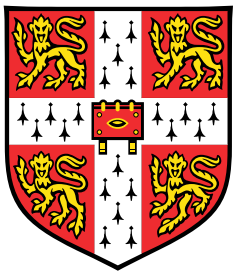
Structural colour from a helicoidal cellulose architecture in the secondary cell wall

Optical properties, cell wall composition, structure and morphology of components, and their assembly and interactions.



Lisa Maria Steiner

Department of Chemistry
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Girton College

September 2018

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the relevant Degree Committee.

Lisa Maria Steiner
September 2018

Abstract

Structural colour from a helicoidal cellulose architecture in the secondary cell wall: Optical properties, cell wall composition, structure and morphology of components, and their assembly and interactions.

Lisa Maria Steiner

Structural colours are produced by constructive interference of light scattered from periodically arranged interfaces within nanostructured materials. A common strategy of plants to achieve structural coloration consists of assembling cellulose microfibrils into helicoidal structures. Examples of these architectures can be found in phylogenetically distant species and in different tissues of plants, such as in the endocarp of the fruit of *Margaritaria nobilis*, an eudicot, and in the fronds of *Microsorum thailandicum*, a fern.

In this thesis, I studied the optical response and anatomical features of the adaxial and abaxial epidermal layers of cells of *M. thailandicum*. The optical variation between fronds, between cells, and within cells, both for the adaxial and abaxial surface, were quantified by polarised optical microscopy and microspectroscopy. The adaxial reflection is mainly in the blue range (400-550 nm), with a narrow distribution of reflection peak maxima and peak widths, while the abaxial reflection spans almost the entire visible range (400-650 nm), and spectra are broader and show more varied features. The anatomical structure responsible for this optical behaviour was confirmed by electron microscopy. Numerical modelling based on the microscopy data indicates that there is much more complexity in the helicoidal architecture than just simple local defects. The significant difference between the adaxial and abaxial epidermal cell walls hints at fundamental differences in their biosynthesis.

For the endocarp of the fruit of *M. nobilis*, I correlated the optical response to the ultrastructure of the secondary cell wall, and the morphology and chemical structure of its main components, cellulose and xylan. The composition of the endocarp was characterised as 9 % extractives, 36 % cellulose, 23 % xylan, and 30 % lignin. The cellulose microfibrils were isolated in a long series of chemical purifications, and found to be extremely short - around 500 nm - and highly crystalline, as determined via nuclear magnetic resonance spectroscopy and X-ray diffraction. Xylan was found to be the most abundant hemicellulose

and its primary structure was characterised via enzyme digestions, gel electrophoresis and mass spectrometry. It has acetyl groups on every other xylose unit, no arabinose and around 4 % of glucuronic acid decoration. By combining the information on the xylan with the morphological characteristics of the cellulose microfibrils, I speculate about their assembly into the helicoidal architecture during the cell wall biosynthesis, based on findings from coarsed-grain modelling, and on their interactions in the final tissue, based on small angle X-ray scattering studies.

Acknowledgements

The work in this thesis has drawn from the outstanding knowledge, help and support from a wide range of people that I would like to thank. First of all, I would like to thank my supervisor, Silvia Vignolini, for teaching me optics and microscopy, for giving me all the freedom and trust I needed to carry out such a fundamental research project, and for letting me attend several conferences that have all been extremely valuable to this work. Thanks also to all of the Vignolini group for emotional support and for teaching me about optics and matlab - and to Alyssa Smith for bio-advices, and in particular the Living Meetings, from which many good discussions and ideas have come. I also want to thank Matt Pond for being the best lab technician ever! I have also felt very supported, also about my pregnancy from all the Vignolini group and the rest of the Melville Lab.

I would also like to thank Paul Dupree (Department of Biochemistry) and his group for allowing me to work in their lab and for being supportive and helpful, in particular Marta Busse-Wicher, who has taught me a lot and been extremely encouraging. Also Tom Simmons and Dora Tryfona should be thanked. In the Department of Chemistry, I am thankful to Dave Reid for having taught me a lot about solid state NMR, and also to Melinda Duer, and Rui and Ieva.

I would also really like to thank Yu Ogawa for his academic support and friendship. His electron microscopy knowledge and skills have been invaluable to this project, and so has his knowledge on cellulose and all his help and support. Also thanks to Cyan Williams for her help, friendship and almost-neighbourship.

I also want to thank my friends in Cambridge that have made my stay much greater and have also been a support during the PhD. I hope we will meet somewhere, somehow in the future. This includes Nicola and Christopher Beauman who took us as tenants in their house in Orchard Street. I would also like to thank my family for never forgetting to ask me when I will finish school and finally get a real job. Lastly, I would also like to thank my partner Villads, who has been extremely supportive and with whom I am excited to start a new chapter with.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Motivation

When I set out for this PhD project, about four years ago, the big question was: How is the helicoidal architecture in plant cell walls, giving rise to structural colour, made? My supervisor, Silvia Vignolini, had been working on characterising the optical properties of *Pollia condensata*, probably the shiniest fruit in the world. She had also been collaborating with some chemists and material scientists, who could replicate similar structures with comparable optical response, starting from cellulose nanocrystal suspensions, via chiral nematic self-assembly. The idea for my project was thus that surely, the whole process is similar in plants. However, soon, I adapted the main question from 'How is the helicoidal architecture in plant cell walls, giving rise to structural colour, made?' to 'What are the main building blocks of this helicoidal architecture, and what requirements of properties are there?'.

The answer to such a big question lies at the interface of many different disciplines, and possibly has implications for even more, like biochemistry, cell biology, chemistry, optics, ecology, animal behaviour, evolution, and biomimetics and biomaterials.

Studying a specific plant cell wall in detail may not have direct applications straight away, but it might open up many avenues and future directions of research, revealing general mechanisms of how polysaccharides interact in the plant cell wall. For example, biomass recalcitrance is a major hurdle on our way to sustainable and economically viable replacements for petrol-based fuels. Many efforts are made in fields like chemical engineering, to improve processing efficiency, and many research groups try to understand fundamental cell wall component interactions, in order to make the different compounds easier to separate and process.

In the scope of this thesis, the discovery that the regularity of the xylan in the *Margaritaria nobilis* endocarp is probably connected to the fact that it is difficult to separate it from the cellulose microfibrils is very important and suggests that a key factor for an efficient separation of cell wall components might rely on trying to identify genes and biosynthetic pathways with respect to xylan biosynthesis [81]. Another major finding of this thesis is that the cellulose microfibrils in the *M. nobilis* endocarp are surprisingly short. Finding out why and whether we could utilise this information by eventually genetically modifying plants used as biofuel feedstock could improve efficiency tremendously by making the cellulose easier to digest.

With respect to the other main topic in this thesis, concerning the variation of optical properties of the adaxial and abaxial cell walls of *Microsorum thailandicum*, many potential future research directions opened up as well, such as how and why plants can control their optical response in the same tissue and possibly adapt to be responsive to specific light environments. And from a cell wall biosynthesis point of view, the finding that the adaxial cell wall biosynthesis is so much tighter controlled and regulated compared to the abaxial wall is intriguing and opens many more exciting questions.

1.2 Structural colour

Structural colour does not derive from pigments, but relies on ordered, quasi-ordered or disordered structures with lattice constants or scattering element sizes of the order of the wavelength of visible radiation (400 - 700 nm) [60]. Colouration caused by this phenomenon is usually more intense than colouration by pigments, and can be dependent on the angle of observation [149, 67].

While structural colour in the animal kingdom has received a lot of attention from scientists from various disciplines [106, 33, 150], structural colour in plants has not been investigated as extensively [77, 146]. Although many examples of structural colour in flowers, leaves and fruits are currently known [153, 90, 147, 145, 76, 131, 50], many fundamental questions are still unanswered, such as how is structural colouration developed in plants [91], and what is the biological function of these colorations? Structural colour is possibly one of the cues by flowers to attract pollinators [153], and in the case of fruits for seed dispersal [147, 145]. However, examples of structural colour in leaves exist, and its purpose is still not well understood [135].

1.2.1 Structural colour from helicoidal architectures in plants

Helicoidal architectures in plants are composed of different layers or pseudolayers of cellulose microfibrils oriented parallel to each other. These layers are stacked up, with a small rotation angle between each of them. At every 180° rotation, the microfibrils have the same orientation (Figure 1.1). The distance between two equally orientated layers is defined as the pitch p, which determines the reflected wavelength λ_r by

$$\lambda_r = 2 \cdot n_a \cdot p \tag{1.1}$$

In Equation 1.1, λ_r is the wavelength of the reflected light, and n_a is the average refractive index of cellulose.

