Glasgow University
Enlighten version

Philosophical Transactions of the Royal Society B
Accepted for publication June 2017

Review

Structure of native cellulose microfibrils, the starting point for nanocellulose manufacture

Michael C. Jarvis

College of Science and Engineering, University of Glasgow, Glasgow G12 8QQ, Scotland, UK

There is an emerging consensus that higher plants synthesise cellulose microfibrils that initially comprise 18 chains. However the mean number of chains per microfibril *in situ* is usually greater than 18, sometimes much greater. Microfibrils from woody tissues of conifers, grasses and dicotyledonous plants, and from organs like cotton hairs, all differ in detailed structure and mean diameter. Diameters increase further when aggregated microfibrils are isolated. Because surface chains differ, the tensile properties of the cellulose may be augmented by increasing microfibril diameter.

Association of microfibrils with anionic polysaccharides in primary cell walls and mucilages leads to *in vivo* mechanisms of disaggregation that may be relevant to the preparation of nanofibrillar cellulose products. For the preparation of nanocrystalline celluloses, the key issue is the nature and axial spacing of disordered domains at which axial scission can be initiated. These disordered domains do not, as has often been suggested, take the form of large blocks occupying much of the length of the microfibril. They are more likely to be located at chain ends or at places where the microfibril has been mechanically damaged, but their structure and the reasons for their sensitivity to acid hydrolysis need better characterisation.

1. Introduction

Nanocelluloses are made from natural cellulosic materials by processes that, in contrast to rayon manufacture, retain much of the native structure of the cellulose microfibrils [1]. Native cellulose fibres normally form part of the quasi-solid, cohesive structure of the plant cell wall [2], and must be dispersed in a solvent to allow their conversion to nanocellulose products. The native structure of cellulose is therefore relevant to the properties of nanocelluloses and to the technology required for their isolation.

A wide variety of spectroscopic, scattering and imaging methods have been used to investigate the structure of cellulose microfibrils. Broadly, NMR spectroscopy is most informative about chain conformation [3]; crystallography, about chain packing [4]; FTIR spectroscopy, about hydrogen bonding [5]; atomic force microscopy, about microfibril dimensions [6]; small-angle scattering, about the aggregation of microfibrils [7]. All these methods have limitations. The drawbacks of each method are understood within its own, specialised physical science community, but for many years there was little sharing of ideas between these communities; divergent evidence from crystallography [8] and NMR spectroscopy [9] was particularly confusing. The debate on cellulose structure has therefore been fragmented and difficult of access for biologists and materials technologists. It is only recently that critical synthesis of evidence from multiple techniques [3, 5, 10, 11] has become more common.

Cellulose is a highly insoluble polymer. Its chains adopt a flat-ribbon, two-fold helical conformation and assemble into sheets through edge-to-edge hydrogen bonding [4] (Figure 1). These sheets then stack into a lattice whose crystallinity [12] increases with size [9, 13]. Cellulose and chitin are unique in being synthesised in the solid state, as microfibrils with lateral dimensions of nanometres and axial dimensions of micrometres [2].

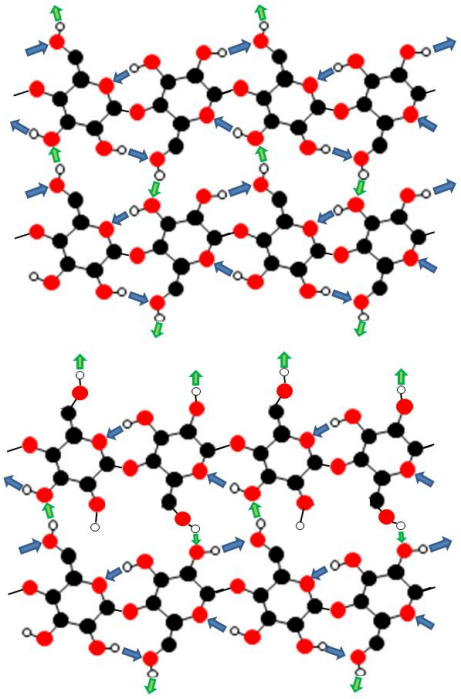

Figure 1. Hydrogen bonding schemes for cellulose. Top: a two-chain segment of 'crystalline' cellulose with all C-6 in the *tg* conformation, permitting a line of intramolecular hydrogen bonds (blue arrows) that run along each side of each cellulose chain (shaded pale blue). Intermolecular hydrogen bonds are shown as green arrows. Bottom: two-chain segment of 'surface' cellulose with the top edge of the upper chain having C-6 in the *gt* conformation, so that the line of intramolecular hydrogen bonds is interrupted and there is an increased number of transversely oriented, intermolecular hydrogen bonds.

The synthetic complex that constructs a cellulose microfibril at the surface of a plant cell is one of the largest protein complexes in

Nature, probably with as many catalytic subunits as there are chains in the emergent microfibril [14, 15]. It follows that our understanding of the structure and functioning of the biosynthetic complex goes hand in hand with our understanding of microfibril structure. In the past decade there have been rapid advances in our knowledge of both cellulose biosynthesis and structure. These advances form the basis of this review.

Structural questions that will be addressed include the diameter of microfibrils, the nature of the disorder within their structure and its relation to tensile properties; the extent to which microfibrils are aggregated into larger bundles in living plants, with potential for this bundling to be carried over into nanocellulose products; and the presence, along the length of the microfibril, of discontinuities that may provide starting points for hydrolysing microfibrils into nanocrystalline celluloses.

2. Different organisms make cellulose microfibrils of different size and structure.

Cellulose biosynthesis arose in prokaryotes [14, 16, 17], is best known in plants [14, 18], and occurs in an oddly scattered range of other eukaryotes [19-22], possibly tracing more than one horizontal transfer of cellulose synthase genes and regulatory elements from bacteria [14, 23]. Cellulose microfibrils vary greatly in thickness and structure depending on the geometry of the arrays of cellulose synthases that assemble them [21]. The largest and most crystalline cellulose microfibrils are those made by tunicates and by some of the green algae, containing several hundred parallel cellulose chains in a cellulose Iα or Iβ lattice up to 15 nm thick [21, 22], with the hydrogen-bonded sheets of chains running diagonally across the rectangular cross-section.

