correlation field splittings is direct evidence for two chains per unit cell, the absence of splittings cannot be regarded as conclusive proof, without augmenting arguments, that a two-chain unit cell should be rejected. However, since there is no reason why correlation field splittings would not occur in a cellulose unit cell with two chains, this lack of splittings raises serious questions about the two-chain unit cell for the I_β form. It could be argued that the lack of splittings

implies that the two molecules in a monoclinic unit cell are non-equivalent with modes at different frequencies such that additional weak couplings would be ineffective in causing splittings. As a counter to such an argument, it would be noted that since the skeletal bands in the *Halocynthia* spectrum are essentially identical to those in the *Valonia* spectrum (whose I_α contributions are unsplit), it would be expected that the two chains in the I_β unit cell would be at least "near-degenerate" and would still show splittings as a result of weak additional couplings. This similarity in the skeletal bands was also reported in the IR spectra observed by Sugiyama et al. cited earlier. Thus, the vibrational spectra, both Raman and IR, strongly suggest that both the I_α and I_β forms have only one molecule per unit cell. This conclusion, of course, raises the question, which is still outstanding, as to why the selection rules used in the analysis of the crystallographic data lead to conclusions that differ from those that are based on the selection rules applicable to spectral observations.

4.3. Solid state ^{13}C NMR spectra

The group at the Kyoto University Institute for Chemical Research carried out important studies that were complementary to those undertaken by VanderHart and Atalla [22,27], and Atalla and VanderHart [24]. More recently, a number of other groups have made valuable contributions. Since questions concerning the nature of the I_α and I_β forms remain outstanding, it is useful to provide 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 [45]. In a careful analysis of the chemical shifts of the C1, C4, and C6 carbons in the (CP/MAS) spectra of monosaccharides and disaccharides for which crystallographic structures were available, Horii et al. 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 C6, they demonstrated a correlation between the chemical shift of the C6 resonance and the value of the dihedral angle χ defining the orientation of the OH group at C6 relative to the C4–C5 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 C6 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 C1 and C4 are correlated with the dihedral angles about the bonds in the glycosidic linkage. In particular, there was a correlation between the shift of C1 and the dihedral angle ϕ about the glycosidic C1–O bond and a correlation between the shift of C4 and the dihedral angle ψ about the O–C4 bond. As the spectra published in the earliest studies did not have sufficient resolution to reveal the splittings of the resonances of C1 and C4, the possibility of occurrence of non-equivalent glycosidic linkages was not addressed at that time.

In addition to the analysis of the correlation between the chemical shifts and the dihedral angles, the Kyoto group also undertook analysis of the lineshapes of the different resonances, particularly that of the C4 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 only a weak justification in terms of our understanding of linewidths in the CP/MAS spectra of organic materials like cellulose [46]. However, deconvolution into Lorentzians can be a useful tool for approximating the amounts of various constituents and for generating a relative ranking of composition. Thus, this method remains in use.

In the early studies by Horii et al., the entire upfield wing of the C4 resonance was attributed to molecules in noncrystalline domains. On this basis, the lineshape analysis of the C4 resonance of different native celluloses did not seem consistent with the model proposed by VanderHart and Atalla [22] 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 C4 in the spectra of higher plant celluloses was attributed to the surface molecules of crystalline domains, Yamamoto and Horii [47] indicated that their results confirm the proposal of VanderHart and Atalla. It is to be noted that in their early reports in this area, Horii 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 more recent studies, they have adopted the I_α and I_β designations that are designed to avoid the confusion with the categories first introduced by Howsmon and Sisson 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 was directed at the effects of annealing, first in saturated steam [31] and later in aqueous alkaline solutions (0.1 N NaOH) selected to avoid hydrolytic decomposition of the cellulose [48,49]. 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 *V. macrophysa*. These results 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. [29], was also annealed in an aqueous alkaline solution at 260°C; it showed little change as a result of the annealing [49].