The handedness of the helicoid defines the handedness of the reflected circularly polarised light. Generally, light can have different polarisation states, and normal light, such as sunlight, contains all polarisation states in equal proportions. The circularly polarised reflection from a helicoid is thus a unique fingerprint for this type of system.

Most plants with this helicoidal arrangement reflect only left-handed circularly polarised (LCP) light, implying that the helicoid is left-handed. (This apparent contradiction is due to opposite naming conventions being used for light polarisation versus helicoidal structures. In reality, light with identical handedness to the helicoid will pass through, and be transmitted, while the opposite handedness is reflected.)

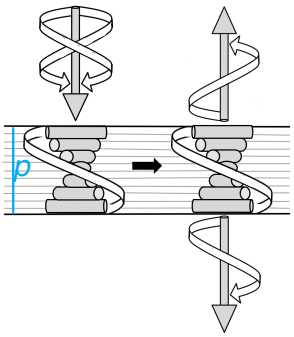


Fig. 1.1 Scheme of a helicoidal structure, with left-handed circularly polarised (LCP) light being reflected and right-handed circularly polarised (RCP) light being transmitted, p is the pitch, the distance between two equally orientated layers. The cylinders represent cellulose microfibrils. Adapted, from the author [85].

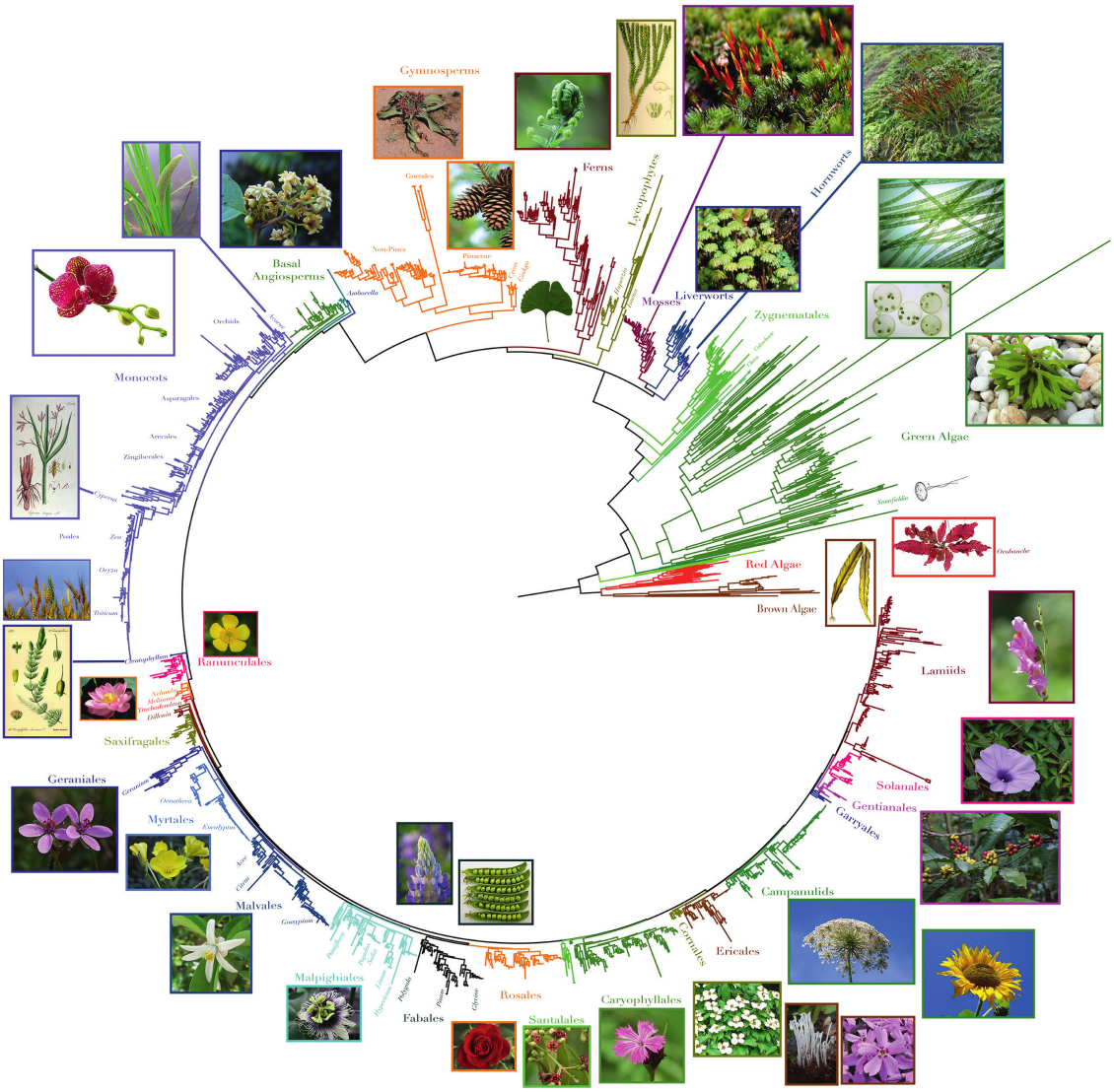
Helicoidal architectures giving rise to structural colouration can be found in many different species which are far apart on the phylogenetic tree. A phylogenetic tree is shown in Figure 1.2. The tree is based on a plastid phylogenomic analysis of Gitzendanner et al.

[45]. Another type of phylogenetic analysis is via nuclear genes, as for example done by Wickett et al. [154], their tree is shown in the appendix, Figure A.1. By now, many of the major clades of plants are defined, but there is still uncertainty about the relationships of some of the major clades to each other. In the past, a small number of samples in some clades was a limiting factor. Furthermore, it is important to compare results from plastid-based and nuclear-based analyses, since both techniques are very complex and have limitations.

Essentially, the first plants were algae, with the red and brown algae branching off the tree early on, and the green algae commonly seen as ancestors to the land plants (starting at 3 o'clock in Figure 1.2 and continuing counter-clockwise, in brown, red and green colours, respectively). The first land plants were the bryophytes, consisting of the hornworts, liverworts and mosses (at around 1 o'clock, in dark blue and magenta). The decisive development for plants to be able to colonise the land was the development of a cuticle covering the surface, so they do not dry out when exposed to air. The next major step in plant evolution was the emergence of vascular plants, which have a vascular system conducting water and nutrients, and which now contain lignin, giving the tissue enough stability to grow higher. The first vascular plants are lycophytes or club mosses (depicted in khaki). Subsequently, the ferns evolved, containing whisk ferns and horsetails (in dark red). They now reproduce via spores, with a gametophyte and dominant sporophyte phase. *Microsorum thailandicum*, one of the plants investigated in this thesis, is a fern, as detailed below in Figure 1.3. The next major developmental step is the development of seeds (following reproduction by spores), with the gymnosperms containing conifers and ginkos, among others (in orange). Finally, angiosperms evolved (at around 11 o'clock), the flowering plants, which further diverged into monocots (to 9 o'clock, in blue, including *Pollia condensata*) and eudicots (including *Margaritaria nobilis*) [121].

Figure 1.3 shows a phylogenetic tree depicting all plants with structural colour from a helicoidal cellulose architecture [60] characterised so far. *Microsorum thailandicum* (discussed in detail in chapter 2 and chapter 3), *Lindsaea lucida* and *Diplazium tomentosum* are all in the order of Polypodiales. This order in turn is in the group of Polypodiopsida, or ferns, which also contains *Danaea nodosa*. So all of these belong to the clade of ferns, depicted in dark red between 12 and 1 o'clock in Figure 1.2 [12, 130, 49, 50].

Next, both *Pollia condensata* and *Mapania caudata* are monocots, shown in Figure 1.2 from 11 to 9 o'clock, in blue [147, 131]. *Pollia condensata* is shown in Figure 1.4 (A-B), it is the only known species with a right-handed helicoid in some of its cells, and *Mapania caudata* is shown in Figure 1.4 (C-D).



Official Publication of the Botanical Society of America, Inc.
www.wileyonlinelibrary.com/journal/AJB

Fig. 1.2 Phylogeny of plants based on plastid phylogenomic analysis, with images of important or well-known members of respective taxa. Reproduced with permission from [45].

Finally, the only eudicot characterised so far is *Margaritaria nobilis*, shown in Figure 1.4 E, which belongs to the order of Malpighiales, depicted in turquoise at 7 o'clock in Figure 1.2 [145].