The economically important cellulose microfibrils from higher plants are typically about 3 nm in diameter [3, 5, 7, 10]. There is a growing consensus that they are synthesised by an enzyme complex containing six groups of three synthase units [24, 25]. That would give 18-chain microfibrils, but partial fusion after synthesis, or accretion of non-cellulosic polymers, means that the average number of cellulose chains can be larger than this. (Here fusion is defined as the formation of a single crystalline unit, whereas aggregation means lateral contact without crystalline continuity). When higher plants make cellulose in the near absence of other polymers, in specialised organs of higher plants such as flax phloem fibres [26], tension wood fibres [11] or cotton hairs [27], a single microfibril can contain up to about 80 parallel chains in a lattice that resembles cellulose Iβ. It has been suggested that some fruits contain cellulose microfibrils with less than 18 chains [28].

3. Each cellulose microfibril contains chains of different types.

It has long been known that the thin microfibrils of cellulose found in higher plants contain both crystalline and disordered chains, or chain segments [29]. The crystallographically characterised Iα and Iβ forms were until recently considered as models for the crystalline fraction of higher-plant cellulose, but crystallography is difficult when disordered regions are present and when the lateral dimensions are small [30].

A long series of solid-state NMR studies [5, 9, 29, 31] showed that at least two forms of cellulose co-exist in microfibrils from higher plants: a crystalline form with the *tg* conformation at C6, as in cellulose Iα and Iβ, and more mobile, less ordered forms with the *gt* or *gg* C6 conformation (Figure 1). The significance of the C6 conformation lies in its controlling effect on the pattern of hydrogen bonding [4]. In the *tg* conformation, O6 points

backward along the chain and is hydrogen-bonded to O-2 on the preceding glucose unit [4] . There is also an O3H-O5 hydrogen bond on the other side of the chain, so that each glycosidic linkage between successive glucose units is flanked by two stabilising hydrogen bonds [4] (Figure 1). The *gg* and *gt* conformations at C-6 favour hydrogen bonding directed outward rather than between glucose units of the same chain [31], with mechanical consequences described below.

Much more detailed structural descriptions of the internal structures of cellulose microfibrils have recently emerged from the application of multidimensional solid-state NMR methods [32]. The distinction between chain types based on C-6 conformation was confirmed by 2D NMR in secondary cell walls of Arabidopsis [33]. In Arabidopsis and grass primary cell walls, two chain types with the *gt* conformation at C-6 were identified: in addition there were no less than five chain types with the *tg* conformation, all slightly different from one another, located in different places within the microfibril and, all but one, slightly different from crystalline cellulose Iα or Iβ [32].

4. Chains of different types have characteristic locations within the microfibril

In wood and primary cell walls, 30-40% of the cellulose has C-6 in the *tg* conformation like the known crystalline forms and the rest does not [5, 8, 9]. Crystalline and non-crystalline cellulose were formerly considered to alternate along the microfibril. This model was consistent with the rather short crystallographic coherence along the fibre axis, some tens of nm in typical wood celluloses [8]. However in the NMR community it has been contended for a long time that this alternating model of the disposition of 30-40% crystalline, C6-*tg* and 60-70% disordered, C6-*gt*/*gg* cellulose is incorrect, and that crystalline and disordered forms co-exist in each cross- section of the microfibril rather than alternating along its axis [9, 29, 31]. This contention is based on spin-diffusion NMR experiments indicating that the spatial separation between these two forms is less than the width of a microfibril, not tens of nm as required by the alternating model. Also, the alternating model requires an assumption that the disordered segments of the microfibril have an open structure permeable to water, to explain the observation that their hydroxyl protons exchange with D_2O; but deuteration-WANS experiments showed that the spacing of the deuterated chains was no greater than in crystalline cellulose [10, 34], consistent with these chains being located at a solid surface. The proportion of crystalline, C6-*tg* chains increases to 60% or more in celluloses like flax and cotton whose larger diameter gives them a smaller surface: volume ratio [29].

Recent multidimensional NMR experiments [32, 33] have confirmed that all seven cellulose chain types in primary cell walls are spatially too close together to be consistent with the alternating model, and that the chains with *gt* C-6 conformation are well hydrated, in agreement with a location at the microfibril surface [5, 31]. The limited axial coherence shown by the X-ray diffraction patterns has other explanations: twist has been suggested [5], bending is possible [35] and the converging and diverging of bound hemicelluloses [33] would also limit the axial coherence of the lattice in which they are partially incorporated.

However an unexpected recent finding was that non-cellulosic polymers – glucuronoarabinoxylans [33] and lignin [36] in Arabidopsis secondary cell walls; pectins and to a lesser extent xyloglucans in Arabidopsis primary cell walls [37]; glucuroarabinoxylans in primary cell walls of grasses [38] – were

more closely associated with a subset of the *tg*-conformation cellulose chains that had been thought to occupy interior locations.

For secondary cell walls the tentative explanation was that the glucurono-arabinoxylans are hydrogen-bonded alongside *tg*-conformation cellulose chains in a similar twofold helical chain conformation (Figure 2), so that the glucuronoarabinoxylans could almost be said to extend the cellulose lattice sideways and the cellulose chains to which they are hydrogen-bonded are then in an interior-like location and adopt the same *tg* C-6 conformation as the interior chains.