In 1990, Newman and Hemingson [50] began to combine some additional methods of processing the ^{13}C NMR spectral data with those that had been used previously such as the monitoring of 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. It must be noted, however, that the application of these methods has been complemented in the work of Newman and Hemingson 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 procedures were introduced into studies of cellulose in the studies by Newman and Hemingson (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 [50]. The second, and perhaps more helpful, procedure was the application of a method [51] that relies on morphological variation in rotating-frame proton relaxation times in order to separate signal contributions from domains where molecules have 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.

The resolution enhancement they used [50] involved convoluting the free induction decay with a function having the form:

$$f(t) = \exp\{at^2 - bt^3\}.$$

This function differs from that used by other investigators in earlier studies in that both time 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 cellulose. The parameters a and b were selected to enhance resolution "without decreasing the signal-to-noise ratio to unacceptable levels".

We have reservations about this method because $f(t)$ has a maximum at a time displaced from $t = 0$, which depends on the choice of a and b. When one convolutes a function with $f(t)$ and does a Fourier transform to the frequency domain, one artificially

introduces a periodicity into the spectrum, which periodicity depends on the displacement of this maximum from $t = 0$. So one must be exceedingly careful that one does not introduce periodicities into the spectrum which are not there. For example, in the case of cellulose, where the separation between multiplet components at both C4 and C1 is about 0.8 ppm, one could choose a and b such that features separated by 0.8 ppm were seemingly enhanced. But this kind of resolution enhancement is biased and one should not expect to be able to analyze, with high accuracy, such resolution-enhanced lineshapes for component intensities. Newman [52,55], Newman and Hemmingson [56] and Newman et al. [53,54] have used resolution-enhanced spectra in attempts to quantify the relative amounts of the I_α and I_β forms in a number of native celluloses. These results might be useful for identifying trends but cannot be regarded as indicative of the true values.

The other method based on variability in rates of proton spin relaxation [56] was used to generate spectra that were described as "proton spin-relaxation-edited" or PSRE spectra. The 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 (C4 in crystal interior cellulose) and 80 ppm (C4 in noncrystalline cellulose) without allowing any signal to become inverted." In a system like cellulose where the lateral size of crystallites is small and where disorder (which is usually associated with higher mobility) exists both at crystallite surfaces and in other three-dimensional regions of disorder, the PSRE spectra are not expected to be very effective in isolating the signals from the crystallite interior from those in the disordered regions, particularly those at the crystallite surfaces. A better separation is expected between signals from three-dimensional regions of disorder and signals from crystallites. A principal reason for this difficulty in separating signals is that polarization transport, called spin diffusion [57], is occurring while the relaxation is going on; hence, when regions with varying relaxation rates are separated by short distances, the polarizations of all regions are influenced by faster intrinsic relaxation in one of the regions.

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) [58]. 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 main spectral signatures proposed by VanderHart and Atalla of the I_α and I_β allomorphs were used as a bias in generating "principal-component subspectra". In addition, they defined the "crystallinity index" in terms of the fractional C4 integral over the interval from 86 to 92 ppm. After extraction of the main principal-component subspectra, using five different calibrating samples, they went on to assay the I_α, I_β, and crystallinity levels in several cellulose samples including several lignocellulosic samples from higher plants. On the basis of this analysis, the I_α/I_β ratio in cotton linters and cotton is, respectively, 0.55 and 0.61 and in the lignocellosics is 0.63–0.73. In other words, 35%–42% of the crystalline material in these higher plant celluloses is the I_α form.

Subsequently, this same group from STFI analyzed CP/MAS spectra of *Halocynthia* [59] using only Lorentzian contributions with assignments consistent with the crystal-composite model. On the basis of this analysis, they reported that *Halocynthia* has an I_α/I_β ratio of about 0.16 as well as possible "paracrystalline" regions. The latter regions have a spectral signature which generally shows up in the region of the "crystal-carbon" resonances at C4; however, there is little to no definition of the usual multiplet structure for this component.