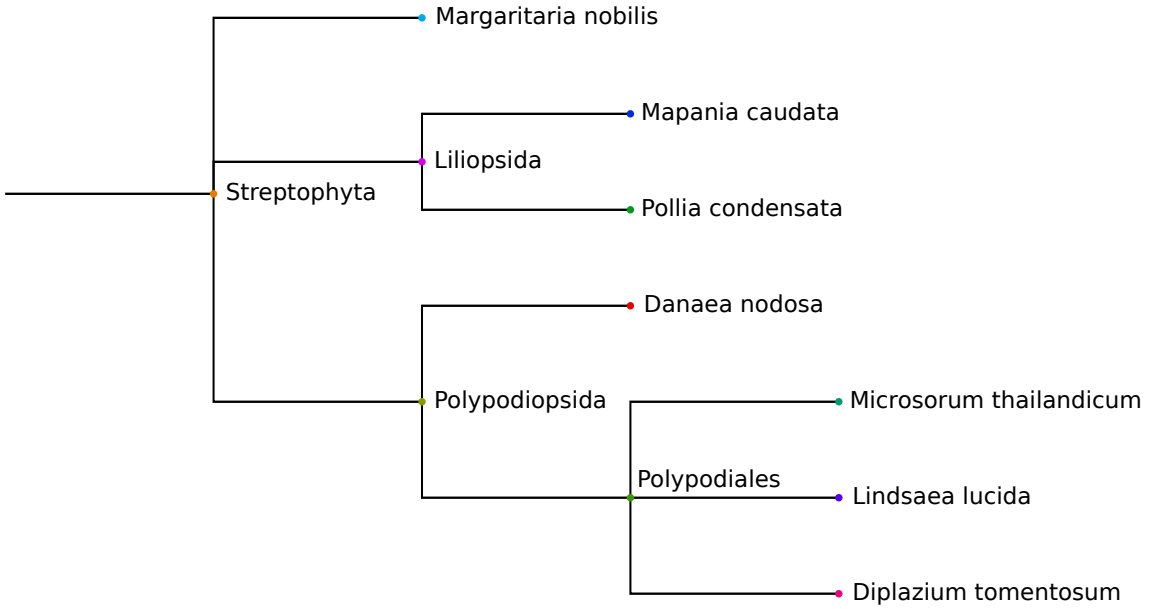


Fig. 1.3 Phylogenetic tree of plants exhibiting structural colour generated by a helicoidal cellulose architecture in their cell walls.

The phylogenetic tree was generated with the NCBI taxonomy browser, synonyms of plant names were searched for on the World Check List of Selected Plant Families from Kew Gardens (*Mapania cuspidata* var. pumila (Uittien) Uittien, J. Arnold Arbor. 20: 214 (1939) for *Mapania caudata*), and the tree was drawn on icytree [57, 43, 142, 143].

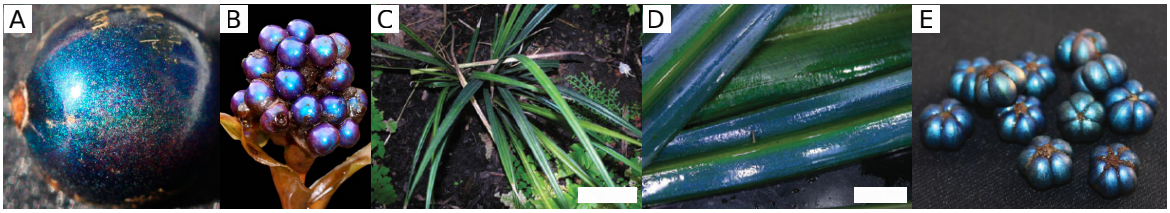


Fig. 1.4 Photos of (A-B) *Pollia condensata*, single fruit and infructescence. Fruit diameter is about 5 mm. From [147]. (C-D) *Mapania caudata*. Scale bars 5 cm, 1 cm. From [131]. (E) *Margaritaria nobilis*. Fruit diameter is about 5 mm. Reproduced with permission from [145].

1.3 Plant cell wall

The helicoidal architecture introduced above is found in the cell wall of the plants investigated in this thesis. Hence, in order to unravel this helicoidal arrangement of cellulose microfibrils, it is essential to study the cell wall in detail. Generally, plant cell walls are a highly complex and dynamic composite balancing multiple functions. In primary cell walls (PCW), they need to withstand turgor pressure from inside the cell, pressure from outside, all while still expanding in size in a highly controlled way. Secondary cell walls (SCW) are deposited once the cell has ceased growth, and provide increased mechanical strength. An interesting aspect is that while in the PCW, cellulose microfibrils orientation is dispersed and irregular, it is often highly organised in the SCW [28]. An example is wood fibre cells, with the highly oriented cellulose arrangement and distinct microfibril angle in the S_1, S_2 and S_3 layers of the SCW, and flax fibres [8, 7].

Figure 1.5 depicts a scheme of the PCW of *Arabidopsis thaliana*. PCW consist of cellulose microfibrils (in light blue), hemicellulose (xyloglucan in dark blue) and pectin (RGI in yellow, HG and RGII in red), which all interact with each other. This cartoon gives a good impression of the three-dimensional complexity of the PCW.

Another model of the PCW is depicted in Figure 1.6, based on the latest findings from atomic force microscopy (AFM), scanning electron microscopy (SEM), solid state nuclear magnetic resonance (ssNMR) spectroscopy and mechanics studies. The importance of the shape and different surfaces of cellulose microfibrils (hydrophobic and hydrophilic, see below), is considered by the blue (hydrophobic) and orange (hydrophilic) colour. Xyloglucan is depicted in green, both in coiled, solvated conformation, and in extended conformation bound to the cellulose microfibrils, where it can also be trapped in between. Finally, pectin is depicted in yellow, as a coiled space filler and as binding to the hydrophilic surfaces of cellulose. Cellulose microfibrils can bundle directly or via areas where xyloglucan is entrapped between them. These spots are highlighted by red arrows and are considered as biomechanical hotspots, where wall-loosening enzymes for PCW extension act [26].

1.3.1 Main components of the plant cell wall

The main components of the plant cell wall are cellulose, hemicellulose, pectin and lignin. Most likely, the helicoidal structure in the secondary cell wall will be based on the interactions of one or more of these components, that are for this reason briefly described here.

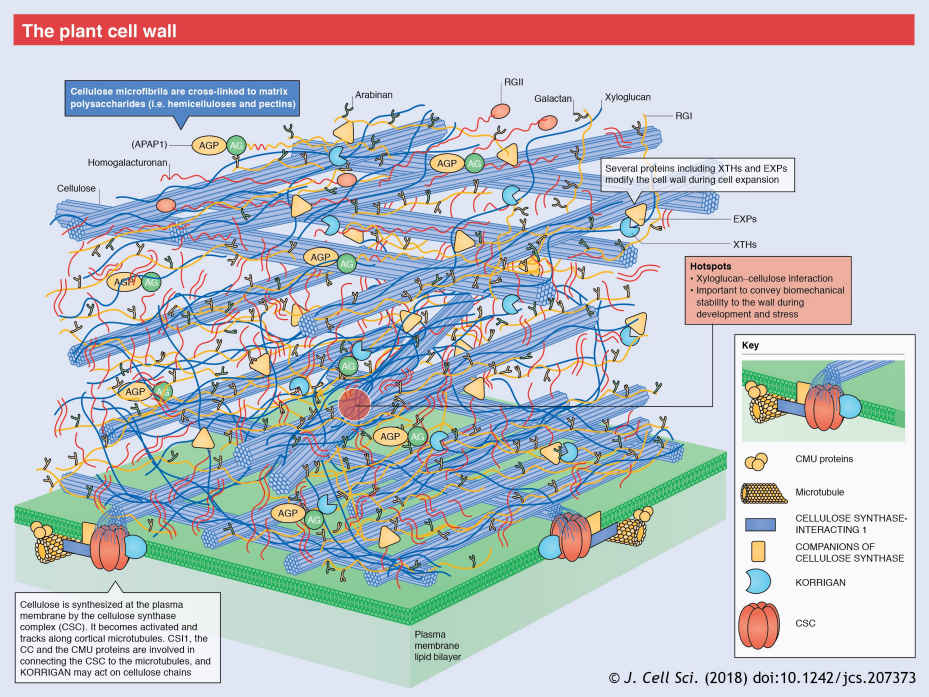


Fig. 1.5 Scheme of the primary cell wall of *Arabidopsis thaliana*. Dark blue XG: xyloglucan, a type of hemicellulose, and yellow RG I: rhamnogalacturonan I, red RG II: rhamnogalacturonan II, and red HG: homogalacturonan, all types of pectin. Reproduced with permission from [73].