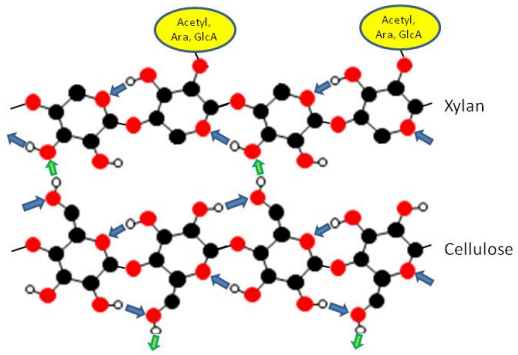

Figure 2. A glucuronoxylan chain in the 2_1 helical conformation can bind edge-on to a cellulose chain without disturbing the C-6 (*tg*) conformation of the 'crystalline' cellulose. Note the gaps left by the absence of C-6 in xylose: these might be filled by water molecules. Other β(1,4')-linked hemicellulosic chains could in principle bind similarly if (1) they can adopt the same 2_1 helical chain conformation with repeat distance 1.03 nm to match cellulose; (2) the substituents on the hemicellulose chain all face away from the cellulose surface; (3) the O3H-O5 intramolecular hydrogen bond in the hemicellulose is retained so that O3 can function as receptor for an intermolecular hydrogen bond from the cellulose O6, with C-6 (*tg*). Hemicellulosic hexosans would not leave gaps like xylans, however.

The old term hemicelluloses is fortuitously appropriate for these glucuronoarabinoxylans because in this bound form, the inner half of the molecule is quite like a cellulose chain; only, lacking C-6, it cannot form the axial O2H-O6 hydrogen bond found in *tg* cellulose chains and is forced to hydrogen bond inward. Other hemicelluloses share this capacity for adopting asymmetric, two-fold helical chain conformations but are wholly or partially unable, for different reasons, to form the O2H-O6 hydrogen bond that stabilises the two-fold helical conformation and increases the tensile modulus in crystalline cellulose [39].

In primary cell walls the hemicelluloses present – xyloglucans in Arabidopsis, arabinoxylans in grasses – appear to interact in a similar way with a smaller subset of *tg* cellulose chains [32]. A different subset of *tg*-conformation chains were in closer proximity to water than would be expected in the microfibril interior [32], as earlier observed [40]. In dicot primary cell walls pectic galacturonans are spatially associated with cellulose chains in the *tg*-conformation [32, 41]. Galacturonans too can adopt a twofold helical conformation, but the glycosidic linkage geometry means that the galacturonan chain is much more puckered than cellulose or hemicellulose chains and the pectic repeat distance [42] is too short to match. How pectins become spatially associated with cellulose is therefore unclear.

While the general idea of a concentric, non-alternating arrangement of chains in the microfibril is supported by recent evidence, the spatial disposition of the chains is unexpectedly complex and is not yet fully clear. However it is evident that hemicellulose chains can be considered as integral parts of the microfibril structure, and that if hemicelluloses are removed during nanocellulose manufacture the native

structure of the microfibrils can be expected to change.

5. **Microfibrils can be enlarged by fusion or by accretion of hemicelluloses**

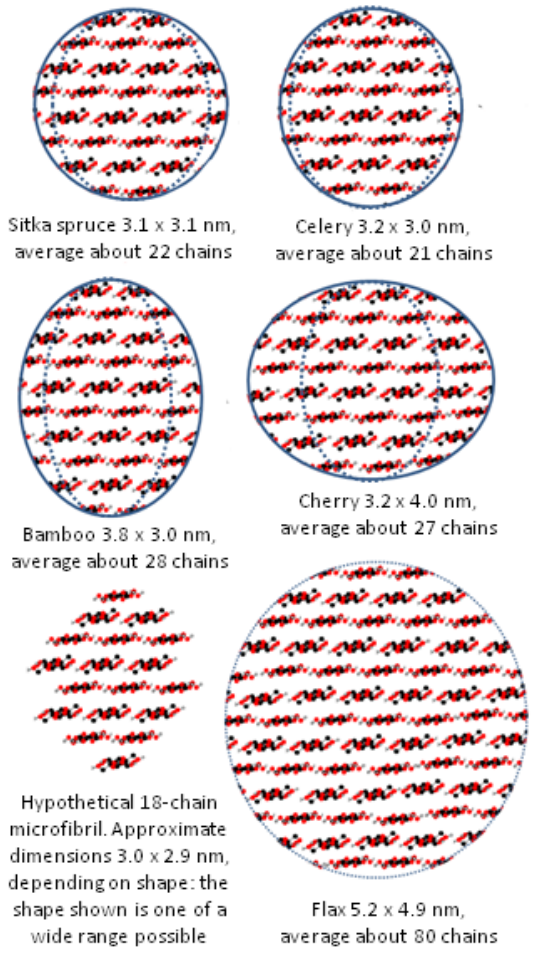

Figure 3. Approximate dimensions of microfibrils from primary and secondary cell walls of diverse higher plants, as estimated by WAXS using the Scherrer equation (dotted) and by SANS (solid line). The SANS estimate is based on the centre-to-centre distance when close-packed through the hydrophilic faces of the microfibrils. The smaller WAXS estimate of width approximates to the distance over which crystallographic coherence is sustained, and the difference between the SANS and WAXS estimates may therefore be due to bound non-cellulosic polymers. Chain conformation and packing are based throughout on the cellulose Iβ structure, ignoring the known conformational heterogeneity and the wider inter-sheet spacing found in small microfibrils. The microfibril shape is diagrammatically shown as elliptical to avoid any assumptions about which crystallite faces are exposed. For comparison, also shown is an example of the range of structures that a newly formed 18-chain microfibril might assume.

In secondary cell walls of dicots, small-angle neutron scattering studies indicated that the microfibrils were around 4 nm wide [34, 43], exceeding the ca-3nm width consistent with an 18-chain structure (Figure 3). The small-angle scattering experiments required deuterium exchange at the interface to generate contrast [34], so the dimension concerned is probably across the hydrophilic faces of the microfibrils. The extra width has been attributed to interpolation of a xylan chain [34], fusion of the microfibrils over part of their length [34], or interpolation of bound water [43]. However in Arabidopsis secondary cell walls, glucuronoarabinoxylan chains in the twofold helical conformation are abundant enough to occupy a considerable fraction of the microfibril surface [33], and an increase in the mean diameter of the microfibrils due to bound xylan therefore appears to be the principal explanation. That would explain why the microfibril width in that direction was smaller when estimated from wide-angle neutron or X-ray scattering [34] (Figure 3), because with wide-angle scattering any increase in the lattice dimensions from xylan masquerading as cellulose would be offset by the simultaneous increase in lattice disorder [26].