More recently, the STFI group analyzed CP/MAS spectra of *Valonia*, *Halocynthia*, and *Cladophora* celluloses along with spectra of cotton and wood. This time, they used both Gaussian and Lorentzian lineshape components, each with variable width, amplitude, and central frequency. They used eight and

six components, respectively, to fit the C1 and the C4 spectral regions of cotton cellulose. The large number of variables that go into the fitting of these lines at least raises the question of the uniqueness or the dependability of the fits. They partially addressed this concern using additional constraints in the fit process, i.e., they used the general features of the proposed I_α and I_β spectra and further demanded that the lineshape analyses at C4 and C1 be consistent with each other in terms of the I_α and I_β contents. Interestingly, on the basis of this analysis, the I_α/I_β ratio for cotton cellulose turned out to be 0.15, and not the 0.61 value derived from the PCA mentioned earlier. Such wide variation in results, however, raises questions concerning the degree of confidence with which the results can be viewed.

Some additional observations are in order with respect to the methods of analysis described above. The extraction of the I_α/I_β ratio, e.g., is not nearly as straightforward a task as might appear at first, especially in the case of the higher plant celluloses where the spectral resolution of the multiplet components is very poor. We note again that the use of symmetric, Lorentzian or Gaussian lineshapes has little justification. Also, the small lateral crystallite dimensions, common to many celluloses, have additional implications. One cannot expect an instantaneous transition from disorder to full order at the crystallite boundary. There will be a finite distance over which the perfection of order increases and it is expected that the corresponding sharpness of resonances will increase for carbons in residues deeper into the crystallite. One sees evidence of this general theme in the spectra of many polymers, notably, in our own experience, polyethylene and polypropylene. Evidence we refer to manifests itself in lineshape changes both with partial rotating-frame proton relaxation and with partial longitudinal carbon relaxation. In the present context, one would like to distinguish material near crystallite surfaces from so-called "paracrystalline" material, proposed by the authors from STFI [59,60]. While these authors did not explicitly suggest a morphological location for the paracrystalline material, there was no direct suggestion that the paracrystalline material lay near the crystallite surfaces. Compared to the evidence from carbon relaxation, evidence from the partial proton rotating frame relaxation is much more convincing that this spectral broadness arises from residues near crystallite surfaces and not from isolated paracrystalline regions. Qualitatively, the amount of reported [60] paracrystalline material varies inversely with lateral crystallite dimension; in fact, for cotton and wood celluloses, respectively, the paracrystalline component has about half and two-thirds the intensity of the crystalline components. This trend is qualitatively consistent with our suggestion involving near-surface chains.

On the basis of what we have noted, one should never expect that a multiplet component in the spectrum of a typical cellulose crystallite of relatively small size could be described by a single linewidth, even if there were justification for a single linewidth in a large, perfect crystallite. As an illustration of this point, we review in Fig. 6 the spectra of moist *Halocynthia*. The top spectrum is the normal 25 MHz CP/MAS spectrum. The middle spectrum is strongly biased towards the most ordered carbons because a period of 75 s of carbon longitudinal relaxation has preceded the acquisition of this spectrum. During that time, the more mobile, less-ordered carbons have preferentially relaxed. In fact, the middle spectrum has a total integral, normalized to the same number of scans, of 0.43 times the upper spectrum. In Fig. 6, the top spectrum has been rescaled so that the difference spectrum, the lower spectrum, displays minimal sharp features. Note, e.g., that in this difference spectrum at C4 (88–92 ppm), one can barely make out that I_β doublet structure so prominent in the middle spectrum. Other regions of the difference spectrum have similar implications, although, in most other resonance regions, contributions from more disordered regions add more confusion to the interpretation of the lineshapes. In any case, it is very clear that even in the spectrum of *Halocynthia* cellulose, whose lateral crystallite dimensions (20 nm) are large for celluloses, the linewidths of all of the carbons contributing, e.g., to the C4 crystalline region (88–92 ppm), are by no means uniform.