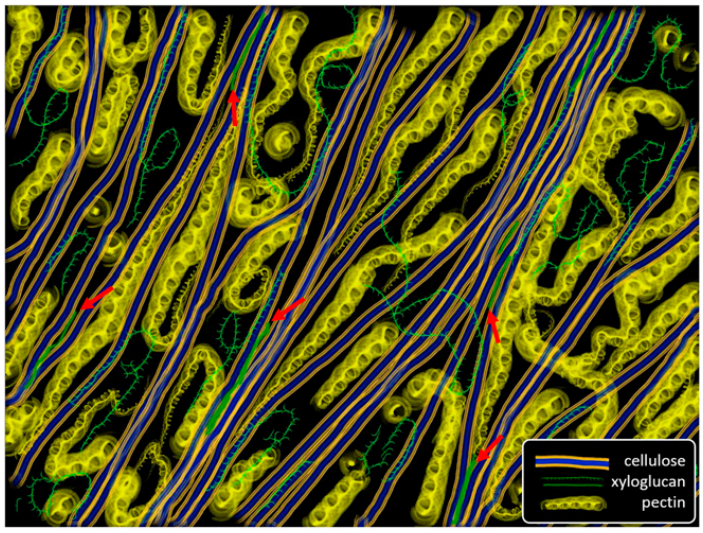


Fig. 1.6 Scheme of the primary cell wall, cellulose microfibrils in blue (hydrophobic) and orange (hydrophilic), xyloglucan in green, pectin in yellow, and biomechanical hotspots are highlighted by red arrows. Reproduced with permission from [26].

Cellulose

Figure 1.7 shows the hierarchical organisation of cellulose from the molecular to the fibrous scale. The monomeric repeating unit of glucose is depicted on the far left, it consists of D-glucose in pyranose ring form, linked via a β-1,4-glycosidic bond, giving rise to a linear chain [41].

These molecular chains come together in crystalline microfibrils, which are depicted in dark blue colour. In higher plants, the predominant type is the Iβ allomorph [98]. It is still under debate how many glucose chains a microfibril consists of [102]. A widespread hypothesis used to be 36 glucose chains [54], but now there is more consensus about 18, which would give them a diameter of approximately 3 nm [100], and also the option of 24 chains is still debated [38]. The inset shows possible cross-sections of cellulose microfibrils, of spruce wood and of a hypothetical 18-chain model, based on small-angle neutron scattering (SANS), and wide-angle X-ray scattering (WAXS) studies [38]. The shape of the cross-sections is not yet known either and might vary by plant and tissue [59]. Subsequently, after biosynthesis by the cellulose-synthesising complexes (CSC) (see below), these microfibrils can aggregate (lateral contact without crystalline continuity). For some sources, larger microfibril sizes are observed. One hypothesis is that they can fuse after biosynthesis, forming a single crystalline unit, leading to a much bigger crystallite size and bigger microfibrils. At the same time it has not been ruled out that they are biosynthesised with larger crystallite size, possibly by a different type of CSC.

Crystallinity is another topic still under debate. For some types of cellulose microfibrils, there might be some regions along the microfibril being crystalline and some regions being amorphous [97]. However, this phenomenon is not observed for many other types of microfibrils, and the exact distribution and location of more or less crystalline areas remains unresolved. Probably, these more disordered regions are located towards ends of glucose chains, or in areas of mechanical damage.

Furthermore, the microfibrils can have different chain conformations, depending on several factors, such as the C-6 conformation for surface chains, resulting in different hydrogen bonding networks. Traditionally, these have been described as cellulose Iα, Iβ, para-crystalline, accessible and inaccessible fibril surface [75, 155, 74], whereas in more recent studies seven different polymorphs, two on the surface and five in the interior, within the same microfibril, were confirmed [152].

Especially in SCW (like the cell wall in *M. nobilis*), microfibrils with diameters larger than 3 nm were observed by SANS and WAXS [38, 136]. Hypotheses to explain these

discrepancies are the interpolation of a xylan chain, the fusion of the microfibrils over part of their length, or the interpolation of bound water. Since xylan in the two-fold helical screw configuration (see below) is very abundant in the SCW (of *Arabidopsis thaliana*), bound xylan seems to be the most likely explanation, leading to fibre bundles as depicted on the far right in Figure 1.7. This effect, the association of microfibrils with matrix polymers, which then aggregate into fibre bundles, has been observed e.g. for conifers.

As with all the concepts explained in section 1.3, all of these points are subject to variation not only within a plant but also between specific tissues in the same plant and between different specimens of the same plant.

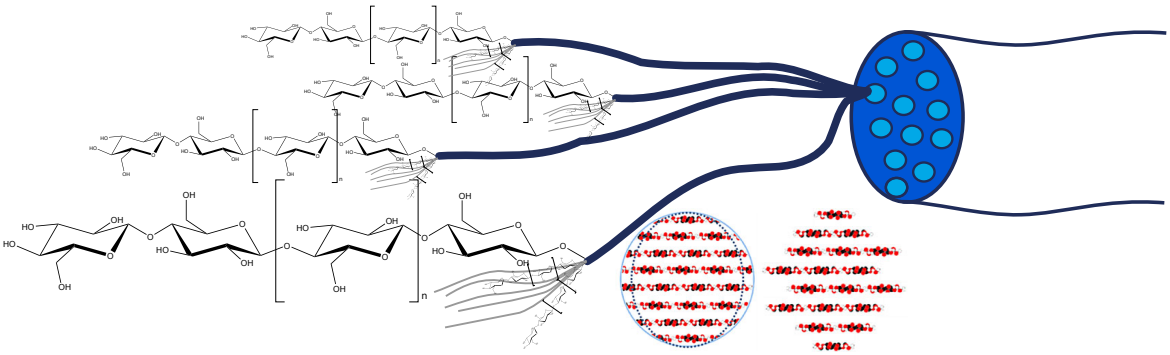


Fig. 1.7 Hierarchical organisation of cellulose, from the molecular scale to fibre bundles: molecular structure of glucose chains, which assemble into cellulose microfibrils, which can in turn aggregate into fibre bundles in a composite with hemicellulose. The insets are tentative cross-sections of microfibrils from spruce wood (left) and a hypothetical microfibril of 18 glucose chains (right). Insets reproduced with permission from [59].

Hemicelluloses

Hemicelluloses are a large group of heteropolysaccharides found in the cell wall, and their definition and what it comprises is still not entirely agreed on. One feature to classify them is the shared equatorial β-(1→4)-linked backbone structure. This definition contains xyloglucans, xylans, mannans and glucomannans, and β-(1→3,1→4)-glucans. An overview of different hemicelluloses is given in Figure 1.8 [61]. Other heteropolysaccharides are galactans, arabinans, and arabinogalactans, but they do not share the same type of glycosidic linkage, and might be included in the group of pectins instead, due to similarities in their initial biosynthesis [118].

Xylan is the main hemicellulose in the SCW of eudicots and grasses, which will be discussed in more detail. Apart from the β-(1→4)-xylose backbone common to all xylans, many different features exist. For example, the backbone can be substituted with α-(1→2)-linked glucuronosyl and 4-O-methyl glucuronosyl residues, in this case it is named glucuronoxylan, being the main hemicellulose in eudicot SCW (like *M. nobilis* endocarp [145]). Furthermore, arabinose residues can be attached to the backbone, being common in both the PCW and SCW of grasses, and being classified as arabinoxylans and glucuronoarabinoxylans. Moreover, many xylans are acetylated to varying degrees, especially in eudicot SCW. Another curious feature is a conserved oligosaccharide domain at the reducing end, a β-D-Xyl-(1→4)-β-D-Xyl-(1→3)-α-L-Rha-(1→2)-α-D-GalA-(1→4)-D-Xyl structure (Xyl: xylose, Rha: rhamnose, GalA: galacturonic acid, all in pyranose configuration) [118].

Xylose is very similar to glucose, it is usually found in pyranose conformation, all hydroxy groups being in equatorial configuration, xylose only missing the C-6. And even though xylan is found in a three-fold helical screw in solution or recrystallised [18, 95], there is now evidence that it can adopt a two-fold helical screw with the same dimensions as cellulose. There is thus great potential for the two polymers to interact, as discussed below [122, 36].