The attachment of cellulose-binding modules to the hydrophobic faces of cellulose microfibrils [44] is evidence that these form part of the microfibril surface in addition to diverse hydrophilic faces, but the fraction of the microfibril surface occupied by each of

these faces is uncertain and need not be constant: we know little about the shape of the small microfibrils found in higher plants. The fraction of the total number of hydroxyls that are accessible to deuterium exchange has been measured [45-47], but the magnitude of this fraction tells us little about microfibril shape because it is influenced also by the inaccessibility of interfaces between microfibrils in contact with one another [34, 47] and by the stability of the O3H-O5 hydrogen bond [46].

Although grass arabinoxylans have also been shown to adopt the twofold helical conformation required for binding to cellulose [32], small-angle neutron scattering experiments on bamboo microfibrils [7] showed that their centre-to-centre spacing was only about 3 nm across the hydrophilic faces, leaving insufficient room for an arabinoxylan chain between. However the bamboo microfibrils were nearly 4 nm in mean diameter across the hydrophobic faces as estimated by wide-angle X-ray and neutron scattering [7] (Figure 3). These dimensions are consistent with measurements by atomic force microscopy of maize cell walls [6], although the hydrophobic and hydrophilic faces of the microfibril cannot be assigned by AFM. Unless grass arabinoxylans bind in a different way from the glucuronoarabinoxylans of dicot secondary cell walls, these observations suggest that the mean width of grass microfibrils is extended more in the direction perpendicular to the sheets of chains (the hydrophobic direction) and thus more probably by local fusion of 18-chain microfibrils than by a xylan sheath [7].

In primary cell walls and in gymnosperm wood, the mean microfibril diameter as measured by scattering studies was only a little greater than would be expected for a 18-chain microfibril [5, 10] and either or both of the above mechanisms for augmenting the diameter could apply.

The much larger (5-7 nm) microfibrils of the textile fibres [26] and tension wood [48] must be formed by fusion of smaller units near the point of synthesis. Their approximately Iβ monocrystalline form [26] implies that the cellulose synthase complexes that assemble them must be grouped close together and travelling in the same direction along the cell membrane, not in antiparallel directions as observed for primary-wall synthesis in Arabidopsis [49]. The absence of hemicellulosic polymers when these large textile microfibrils are laid down suggests a role for the hemicelluloses, in other cell walls, keeping microfibrils apart as well as bridging between them. Interestingly, cactus spines have equally large microfibrils accompanied by an α(1,5') L-arabinan [50], a pectic polymer normally associated with non-elongating primary cell walls [51] and incapable of assuming the cellulose-like chain conformations of the hemicelluloses.

6. The mechanical properties of cellulose depend on microfibril structure

The tensile stiffness of crystalline cellulose Iβ is about 140 GPa [52]. Since the density is only 1.56 [4] it is fair to call cellulose a high-performance material. It is much stiffer than spider silk, for example [53], but has less extensibility and fracture energy. Modelling studies have shown that the tensile stiffness is heavily dependent on the two continuous lines of intramolecular hydrogen bonds flanking each cellulose chain (Figure 1), and is reduced when these are interrupted [54]. This was a puzzle, since the force constant of a hydrogen bond is almost negligible compared to a covalent bond under tension. The explanation lies in molecular leverage: tension on the cellulose chain tends to straighten its slightly zig-zag conformation, and this is

resisted by hydrogen bonds at the end of quite long levers [55, 56].

It follows that cellulose microfibrils from higher plants, usually with less than half of their chains adopting the *tg* C6 conformation required for the molecular leverage effect, probably have less than the 140 GPa tensile modulus that is normally assumed for nanocelluloses [55]. So do hemicelluloses. A further consequence is that large-diameter cellulose microfibrils like those from flax, with smaller relative surface area and approximately 70% of crystalline C6-*tg* chains, will have a higher tensile modulus than cellulose from wood or from annual plants. The adaptive significance is unclear - what plants gain from having thin microfibrils with low tensile stiffness, and why sometimes they make thicker and stiffer microfibrils [e.g. [11]].

A practical consequence for nanocellulose technology is that the small increase in microfibril diameter observed during pulping processes [57] may slightly improve mechanical performance. More substantial increases in microfibril diameter have been observed following more extreme hydrothermal treatment [31], and may also occur during some processes for isolating cellulose nanocrystals [58].

7. Axial disorder is limited, but may facilitate cellulose hydrolysis

It has often been suggested that the formation of nanocrystalline cellulose involves the acid hydrolysis of non-crystalline microfibril segments alternating with segments that are crystalline [58]. This proposed mechanism arose from the assumption that the large quantity of disordered cellulose present alternated with crystalline domains along the microfibril, a model of microfibril structure that in its simplest form is now superseded as explained above.

Nevertheless, local deviations from crystalline regularity, that could form starting points for acid or enzymatic hydrolysis, may exist. Indeed, they must exist; for cellulose molecules are much shorter than microfibrils [59]. We know nothing of the structural details where one chain ends and another begins, but the neatest arrangement would be a gap of one glucose or one cellobiose unit, with on either side some mobility of the broken chain that might allow a glycosidic linkage to take up a conformation permitting hydrolysis. We know too little about whether chain ends are randomly located or staggered, or how this is controlled [60] to let us calculate how far apart they are along the microfibril, but several hundred nm would be consistent with the degree of polymerisation of cellulose.

Additional sources of local axial disorder have been described. Mechanical disruption of cellulose fibres during their isolation leads to damaged segments [61, 62], and it is possible that microfibrils are sometimes kinked even without such damage [63]. Such sites could provide starting points for hydrolysis without comprising a large fraction of the non-crystalline cellulose present. Hydrolysis could then propagate in both directions along the microfibril, with the necessary chain flexibility coming from proximity to the broken end. For a better understanding of the preparation of nanocrystalline cellulose, we need to understand the nature and origins of the axial discontinuities in the structure of the original microfibrils. The abundant disordered material at the microfibril surface does not help us to understand these axial discontinuities.