A second point evident in Fig. 6 is that, although the top spectrum is nearly identical to that analyzed by the STFI group for which an I_α/I_β ratio of 0.15 was reported, the middle spectrum gives no evidence whatsoever, of any I_α component in the interior of the *Halocynthia* crystallites. This deduction is most

evident in the C1 region of the middle spectrum. Perhaps, this means that if the I_α crystallites exist, they are smaller and have shorter carbon relaxation times than their I_β counterparts. In view of the fact that in the algal celluloses, the carbons in the I_α and I_β crystallites have comparable carbon relaxation times, the foregoing explanation seems less likely. It would appear more likely that it is a reflection of the inadequacy of the assumption that the CP/MAS lineshape of C1 can be decomposed into Lorentzian components and each component assigned meaningful morphological identities.

Another point regarding the middle spectrum of Fig. 6 is that the linewidths of the two multiplet components at C4 are in a ratio of approximately 10:7; hence, although their integrals are about equal, their peak intensities are not. It is a general finding that the I_β doublet components at C4 have different linewidths (see Figs. 4 and 5). The reason for the two linewidths at C4 is not obvious; one possibility is that the conformational variation in the lattice is wider for one C4 site than the other. In this connection, we also recognize that the two C1 resonances have negligible linewidth difference so they offer no support for site-based variation in the glycosidic linkage. We are presently considering whether this observation has any relevance to the question of whether the twofold screw axis coincides with the polymer chain. Nothing in the NMR measurements performed allows for a direct assignment of the two sites to either consecutive residues on the same chain or to separate chains.

A final consideration concerning lineshape analyses relates to the higher plant celluloses where we have already pointed to strongly varied I_α/I_β ratios [60,61] for cotton cellulose using different lineshape analysis methods. If one looks at Fig. 9 of Ref. [27] where one can, for acid-hydrolyzed cotton linters, follow the lineshape changes for the 50 MHz C4 and C1 resonances as a function of carbon longitudinal relaxation time (1 ms, 20 s and 200 s), one will see almost contradictory changes at C1 and C4. At C1, the central component (I_α by assignment) steadily loses intensity while the relative intensities of the multiplet components at C4 remain fixed, although the linewidth of each surviving multiplet component gets slightly narrower at the longer delay times. Thus, these observations seem contradictory in that the relative I_α lineshape contribution at C1 diminishes, but at C4 remains fixed as the delays increase. Since we expect and usually find that the relaxation times of C4 and C1 are comparable when crystallite size is larger, we do not expect that differential relaxation times is the correct explanation for the differential intensity change. What seems more reasonable to us is that the lineshape contributions from less-ordered cellulose at C1 strongly overlap with the I_α multiplet component and these C1 carbons relax preferentially, thereby causing the central region of C1 to decrease. It is particularly without foundation to suggest that the lineshape contribution from the disordered C1 carbons is Lorentzian or Gaussian. Hence, for the higher plant celluloses, it is very dangerous to use the normal C1 lineshape to analyze for the I_α/I_β ratio. Moreover, on the basis of the C1 lineshape at 200 s, coupled with the stability of the C4 lineshape, we would expect that the I_α/I_β ratio in this acid-hydrolyzed cotton linters is less than 0.25. It was suggested earlier that in the C4 region of the spectra of higher plant celluloses, the downfield multiplet component may be evidence for an imperfect crystallite structure (owing to small lateral size), rather than an I_α component. The stability of the C4 lineshape with carbon relaxation time for the hydrolyzed cotton linters suggests that the downfield wing does not arise preferentially from cellulose molecules near the crystallite surface. Hence, it is more likely that there indeed is some I_α present and that the relative intensity ratio of the upfield to the downfield shoulders, which generally looks to be about 2:1, is really more like 4:1 or higher. A possible explanation for this is that the downfield multiplet at C4 has a narrower intrinsic component linewidth than the other multiplet components at C4.