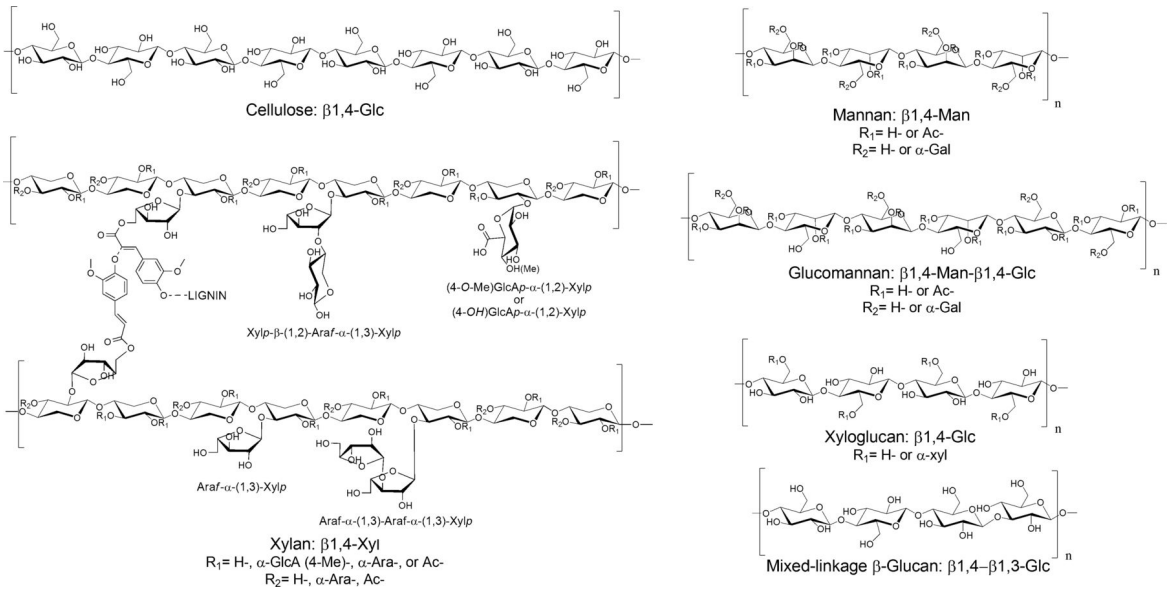


Fig. 1.8 Possible molecular structures of xyloglucans, xylans, mannans and glucomannans, and molecular structure of cellulose. From [61].

Pectin

Pectin is a major component in the PCW, accounting for up to 35 %. Furthermore, it is deposited in the middle lamellae and cell corners, where it plays an important role in cell-cell adhesion. Finally, it is also found in the junction zone between cells with secondary walls including xylem and fibre cells in woody tissue, albeit in smaller amounts. Similar to the hemicelluloses, it is biosynthesised in the Golgi. It fulfills various roles in the plant cell wall and beyond, including most importantly cell-cell adhesion via, for example, RG-II dimerization and HG cross-linking. Moreover, it is found to provide structural support in soft tissues in PCW, it might influence secondary wall formation in fibres and woody tissues, and acts as a hydration polymer affecting wall rheology, among others.

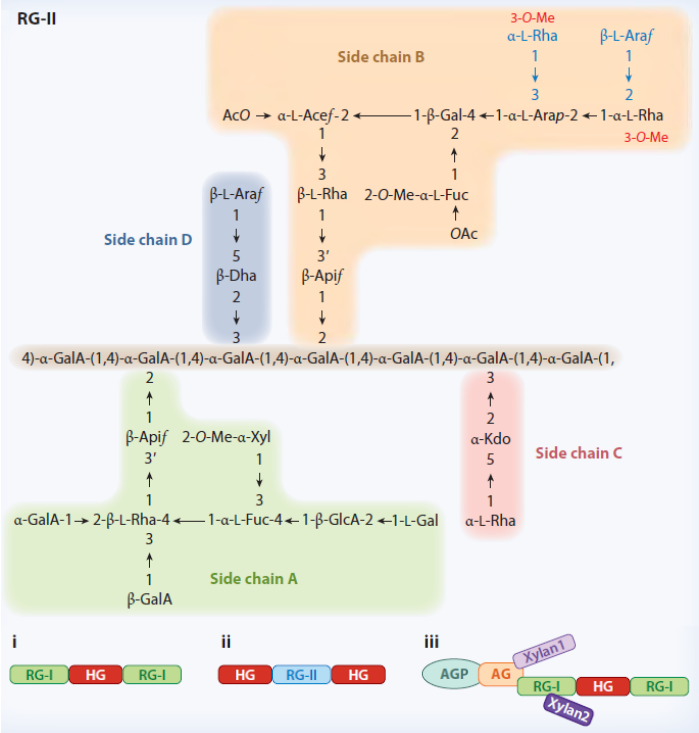


Fig. 1.9 Overview of the structure of RG-II (top), and pectin interconnections (bottom). Ace*f*: aceric acid, Api*f*: apiofuranose, Ara*f*: arabinofuranose, Ara*p*: arabinopyranose, Dha: 3-deoxy-D-lyxo-2-heptulosaric acid, Fuc: fucose, Gal: galactose, GalA: galacturonic acid, GlcA: glucuronic acid, Kdo: 3-deoxy-D-manno-octulosonic acid, Rha: rhamnose, Xyl: xylose. Abbreviations for nonpectin portions of APAP1: AG: arabinogalactan, AGP: arabinogalactan protein, Xylan1 and Xylan2: hemicellulosic glycan domains. Reproduced with permission from [5].

There is considerable structural variation, and the full structure and length has yet to be investigated, intact extraction of pectin for structural elucidation remains impossible.

Generally, pectin is divided into homogalacturonan (HG), rhamnogalacturonan I (RG-I) and rhamnogalacturonan II (RG-II). HG and RG-II both have a backbone of galacturonic acid, and all glycosidic linkages in the backbone are α-1,4-linkages, while the backbone of RG-I consists of alternating galacturonic acid and rhamnose residues linked via [,4)-α-GalA-(1,2)-α-L-Rha-(1,4)-α-GalA-(1,2)-α-L-Rha-(1,]$_n$ glycosidic bonds.

For HG, some modifications are found, like methylesterifications of the carboxylic acid groups, O-acetylation of the O-2 or O-3, some xylose oligosaccharide branches attached to the GalA backbone, etc. The structure of RG-II is shown in Figure 1.9, the occurence of the different side chains A-D varies with species. RG-I can also have different modifications like acetylation, and it can have side chains branching of the rhamnose residues, for example containing arabinose or galactose residues of various length. Finally, there is growing evidence of interconnections between different types of pectin, and to proteins and hemicellulose, as depicted in Figure 1.9 bottom [86, 5, 117]. Furthermore, close interactions of pectin with cellulose have been observed in the PCW of *A. thaliana* via NMR investigations [151].

Pectin is not expected in major amounts in the helicoidal SCW of *M. nobilis*, but there will still be small amounts found in the PCW and middle lamellae.

Lignin

Lignin is not a polysaccharide, but a polyphenol, and there is tremendous heterogeneity in its structure across different plants and tissues [10, 11, 140, 89]. The production of monolignols is tightly controlled biochemically, with the most abundant monomers being p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol. They are produced in the plastid and cytosol, and three hypotheses about their transport to the cell wall exist: passive diffusion (the concentration is higher at the site of biosynthesis, and the equilibrium is continuously shifted by polymerisation), active transport, or vesicle-mediated exocytosis. Additionally, peroxidases and laccases are placed strategically in the cell wall, and will then generate the monolignol radicals, which are stabilised by electron delocalisation. Polymerisation is only chemically, but not biochemically controlled, hence the huge variation in the structure of lignins. It consists mainly of dimerisation and lignification, and β-O-4-aryl-ethers, β-5, β-β, 5-5, 5-O-4, and β-5 linkages are the most common linkages connecting the monolignols to form the lignin polymer, depicted in Figure 1.10 A [10].

Lignification usually takes place late in the formation of the SCW, starting from the cell corners of the PCW and middle lamellae. It can still continue after programmed cell death, and different types of cells can either only use their own monolignols and lignification would arrest after all monolignol is used up, or cells can also accept monolignols from other cells after their own death, or cells can also donate monolignols. The entire process is depicted

in the scheme in Figure 1.10 B, from [11]. Finally, lignin biosynthesis is very flexible with respect to accepting changes in monomers and polymer composition with little or no obvious phenotypic effects. At the same time, conservation of composition and spatial and temporal control indicate that regulation does have advantages.