8. Microfibrils of wood cellulose are aggregated

It is well established that in conifer wood, the microfibrils are aggregated into bundles that are typically 10-20 nm thick [5, 64, 65]. The

dimensions of cellulose nanofibrils made from conifer wood indicate that they comprise these or larger bundles of microfibrils. The microfibril spacing within the bundles can be estimated at about 3 nm by small-angle neutron scattering [5], consistent with at least some regions of direct microfibril-to-microfibril contact. The spacing widens a little on hydration, [5], and presumably the microfibrils adhere by hydrogen bonding. Conifer hemicelluloses include acetylated and non-acetylated glucomannans as well as xylans, but their spatial distribution inside or outside the microfibril aggregates is uncertain: the xylans are structurally capable of binding to cellulose microfibrils [66]. In angiosperm wood there is more limited evidence for similar aggregation of microfibrils [64]. The small-angle scattering signal implying aggregation is weaker than in conifers [34] but this is not necessarily inconsistent with aggregation: it could simply mean that the microfibril spacing within the aggregates is less uniform. Bamboo microfibrils, in contrast, are strongly and regularly aggregated [7].

In celery collenchyma cell walls, small-angle X-ray and neutron scattering readily demonstrated aggregation of the microfibrils at a spacing that expanded much further on hydration than in any secondary cell walls [10]. Celery collenchyma cell walls, on account of their polymer composition, are in most respects a good model for the angiosperm primary cell wall [67]. However their cellulose orientation is much more uniform than in other primary cell walls [10] and it is likely that in more conventional primary cell walls, microfibril aggregation is weak or confined to single wall layers [68].

9. **Nanocelluloses are found in Nature**

When the coat of a quince seed splits and water penetrates, the helicoidal structure of the inner seed coat hydrates almost explosively and a sticky mucilage envelops the seed, protecting it from dehydration and microbial attack [69]. The mucilage consists of single microfibrils of cellulose [70], coated with a remarkable glucuronoxylan having half of its xylosyl units substituted with 4-O-methyl glucuronic acid [71]. The surface charges appear to be sufficient to disperse the microfibrils spontaneously in exactly the way that is required of a nanofibrillated cellulose, and the properties of the mucilage layer around the seed match some of the technological applications targeted: in particular, the mucilage is a highly effective water-based lubricant [72]. The long, dispersed microfibrils have a smooth and uniform cross-section [69] that is a visual demonstration of the lack of alternating crystallinity and disorder along the microfibril axis.

Quince seeds are not commercially harvested, but other seed mucilages containing cellulose exist. For example the mucilage from Arabidopsis seeds is a complex bilayered structure that depends on pectic rhamnogalacturonan for its charges, but incorporates cellulose, xylan and glucomannan [73]. A wider search might reveal seed mucilages with more favourable prospects for commercial utilisation.

[1] Klemm, D., Kramer, F., Moritz, S., Lindstrom, T., Ankerfors, M., Gray, D. & Dorris, A. 2011 Nanocelluloses: a new family of Nature-based materials. *Angewandte Chemie-International Edition* **50**, 5438-5466. (DOI:10.1002/anie.201001273).

[2] Fratzl, P. & Weinkamer, R. 2007 Nature's hierarchical materials. *Progress in Materials Science* **52**, 1263-1334. (DOI:10.1016/j.pmatsci.2007.06.001).

[3] Newman, R. H., Hill, S. J. & Harris, P. J. 2013 Wide-angle X-ray scattering and solid-state nuclear magnetic resonance data combined to test models for cellulose microfibrils in mung bean cell walls. *Plant Physiology* **163**, 1558-1567. (DOI:10.1104/pp.113.228262).