Though the detailed methods developed by Larsson et al. for analyzing cellulose fractions cannot be regarded as indicating absolute values, they represent an important step forward in recognizing the different degrees of order along with some correlation to the associated levels of molecular mobility. Furthermore, it highlights the fact that the celluloses that 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.

5. Conclusions

The analysis by VanderHart and Atalla that led to the recognition of the composite nature of native celluloses has become the basis of much of the research seeking to characterize the nature of native celluloses. The most fundamental change that has arisen as a result of the findings is that native celluloses are no longer viewed as variations on a single structure differing only in the degree of their crystallinity. Rather, it is now recognized that the structures are species-specific blends of the two forms of native cellulose, I_α and I_β. The key variables characterizing a particular native cellulose appear to be the lateral dimensions of the microfibrils and the relative amounts of the two forms. These relative amounts are indicated by distinctive features in the solid state ^{13}C NMR spectra, particularly the shoulders associated with the resonance of C4 and, to a lesser extent, features at C6 and C1.

Other investigations of the two forms based on Raman and IR spectroscopy indicate that the two forms of native cellulose have molecular chains that have the same skeletal conformations but different hydrogen bonding patterns. In spite of these similar conformations, other studies, based on electron diffraction of the more highly crystalline native celluloses, conclude that the lattices are different. A triclinic lattice with one chain per unit cell is assigned to the I_α form and a monoclinic lattice with two chains per unit cell is assigned to the I_β form. The latter assignment is challenged by the lack of any observed correlation field splittings in the vibrational spectra of this form. One expects such splittings in unit cells with two equivalent chains but not in unit cells with only a single chain; thus, the extensive observed similarity of the vibrational spectra for the I_α and I_β forms supports the idea that these splittings are absent for the I_β form. In any case, it remains a challenge to establish the I_β crystal structure and to reconcile the different conclusions, based on different types of measurements, about the I_β form.

The use of *Halocynthia* cellulose as a proper model for the I_β form has clarified certain issues about the true CP/MAS spectrum belonging to the I_β form. In particular, the number of identifiable inequivalent anhydroglucose units in the I_β unit cell is, from the CP/MAS lineshape, limited to 2. At the same time, for the higher plant celluloses, the typical CP/MAS lineshape at C4 arising from the crystalline regions includes a downfield wing which is not a feature of the I_β spectrum but could belong to the I_α spectrum. However, given the small size of crystallites in higher plant celluloses, one cannot dismiss the possibility that the more complex shape of the C4 resonance indicates imperfections in crystalline order. Yet, it is more probable that the higher plant celluloses are composites of I_α and I_β, just like the algal and bacterial celluloses. As to the relative amounts of the I_α and I_β constituents, it is likely that the I_α/I_β ratio in the higher plant celluloses is less than 0.25 even though the apparent ratio of intensity of the upfield and downfield shoulders at C4 can suggest a I_α/I_β ratio closer to 0.5. The proper rationalization for the shape of the C4 resonance remains an open issue at this time. Certainly, it is true that a determination of I_α/I_β ratios for higher plant celluloses via CP/MAS lineshape deconvolution of the C4 and C1 resonances is not readily accomplished. It would be helpful if other spectral signatures of the I_α form could be identified in the CP/MAS spectra of the higher plants.

Further studies of the solid state ^{13}C NMR lineshapes have been helpful for validating methods of converting I_α to I_β and for comparing crystallinity and I_α/I_β ratios in different native celluloses, especially those celluloses characterized by larger crystallites. It is important, however, that the particular mathematical approach to lineshape analysis be understood, especially if one wishes to carry out quantitative comparisons of results from different laboratories.

A major challenge for future work is development of methods for more detailed characterization of the unique blends of the two forms I_α and I_β in a manner that is useful for understanding the biological implications of structural differences. Future advances in understanding the biogenesis of cellulose will require more effective and more comprehensive characterization of the secondary and tertiary struc-

tures of native celluloses. A more complete characterization of the variability of native celluloses is also essential to advancing our understanding of the action of microorganism-borne cellulases that chemically disassemble cellulose. These microorganisms and their cellulases play a key role in recycling plant biomass into the biosphere.