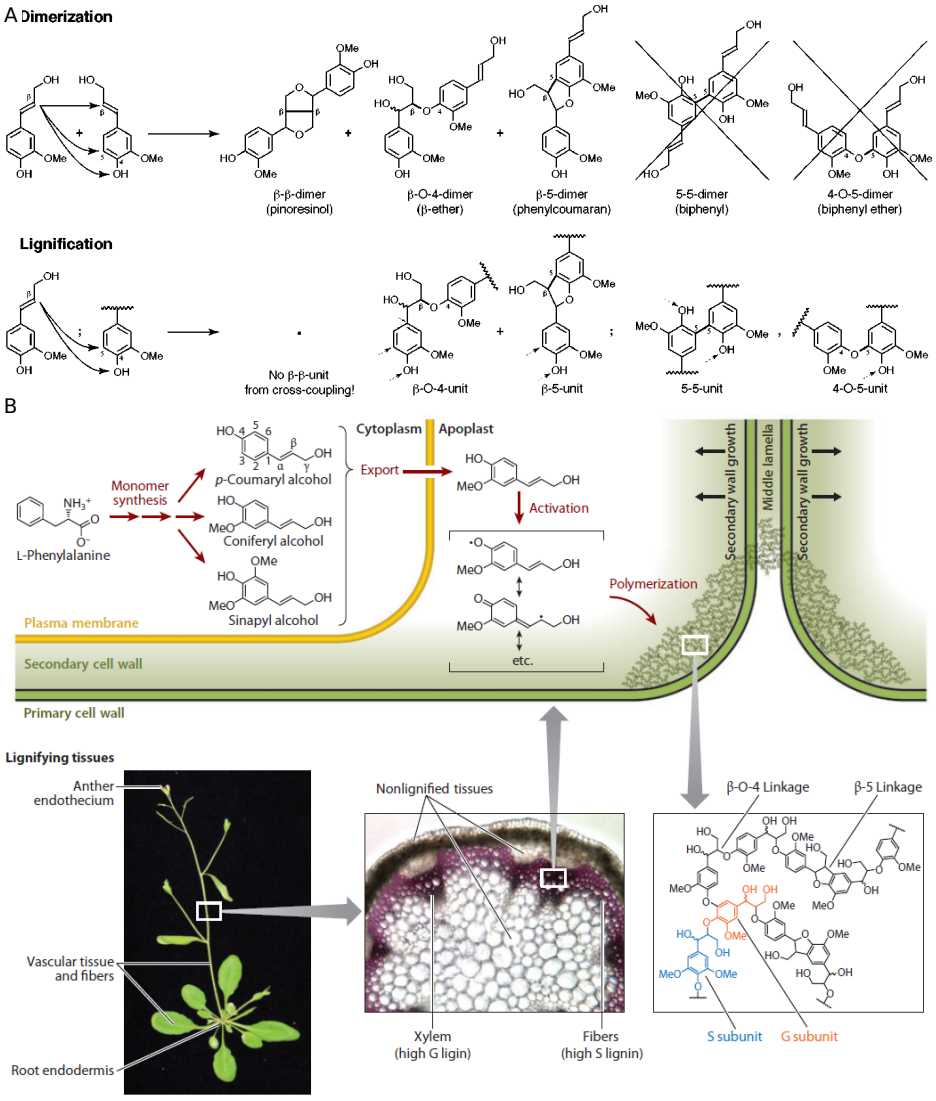


Fig. 1.10 Scheme showing (A) the most common lignin linkages. Reproduced with permission from [10]. (B) the biosynthesis of lignin: the three main monolignols p-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol, derived from phenylalanine, are activated and polymerised, starting from the middle lamellae and primary cell wall corners, and with different subunits and linkages depicted. Reproduced with permission from [11].

Since lignin biosynthesis and incorporation into the wall only happens towards the end of the SCW being laid down, it is most likely that in the endocarp of *M. nobilis*, the helicoidal

structure is already implemented before lignification. It is thus not expected that lignin contributes to the twisting arrangement of cellulose and hemicellulose.

Then, there is still the option that lignin could play a role in adjusting the helicoidal pitch p to the right length to give rise to reflection in the visible range. This hypothesis could be investigated by looking at fruit development. For its lack of influence in the helicoidal arrangement, and its complexity and major difference from cellulose and hemicellulose, lignin was not investigated further in this thesis.

1.3.2 Biosynthesis of the plant cell wall

To address the question 'What are the main building blocks of this helicoidal architecture, and what requirements of properties are there?', it is essential to know the basics of plant cell wall biosynthesis. They will also be required for the discussion, section 5.7, which concludes this thesis, and are hence briefly presented here.

The biosynthesis of the cell wall is a highly complex and controlled process, with most research being focused on the primary cell wall [25, 66]. Cell walls are generally organised in layers, which are laid down sequentially during their biosynthesis, as sketched in Figure 1.11 [2].

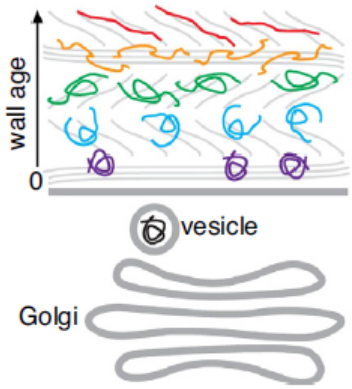


Fig. 1.11 Cartoon sketching the sequential assembly of layers in the plant cell wall, with grey cellulose microfibrils, and matrix polysaccharides in colour, with the most recently deposited in purple and the oldest in red. Reproduced with permission from [2].

The cellulose microfibrils are biosynthesised directly at the plasma membrane, by cellulose-synthesising complexes (CSC) consisting of six subunits or lobes arranged in a hexameric array, containing a still undetermined number of cellulose synthase (CESA) proteins. Hypotheses include 3 CESA per lobe, which would lead to cellulose microfibrils

consisting of 18 glucose chains, 4 CESA per lobe, leading to 24 glucose chains per microfibril, or 6, leading to 36, as discussed in subsection 1.3.1. The cellulose microfibrils are directly extruded from the plasma membrane into the cell wall, as depicted in purple in Figure 1.12.

Hemicellulose (and pectin, since this is a model for the PCW), on the other hand, is synthesised in the Golgi apparatus and transported to the cell wall by vesicles, depicted in blue and grey. (Pectins are depicted in red, orange and green. In reality, hemicelluloses and pectins are distributed everywhere and mixed, they are only separated in the scheme for simplicity.)

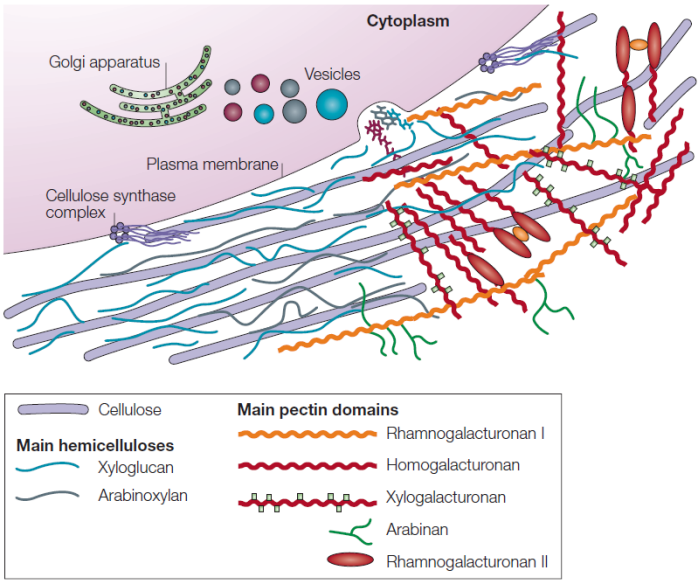


Fig. 1.12 Scheme depicting the biosynthesis of the primary cell wall. Cellulose microfibrils are extruded by CSC directly into the wall, while hemicellulose is synthesised in the Golgi and transported to the wall by vesicles. Reproduced with permission from [25].

During cellulose microfibril biosynthesis, the CSC move along cortical microtubules (CMT), with approximately 200–300 nm/min velocity for the PCW. In the PCW, the CSC are more widely spaced apart, their density on the plasma membrane is lower (approximately 0.3 to 1.3 particles/μm^2), while they seem to occur more densely placed and in clusters in the SCW plasma membrane [2].

For the PCW biosynthesis, a small amount of CSC is delivered to the plasma membrane (PM), and they are evenly distributed, well spaced apart and uncoordinated. Furthermore, the CSC can move in either direction along the CMT, causing bidirectional cellulose microfibril extrusion, see Figure 1.13 A and B. For the SCW biosynthesis, on the other hand, a higher amount of CSC are transported to the PM, and they thus coordinate and move along the CMT

in the same direction. Due to this coordinated extrusion of cellulose microfibrils, they can also aggregate into larger fibre bundles, as depicted in Figure 1.13 C and D [78].