[4] Nishiyama, Y., Langan, P. & Chanzy, H. 2002 Crystal structure and hydrogen-bonding system in cellulose 1 beta from synchrotron X-ray and neutron fiber diffraction. *Journal of the American Chemical Society* **124**, 9074-9082. (DOI:10.1021/ja0257319).
[5] Fernandes, A. N., Thomas, L. H., Altaner, C. M., Callow, P., Forsyth, V. T., Apperley, D. C., Kennedy, C. J. & Jarvis, M. C. 2011 Nanostructure of cellulose microfibrils in spruce wood. *Proceedings of the National Academy of Sciences of the United States of America* **108**, E1195-E1203. (DOI:10.1073/pnas.1108942108).
[6] Ding, S. Y. & Himmel, M. E. 2006 The maize primary cell wall microfibril: A new model derived from direct visualization. *Journal of Agricultural and Food Chemistry* **54**, 597-606. (DOI:10.1021/jf051851z).
[7] Thomas, L. H., Forsyth, V. T., Martel, A., Grillo, I., Altaner, C. M. & Jarvis, M. C. 2015 Diffraction evidence for the structure of cellulose microfibrils in bamboo, a model for grass and cereal celluloses. *Bmc Plant Biology* **15**. (DOI:10.1186/s12870-015-0538-x).
[8] Andersson, S., Serimaa, R., Paakkari, T., Saranpaa, P. & Pesonen, E. 2003 Crystallinity of wood and the size of cellulose crystallites in Norway spruce (*Picea abies*). *Journal of Wood Science* **49**, 531-537. (DOI:10.1007/s10086-003-0518-x).
[9] Wickholm, K., Larsson, P. T. & Iversen, T. 1998 Assignment of non-crystalline forms in cellulose I by CP/MAS C-13 NMR spectroscopy. *Carbohydrate Research* **312**, 123-129. (DOI:10.1016/s0008-6215(98)00236-5).
[10] Thomas, L. H., Forsyth, V. T., Sturcova, A., Kennedy, C. J., May, R. P., Altaner, C. M., Apperley, D. C., Wess, T. J. & Jarvis, M. C. 2013 Structure of cellulose microfibrils in primary cell walls from collenchyma. *Plant Physiology* **161**, 465-476. (DOI:10.1104/pp.112.206359).
[11] Wada, M., Okano, T., Sugiyama, J. & Horii, F. 1995 Characterization of tension and normally lignified mood cellulose in *Populus maximowiczii*. *Cellulose* **2**, 223-233. (DOI:10.1007/bf00811814).
[12] Park, S., Baker, J. O., Himmel, M. E., Parilla, P. A. & Johnson, D. K. 2010 Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnology for Biofuels* **3**. (DOI:10.1186/1754-6834-3-10).
[13] Baker, A. A., Helbert, W., Sugiyama, J. & Miles, M. J. 2000 New insight into cellulose structure by atomic force microscopy shows the I-alpha crystal phase at near-atomic resolution. *Biophysical Journal* **79**, 1139-1145.
[14] Kumar, M. & Turner, S. 2015 Plant cellulose synthesis: CESA proteins crossing kingdoms. *Phytochemistry* **112**, 91-99. (DOI:10.1016/j.phytochem.2014.07.009).
[15] Purushotham, P., Cho, S. H., Diaz-Moreno, S. M., Kumar, M., Nixon, B. T., Bulone, V. & Zimmer, J. 2016 A single heterologously expressed plant cellulose synthase isoform is sufficient for cellulose microfibril formation in vitro. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 11360-11365. (DOI:10.1073/pnas.1606210113).
[16] Nobles, D. R., Romanovicz, D. K. & Brown, R. M. 2001 Cellulose in cyanobacteria. Origin of vascular plant cellulose synthase? *Plant Physiology* **127**, 529-542. (DOI:10.1104/pp.010557).
[17] Ross, P., Mayer, R. & Benziman, M. 1991 Cellulose biosynthesis and function in bacteria. *Microbiological Reviews* **55**, 35-58.
[18] Somerville, C. 2006 Cellulose synthesis in higher plants. In *Annual Review of Cell and Developmental Biology* (pp. 53-78.
[19] Blanton, R. L., Fuller, D., Iranfar, N., Grimson, M. J. & Loomis, W. F. 2000 The cellulose synthase gene of Dictyostelium. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 2391-2396. (DOI:10.1073/pnas.040565697).
[20] Dudley, R., Jarroll, E. L. & Khan, N. A. 2009 Carbohydrate analysis of *Acanthamoeba castellanii*. *Experimental Parasitology* **122**, 338-343. (DOI:10.1016/j.exppara.2009.04.009).
[21] Koyama, M., Sugiyama, J. & Itoh, T. 1997 Systematic survey on crystalline features of algal celluloses. *Cellulose* **4**, 147-160. (DOI:10.1023/a:1018427604670).
[22] Nakashima, K., Sugiyama, J. & Satoh, N. 2008 A spectroscopic assessment of cellulose and the molecular mechanisms of cellulose biosynthesis in the ascidian *Ciona intestinalis*. *Marine Genomics* **1**, 9-14. (DOI:10.1016/j.margen.2008.01.001).
[23] Sasakura, Y., Ogura, Y., Treen, N., Yokomori, R., Park, S.-J., Nakai, K., Saiga, H., Sakuma, T., Yamamoto, T., Fujiwara, S., et al. 2016 Transcriptional regulation of a horizontally transferred gene from bacterium to chordate. *Proceedings of the Royal Society B-Biological Sciences* **283**. (DOI:10.1098/rspb.2016.1712).
[24] Nixon, B. T., Mansouri, K., Singh, A., Du, J., Davis, J. K., Lee, J. G., Slabaugh, E., Vandavasi, V. G., O'Neill, H., Roberts, E. M., et al. 2016 Comparative structural and computational analysis supports eighteen cellulose synthases in the plant cellulose synthesis complex. *Scientific Reports* **6**. (DOI:10.1038/srep28696).
[25] Vandavasi, V. G., Putnam, D. K., Zhang, Q., Petridis, L., Heller, W. T., Nixon, B. T., Haigler, C. H.,