References

[1] C.F. Cross, E.J. Bevan, Researches on Cellulose (III), Longmans, Green and Co., London, 1912.
[2] R.H. Atalla, in: B.M. Pinto (Ed.), Comprehensive Natural Products Chemistry, Vol. 3, Carbohydrates, Elsevier, New York, 1999, p. 529.
[3] K.H. Gardner, J. Blackwell, Biopolymers 13 (1974) 1975.
[4] A. Sarko, R. Muggli, Macromolecules 7 (1974) 486.
[5] K.H. Meyer, L. Misch, Helv. Chim. Acta 20 (1937) 232.
[6] K.H. Meyer, H.Z. Mark, Phys. Chem. B 2 (1929) 115.
[7] P.H. Hermans, Physics and Chemistry of Cellulose Fibers, Elsevier, New York, 1949.
[8] A. French, in: R.H. Atalla (Ed.), The Structures of Celluloses, ACS Symposium Series 340, American Chemical Society, Washington, DC, 1987, p. 15.
[9] J.J. Hebert, L.L. Muller, J. Appl. Polym. Sci. 18 (1974) 3373.
[10] R.H. Atalla, Appl. Polym. Symp. 28 (1976) 659.
[11] R.H. Atalla, Adv. Chem. Ser. 181 (1979) 55.
[12] K. Hieta, S. Kuga, M. Usuda, Biopolymers 23 (1984) 1807.
[13] H. Chanzy, B. Henrissat, FEBS Lett. 184 (1985) 285.
[14] R.H. Atalla, J.C. Gast, D.W. Sindorf, V.J. Bartuska, G.E. Maciel, J. Am. Chem. Soc. 102 (1980) 3249.
[15] W.L. Earl, D.L. VanderHart, J. Am. Chem. Soc. 102 (1980) 3251.
[16] W.L. Earl, D.L. VanderHart, Macromolecules 14 (1981) 570.
[17] D.L. VanderHart, Deductions about the morphology of wet and wet beaten cellulose from solid state ^{13}C NMR, NBSIR 82-2534.
[18] C.A. Fyfe, R.L. Dudley, P.J. Stephenson, Y. Deslandes, G.K. Hamer, R.H. Marchessault, J. Am. Chem. Soc. 105 (1983) 2469.
[19] F. Horii, A. Hirai, R. Kitamaru, Polym. Bull. 8 (1982) 163.
[20] G.E. Maciel, W.L. Kolodziejski, M.S. Bertran, B.R. Dale, Macromolecules 15 (1982) 686.
[21] J.C. Gast, R.H. Atalla, R.D. McKelvey, Carbohydr. Res. 84 (1980) 137.
[22] D.L. VanderHart, R.H. Atalla, Macromolecules 17 (1984) 1465.
[23] R.H. Atalla, Proceedings of the International Symposium on Wood and Pulping Chemistry, SPCI Rep. No. 38, Vol. 1, Stockholm, 1981, p. 57.
[24] R.H. Atalla, D.L. VanderHart, Science 223 (1984) 283.
[25] J.A. Howsmon, W.A. Sisson, in: E. Ott, H.M. Spurlin, M.W. Graffline (Eds.), Cellulose and Cellulose Derivatives, Part I, Wiley-Interscience, New York, 1954, p. 237.
[26] J. Blackwell, R.H. Marchessault, in: N.M. Bikales, L. Segal (Eds.), Cellulose and Cellulose Derivatives, Part IV, Wiley-Interscience, New York, 1971, p. 1.
[27] D.L. VanderHart, R.H. Atalla, in: R.H. Atalla (Ed.), The Structures of Celluloses, ACS Symposium Series 340, American Chemical Society, Washington, DC, 1987, p. 88.
[28] D.L. VanderHart, W.L. Earl, A.N. Garroway, J. Magn. Reson. 44 (1981) 261.