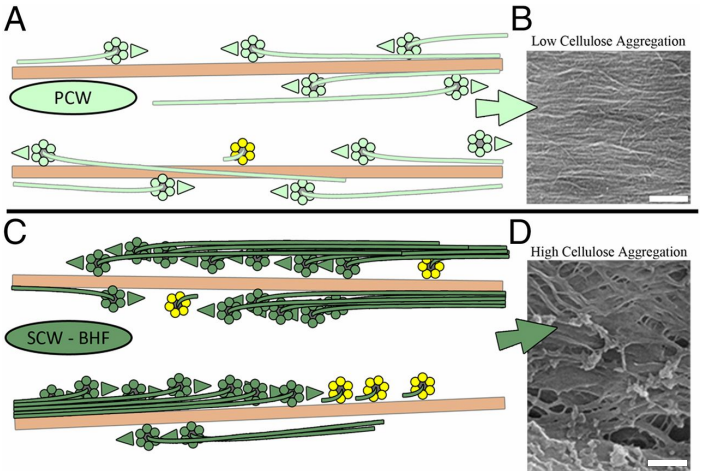


Fig. 1.13 Scheme and SEM images of cellulose biosynthesis for the primary (A,B) and secondary (C,D) cell wall. Active CSC in green, travelling along the CMT in brown. Recently delivered CSC in yellow, more frequent for the SCW. Scale bars: 200 nm. From [78].

Additionally to the rearrangement and bundling of CMT during the transition from PCW to SCW biosynthesis, many other adaptations are made when the cell has ceased its growth and transitions to build a SCW. An overview is given in Figure 1.14 [84].

Firstly, different CESA make up the CSC between the PCW and the SCW, therefore enzyme production and assembly into the CSC has to be altered. Secondly, in the Golgi apparatus, pectin does not need to be produced anymore, since it is not incorporated into the SCW, only into the PCW. Furthermore, different types of hemicelluloses are now produced: while e.g. xyloglucan and arabinoxylans dominate the PCW, the SCW contains e.g. xylans and mannans instead. Thirdly, as preparation for lignification, monolignols and the necessary oxidative enzymes, laccases and peroxidases, are produced and secreted. For cells that are dead at maturity (like in the *M. nobilis* endocarp), programmed cell death eventually occurs, after which lignfication can still continue, if there are still monolignols and the necessary oxidative enzymes, or if they are donated by neighbouring cells.

Moreover, the composition of PCW to SCW is different (with generally large variation between species and tissues). PCW usually contain 15-40 % cellulose, 30-50 % pectin, and 20-30 % xyloglucans, as well as smaller amounts of arabinoxylans and structural proteins (on a dry weight basis) [28]. SCW are usually comprised of up to 60 % cellulose, 10-40

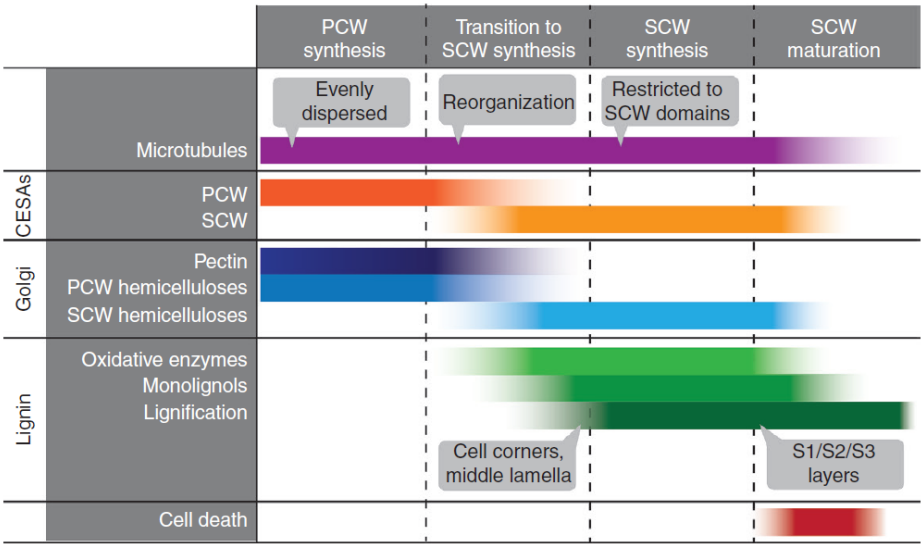


Fig. 1.14 Overview of biosynthetic processes during PCW biosynthesis, transition to SCW biosynthesis, and finally programmed cell death. Reproduced with permission from [84].

% hemicellulose, and about 20-30 % lignin [84]. At this point, it is also worth mentioning that the development of the SCW enabled land plants to grow tall, thanks to its mechanical stability.

How the different types of polysaccharides assemble into the plant cell wall is a current topic of research [27, 26, 148].

Figure 1.15 shows how cellulose microfibrils and hemicelluloses could interact in the PCW. Two different possible cross-sections of cellulose microfibrils are depicted in (b), and the surfaces of the cellulose microfibril exposing the sugar rings are described as the hydrophobic surfaces (in blue), while the surfaces exposing the equatorial hydroxy groups are described as hydrophilic (in red). It is hypothesised that xyloglucans interact well with the hydrophobic surfaces, while xylan interacts with the hydrophilic surfaces (and the hydrophobic surfaces as well, see below). (a) and (c) summarise findings from recent in-depth NMR studies [152], where seven different cellulose microfibril polymorphs were distinguished, comprising surface and interior, and more and less ordered environments (as explained in subsection 1.3.1). Finally, (d) shows hyphothetical models of how bundles of cellulose microfibrils could be arranged, and how they could be embedded in a hemicellulose matrix. They could bundle via their hydrophobic surfaces, as shown, but also via their hydrophilic surfaces, and with or without xyloglucan in-between.

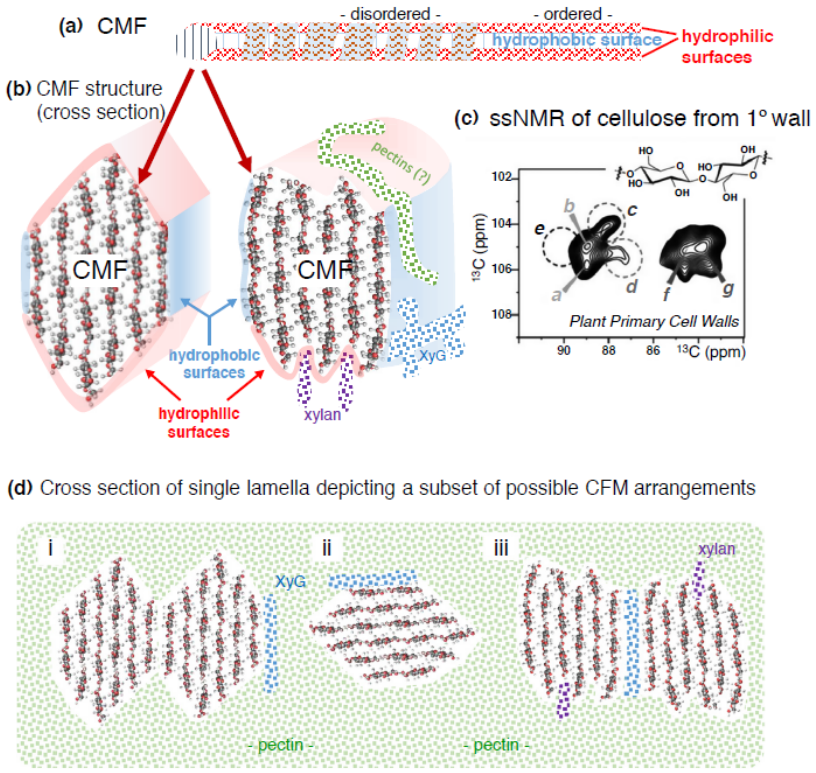


Fig. 1.15 Hypothetical cross-sections of cellulose microfibrils and their possible bundling and interactions with xyloglucan, xylan and pectin, depending on hydrophobic and hydrophilic surfaces. Reproduced with permission from [27].

More details about the possible interactions between cellulose and xylan are depicted in Figure 1.16 [19]. Xylan can adopt a two-fold helical screw in the plant cell wall, interacting tightly with cellulose microfibrils [122, 36]. (In solution or recrystallised, it adopts a three-fold helical screw configuration.) Furthermore, in the SCW of *A. thaliana*, distinct domains of xylan have been identified: the compatible and incompatible domain [18]. While the incompatible domain exhibits random decorations and acetylation, the acetylation on the compabtible domain is found on every other xylose unit of the backbone, and also decorations like glucuronic acid are all found with even spacing. Hence, in a two-fold helical screw, all decorations will point to the same side of the chain, and there will be no steric hindrance on the other side, enabling intimate interaction with the hydrophilic surface of the cellulose microfibrils. Irrespective of decorations, all xylan domains will also be able to interact with the hydrophobic surface, allowing xylan to cover the entire surface of cellulose microfibrils [19]. Many different interaction hypotheses are depicted in the scheme in Figure 1.16.