Kalluri, U., Coates, L., Langan, P., et al. 2016 A structural study of CESA1 catalytic domain of Arabidopsis cellulose synthesis complex: evidence for CESA trimers. *Plant Physiology* **170**, 123-135. (DOI:10.1104/pp.15.01356).
[26] Thomas, L. H., Altaner, C. M. & Jarvis, M. C. 2013 Identifying multiple forms of lateral disorder in cellulose fibres. *Journal of Applied Crystallography* **46**, 972-979. (DOI:10.1107/s002188981301056x).
[27] Nam, S., French, A. D., Condon, B. D. & Concha, M. 2016 Segal crystallinity index revisited by the simulation of X-ray diffraction patterns of cotton cellulose I beta and cellulose II. *Carbohydrate Polymers* **135**, 1-9. (DOI:10.1016/j.carbpol.2015.08.035).
[28] Niimura, H., Yokoyama, T., Kimura, S., Matsumoto, Y. & Kuga, S. 2010 AFM observation of ultrathin microfibrils in fruit tissues. *Cellulose* **17**, 13-18. (DOI:10.1007/s10570-009-9361-6).
[29] Atalla, R. H. & VanderHart, D. L. 1999 The role of solid state C-13 NMR spectroscopy in studies of the nature of native celluloses. *Solid State Nuclear Magnetic Resonance* **15**, 1-19. (DOI:10.1016/s0926-2040(99)00042-9).
[30] Chanzy, H., Imada, K., Mollard, A., Vuong, R. & Barnoud, F. 1979 Crystallographic aspec ts of sub-elementary cellulose fibrils occurring in the wall of rose cells cultured in vitro. *Protoplasma* **100**, 303-316. (DOI:10.1007/bf01279318).
[31] Sturcova, A., His, I., Apperley, D. C., Sugiyama, J. & Jarvis, M. C. 2004 Structural details of crystalline cellulose from higher plants. *Biomacromolecules* **5**, 1333-1339. (DOI:10.1021/bm034517p).
[32] Wang, T., Yang, H., Kubicki, J. D. & Hong, M. 2016 Cellulose structural polymorphism in plant primary cell walls investigated by high-field 2D solid-state NMR spectroscopy and density functional theory calculations. *Biomacromolecules* **17**, 2210-2222. (DOI:10.1021/acs.biomac.6b00441).
[33] Simmons, T. J., Mortimer, J. C., Bernardinelli, O. D., Poppler, A.-C., Brown, S. P., deazevedo, E. R., Dupree, R. & Dupree, P. 2016 Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nature Communications* **7**. (DOI:10.1038/ncomms13902).
[34] Thomas, L. H., Forsyth, V. T., Martel, A., Grillo, I., Altaner, C. M. & Jarvis, M. C. 2014 Structure and spacing of cellulose microfibrils in woody cell walls of dicots. *Cellulose* **21**, 3887-3895. (DOI:10.1007/s10570-014-0431-z).
[35] Jarvis, M. C. 2000 Interconversion of the I alpha and I beta crystalline forms of cellulose by bending. *Carbohydrate Research* **325**, 150-154. (DOI:10.1016/s0008-6215(99)00316-x).
[36] Dupree, R., Simmons, T. J., Mortimer, J. C., Patel, D., Iuga, D., Brown, S. P. & Dupree, P. 2015 Probing the molecular architecture of *Arabidopsis thaliana* secondary cell walls using two- and three-dimensional C-13 solid state nuclear magnetic resonance spectroscopy. *Biochemistry* **54**, 2335-2345. (DOI:10.1021/bi501552k).
[37] Wang, T., Park, Y. B., Cosgrove, D. J. & Hong, M. 2015 Cellulose-pectin spatial contacts are inherent to never-dried Arabidopsis primary cell walls: evidence from solid-state nuclear magnetic resonance. *Plant Physiology* **168**, 871-+. (DOI:10.1104/pp.15.00665).
[38] Wang, T., Salazar, A., Zabotina, O. A. & Hong, M. 2014 Structure and dynamics of *Brachypodium* primary cell wall polysaccharides from two-dimensional C-13 solid-state nuclear magnetic resonance spectroscopy. *Biochemistry* **53**, 2840-2854. (DOI:10.1021/bi500231b).
[39] Scheller, H. V. & Ulvskov, P. 2010 Hemicelluloses. In *Annual Review of Plant Biology, Vol 61* (eds. S. Merchant, W. R. Briggs & D. Ort), pp. 263-289.
[40] Hediger, S., Lesage, A. & Emsley, L. 2002 A new NMR method for the study of local mobility in solids and application to hydration of biopolymers in plant cell walls. *Macromolecules* **35**, 5078-5084. (DOI:10.1021/ma020065h).
[41] Thimm, J. C., Burritt, D. J., Ducker, W. A. & Melton, L. D. 2009 Pectins influence microfibril aggregation in celery cell walls: An atomic force microscopy study. *Journal of Structural Biology* **168**, 337-344. (DOI:10.1016/j.jsb.2009.06.017).
[42] Makshakova, O. N., Gorshkova, T. A., Mikshina, P. V., Zuev, Y. F. & Perez, S. 2017 Metrics of rhamnogalacturonan I with beta-(1 -> 4)-galactan side chains and structural basis for its self-aggregation. *Carbohydrate Polymers* **158**, 93-101. (DOI:10.1016/j.carbpol.2016.11.082).
[43] Langan, P., Petridis, L., O'Neill, H. M., Pingali, S. V., Foston, M., Nishiyama, Y., Schulz, R., Lindner, B., Hanson, B. L., Harton, S., et al. 2014 Common processes drive the thermochemical pretreatment of lignocellulosic biomass. *Green Chemistry* **16**, 63-68. (DOI:10.1039/c3gc41962b).
[44] Bubner, P., Plank, H. & Nidetzky, B. 2013 Visualizing cellulase activity. *Biotechnology and Bioengineering* **110**, 1529-1549. (DOI:10.1002/bit.24884).
[45] Lemke, C. H., Dong, R. Y., Michal, C. A. & Hamad, W. Y. 2012 New insights into nano-crystalline cellulose structure and morphology based on solid-state NMR. *Cellulose* **19**, 1619-1629. (DOI:10.1007/s10570-012-9759-4).
[46] Lindh, E. L., Bergenstrahle-Wohlert, M., Terenzi, C., Salmen, L. & Furo, I. 2016 Non-exchanging hydroxyl groups on the surface of