[29] P.S. Belton, S.F. Tanner, N. Cartier, H. Chanzy, Macromolecules 22 (1989) 1615.
[30] R.H. Atalla, R.E. Whitmore, D.L. VanderHart, Biopolymers 24 (1985) 421.
[31] F. Horii, H. Yamamoto, R. Kitamaru, M. Tanahashi, T. Higuchi, Macromolecules 20 (1987) 2946.
[32] D.L. VanderHart, J. Magn. Reson. 72 (1987) 13.
[33] D.A. Torchia, J. Magn. Reson. 30 (1978) 613.
[34] J.H. Wiley, R.H. Atalla, in: R.H. Atalla (Ed.), The Structures of Celluloses, ACS Symposium Series 340, American Chemical Society, Washington, DC, 1987, p. 151.
[35] J. Sugiyama, J. Presson, H. Chanzy, Macromolecules 24 (1990) 4168.
[36] H. Yamamoto, F. Horii, H. Odani, Macromolecules 22 (1989) 4130.
[37] J. Sugiyama, T. Okano, H. Yamamoto, F. Horii, Macromolecules 23 (1990) 3196.
[38] R.H. Atalla, J. Hackney, U.P. Agarwal, A. Isogai, to be published.
[39] J. Sugiyama, H. Harada, Y. Fujiyoshi, N. Uyeda, Mokuzai Gakkaishi 30 (1984) 98.
[40] J. Sugiyama, H. Harada, Y. Fujiyoshi, N. Uyeda, Mokuzai Gakkaishi 31 (1985) 61.
[41] J. Sugiyama, H. Harada, Y. Fujiyoshi, N. Uyeda, Planta 166 (1985) 161.
[42] R.D. Preston, The Physical Biology of Plant Cell Walls, Chapman&Hall, London, 1974.
[43] A. Frey-Wyssling, The Plant Cell Wall, Gebruder Borntrager, Berlin, 1976.
[44] D.I. Bower, W.F. Maddams, The Vibrational Spectroscopy of Polymers, Cambridge Univ. Press, Cambridge, UK, 1989.
[45] F. Horri, A. Hirai, R. Kitamaru, in: R.H. Atalla (Ed.), The Structures of Celluloses, ACS Symposium Series 340, American Chemical Society, Washington, DC, 1987, p. 119.
[46] D.L. VanderHart, G.C. Campbell, J. Magn. Reson. 134 (1998) 88.
[47] H. Yamamoto, F. Horii, Macromolecules 26 (1993) 1313.
[48] H. Yamamoto, F. Horii, H. Odani, Macromolecules 22 (1989) 4130.
[49] H. Yamamoto, F. Horii, Macromolecules 26 (1993) 1313.
[50] R.H. Newman, J.A. Hemmingson, Holzforschung 44 (1990) 351.
[51] D.L. VanderHart, E. Perez, Macromolecules 19 (1986) 1902.
[52] R.H. Newman, J. Wood Chem. Technol. 14 (1994) 451.
[53] R.H. Newman, M.A. Ha, L.D. Melton, J. Agric. Food Chem. 42 (1994) 1402.
[54] R.H. Newman, L.M. Davies, P.J. Harris, Plant Physiol. 111 (1996) 474.
[55] R.H. Newman, Cellulose 4 (1997) 269.
[56] R.H. Newman, J.A. Hemmingson, Cellulose 2 (1995) 95.

[57] A. Abragam, The Principles of Nuclear Magnetism, Chap. V, Oxford Univ. Press, London, 1961.
[58] H. Lenholm, T. Larsson, T. Iversen, Carbohydr. Res. 261 (1994) 119.
[59] T. Larsson, U. Westermark, T. Iversen, Carbohydr. Res. 278 (1995) 33.
[60] T. Larsson, K. Wickholm, T. Iversen, Carbohydr. Res. 302 (1997) 19.
[61] H. Lennholm, T. Larsson, T. Iversen, Carbohydr. Res. 261 (1994) 119.