cellulose fibrils: The role of interaction with water. *Carbohydrate Research* **434**, 136-142. (DOI:10.1016/j.carres.2016.09.006).
[47] Lindh, E. L. & Salmen, L. 2017 Surface accessibility of cellulose fibrils studied by hydrogen-deuterium exchange with water. *Cellulose* **24**, 21-33. (DOI:10.1007/s10570-016-1122-8).
[48] Clair, B., Almeras, T., Pilate, G., Jullien, D., Sugiyama, J. & Riekel, C. 2011 Maturation stress generation in poplar tension wood studied by synchrotron radiation microdiffraction. *Plant Physiology* **155**, 562-570. (DOI:10.1104/pp.110.167270).
[49] Li, S., Bashline, L., Zheng, Y., Xin, X., Huang, S., Kong, Z., Kim, S. H., Cosgrove, D. J. & Gu, Y. 2016 Cellulose synthase complexes act in a concerted fashion to synthesize highly aggregated cellulose in secondary cell walls of plants. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 11348-11353. (DOI:10.1073/pnas.1613273113).
[50] Vignon, M. R., Heux, L., Malainine, M. E. & Mahrouz, M. 2004 Arabinan-cellulose composite in *Opuntia ficus-indica* prickly pear spines. *Carbohydrate Research* **339**, 123-131. (DOI:10.1016/j.carres.2003.09.023).
[51] Willats, W. G. T., Steele-King, C. G., Marcus, S. E. & Knox, J. P. 1999 Side chains of pectic polysaccharides are regulated in relation to cell proliferation and cell differentiation. *Plant Journal* **20**, 619-628. (DOI:10.1046/j.1365-313X.1999.00629.x).
[52] Sturcova, A., Davies, G. R. & Eichhorn, S. J. 2005 Elastic modulus and stress-transfer properties of tunicate cellulose whiskers. *Biomacromolecules* **6**, 1055-1061. (DOI:10.1021/bm049291k).
[53] Vollrath, F. & Porter, D. 2006 Spider silk as archetypal protein elastomer. *Soft Matter* **2**, 377-385. (DOI:10.1039/b600098n).
[54] Cintron, M. S., Johnson, G. P. & French, A. D. 2011 Young's modulus calculations for cellulose I-beta by MM3 and quantum mechanics. *Cellulose* **18**, 505-516. (DOI:10.1007/s10570-011-9507-1).
[55] Altaner, C. M., Thomas, L. H., Fernandes, A. N. & Jarvis, M. C. 2014 How cellulose stretches: synergism between covalent and hydrogen bonding. *Biomacromolecules* **15**, 791-798. (DOI:10.1021/bm401616n).
[56] Djahedi, C., Bergenstrahle-Wohlert, M., Berglund, L. A. & Wohlert, J. 2016 Role of hydrogen bonding in cellulose deformation: the leverage effect analyzed by molecular modeling. *Cellulose* **23**, 2315-2323. (DOI:10.1007/s10570-016-0968-0).
[57] Hult, E. L., Iversen, T. & Sugiyama, J. 2003 Characterization of the supermolecular structure of cellulose in wood pulp fibres. *Cellulose* **10**, 103-110. (DOI:10.1023/a:1024080700873).
[58] Bondeson, D., Mathew, A. & Oksman, K. 2006 Optimization of the isolation of nanocrystals from microcrystalline cellulose by acid hydrolysis. *Cellulose* **13**, 171-180. (DOI:10.1007/s10570-006-9061-4).
[59] Shinoda, R., Saito, T., Okita, Y. & Isogai, A. 2012 Relationship between length and degree of polymerization of TEMPO-oxidized cellulose nanofibrils. *Biomacromolecules* **13**, 842-849. (DOI:10.1021/bm2017542).
[60] Lei, L., Zhang, T., Strasser, R., Lee, C. M., Gonneau, M., Mach, L., Vernhettes, S., Kim, S. H., Cosgrove, D. J., Li, S. D., et al. 2014 The jiaoyao1 mutant Is an allele of korrigan1 that abolishes endoglucanase activity and affects the organization of both cellulose microfibrils and microtubules in Arabidopsis. *Plant Cell* **26**, 2601-2616. (DOI:10.1105/tpc.114.126193).
[61] Hidayat, B. J., Felby, C., Johansen, K. S. & Thygesen, L. G. 2012 Cellulose is not just cellulose: a review of dislocations as reactive sites in the enzymatic hydrolysis of cellulose microfibrils. *Cellulose* **19**, 1481-1493. (DOI:10.1007/s10570-012-9740-2).
[62] Thygesen, L. G., Bilde-Sorensen, J. B. & Hoffmeyer, P. 2006 Visualisation of dislocations in hemp fibres: A comparison between scanning electron microscopy (SEM) and polarized light microscopy (PLM). *Industrial Crops and Products* **24**, 181-185. (DOI:10.1016/j.indcrop.2006.03.009).
[63] Usov, I., Nystroem, G., Adamcik, J., Handschin, S., Schutz, C., Fall, A., Bergstrom, L. & Mezzenga, R. 2015 Understanding nanocellulose chirality and structure-properties relationship at the single fibril level. *Nature Communications* **6**. (DOI:10.1038/ncomms8564).
[64] Donaldson, L. 2007 Cellulose microfibril aggregates and their size variation with cell wall type. *Wood Science and Technology* **41**, 443-460. (DOI:10.1007/s00226-006-0121-6).
[65] Fahlen, J. & Salmen, L. 2005 Pore and matrix distribution in the fiber wall revealed by atomic force microscopy and image analysis. *Biomacromolecules* **6**, 433-438. (DOI:10.1021/bm040068x).
[66] Busse-Wicher, M., Li, A., Silveira, R. L., Pereira, C. S., Tryfona, T., Gomes, T. C. F., Skaf, M. S. & Dupree, P. 2016 Evolution of Xylan Substitution Patterns in Gymnosperms and Angiosperms: Implications for Xylan Interaction with Cellulose. *Plant Physiology* **171**, 2418-2431. (DOI:10.1104/pp.16.00539).

[67] Zujovic, Z., Chen, D. & Melton, L. D. 2016 Comparison of celery (Apium graveolens L.) collenchyma and parenchyma cell wall polysaccharides enabled by solid-state C-13 NMR. *Carbohydrate Research* **420**, 51-57. (DOI:10.1016/j.carres.2015.11.006).
[68] Zhang, T., Zheng, Y. & Cosgrove, D. J. 2016 Spatial organization of cellulose microfibrils and matrix polysaccharides in primary plant cell walls as imaged by multichannel atomic force microscopy. *Plant Journal* **85**, 179-192. (DOI:10.1111/tpj.13102).
[69] Reis, D., Vian, B., Chanzy, H. & Roland, J. C. 1991 Liquid-crystal-type assembly of native cellulose-glucuronoxylans extracted from plant cell walls. *Biol. Cell* **73**, 173-178. (DOI:10.1016/0248-4900(91)90100-2).
[70] Vietor, R. J., Newman, R. H., Ha, M. A., Apperley, D. C. & Jarvis, M. C. 2002 Conformational features of crystal-surface cellulose from higher plants. *Plant Journal* **30**, 721-731. (DOI:10.1046/j.1365-313X.2002.01327.x).
[71] Vignon, M. R. & Gey, C. 1998 Isolation, H-1 and C-13 NMR studies of (4-O-methyl-D-glucurono)-D-xylans from luffa fruit fibres, jute bast fibres and mucilage of quince tree seeds. *Carbohydrate Research* **307**, 107-111. (DOI:10.1016/s0008-6215(98)00002-0).
[72] Hakala, T. J., Saikko, V., Arola, S., Ahlroos, T., Helle, A., Kuosmanen, P., Holmberg, K., Linder, M. B. & Laaksonen, P. 2014 Structural characterization and tribological evaluation of quince seed mucilage. *Tribology International* **77**, 24-31. (DOI:10.1016/j.triboint.2014.04.018).
[73] Voiniciuc, C., Yang, B., Schmidt, M. H.-W., Guenl, M. & Usadel, B. 2015 Starting to gel: how Arabidopsis seed coat epidermal cells produce specialized secondary cell walls. *International Journal of Molecular Sciences* **16**, 3452-3473. (DOI:10.3390/ijms16023452).