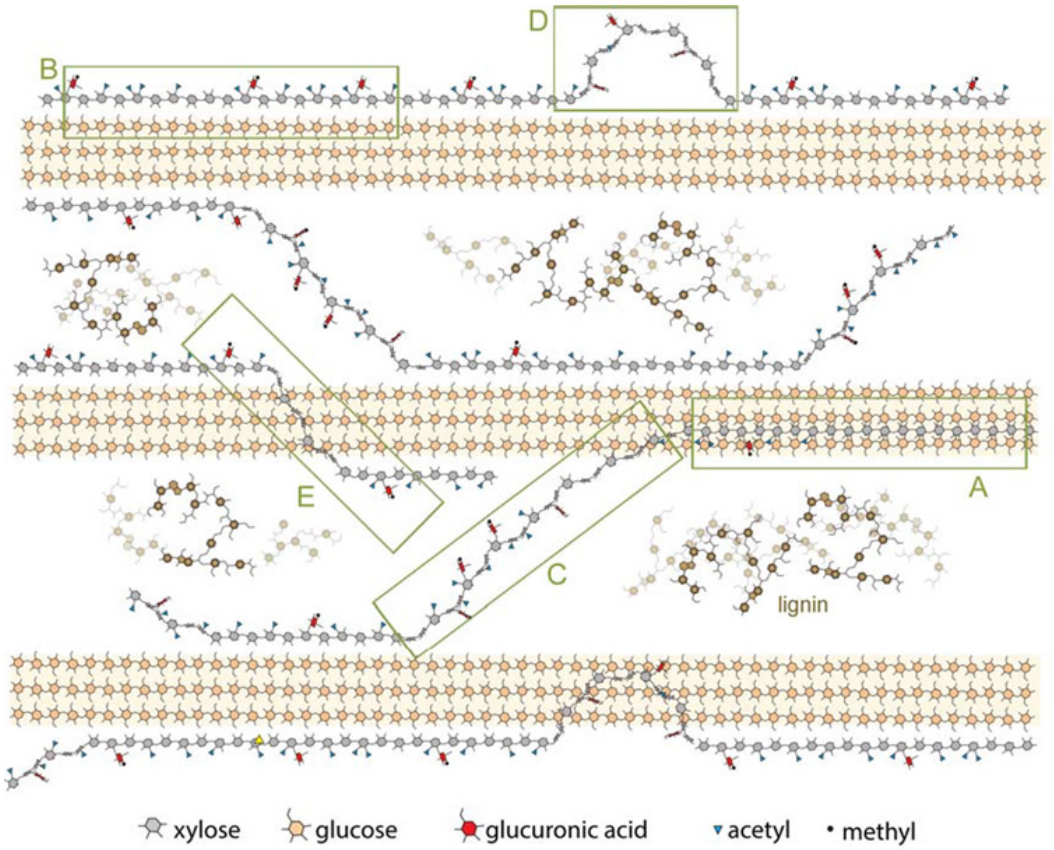


Fig. 1.16 Scheme showing possible interactions of the compatible and incompatible xylan domain with the cellulose microfibril: (A) interaction of either domain with the hydrophobic surface. (B) Interaction of the compatible domain with the hydrophilic surface. (C-E) The incompatible domain can span the distance between adjacent cellulose microfibrils (C), detach and form loops into the matrix and back onto the fibril (D) and span the hydrophobic surface before docking on the opposite hydrophilic surface again (E). From [19]

Furthermore, more findings about how lignin interacts with the polysaccharides and how it is integrated in the SCW emerge. Via in-depth NMR studies, it was observed that lignin is in close physical contact with xylan by electrostatic contact, but not as much with cellulose, see Figure 1.17 top. Just like cellulose bundles, it is hydrophobic and does not hold much water, while three-fold xylan is considered as the hydrated matrix. Generally, the SCW is much less hydrated than the PCW, only about 40 % of the volume are filled by water [28]. Another option of interaction of lignin with hemicellulose is covalent bonding in so-called lignin-carbohydrate complexes (LCC). Two mechanisms are proposed, radical coupling of a monolignol to ferulated xylan, and re-aromatisation of the quinone methide intermediate by hemicellulose nucleophiles. While the former mechanism is only relevant for commelinids (a clade within the monocots, comprising e.g. Poales including grasses and sedges) which have

ferulic acid substitution, the later is also possible in the SCW investigated here. Depending on whether a hydroxy group or a carboxylic acid group acts as a nucleophile, an ether or ester linkage is formed, respectively. Possible structures include phenyl glycosides, γ-esters and benzyl ethers, as depicted in Figure 1.17 bottom.

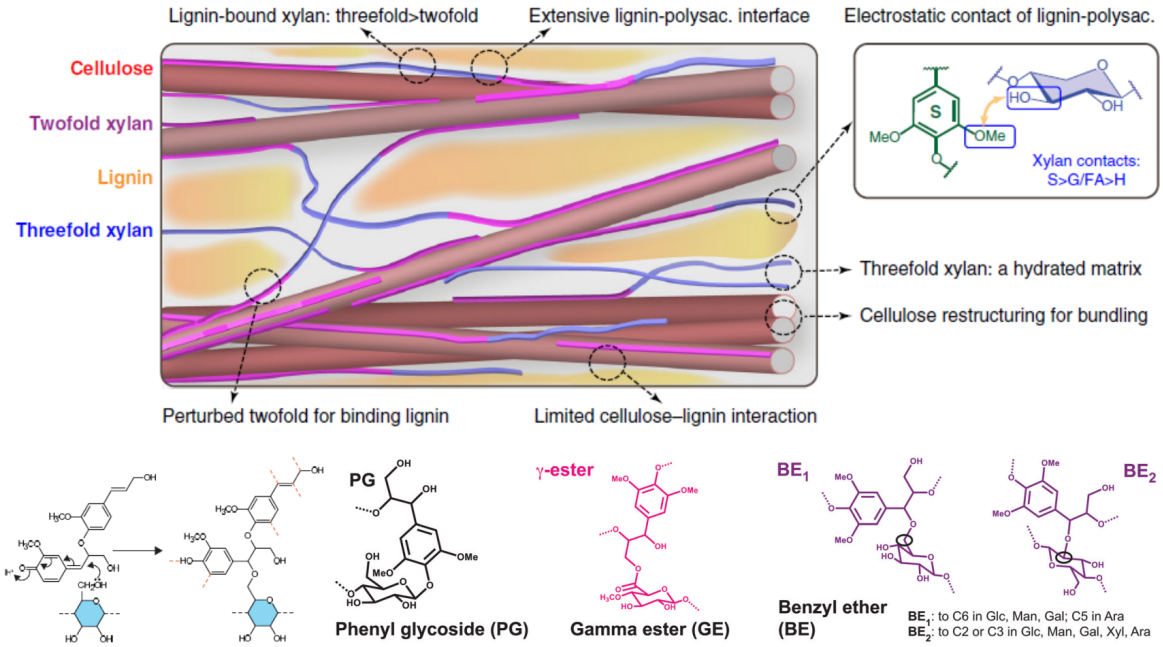


Fig. 1.17 Top: Scheme of lignin-polysaccharide packing and secondary cell wall architecture based on ssNMR. SCW is tighter packed than PCW, and xylan bridges the hydrophobic cellulose and lignin areas. Xylan adopts a twofold or threefold helical screw conformation. From [62]. Bottom left: Proposed mechanism of covalent bond formation between lignin and hemicellulose: Hemicellulose acts as a nucleophile, causing re-aromatisation of the quinone methide intermediate, forming a β-O-4 linkage. Further lignin polymerisation can occur at carbons marked by orange dotted lines. Reproduced with permission from [134]. Bottom right: Lignin-carbohydrate complexes have been identified as phenyl glycosides, gamma esters, and benzyl ethers. From [82].

1.3.3 Helicoidal assembly in the plant cell wall

The helicoidal arrangement of cellulose microfibrils is very common in many species. However, in most cases, the length of the pitch p along the helicoidal axis is not in the right length range to give rise to any optical effects. The biological significance of the helicoidal architecture could simply be circumstantial, be dictated for biosynthetic reasons, or because of improved mechanical properties.

Previous work focused on the assembly of cellulose microfibrils in a helicoidal architecture, and how it could possibly be compared to a cholesteric chiral nematic liquid crystal

system, like cellulose nanocrystal suspensions forming a chiral nematic phase in the right conditions. Quince seed mucilage is an example [113, 144]. More recently, the role of xylan was investigated with respect to this helicoidal arrangement, suggesting that the charges it can provide via its glucuronic acid decorations could play a role in mediating the cellulose microfibrils into the helicoidal arrangement, and also serving as a host structure during lignification, as shown in Figure 1.18 [112]. However these findings are quite sparse and speculative, leaving the formation of this helicoidal architecture far from being understood.

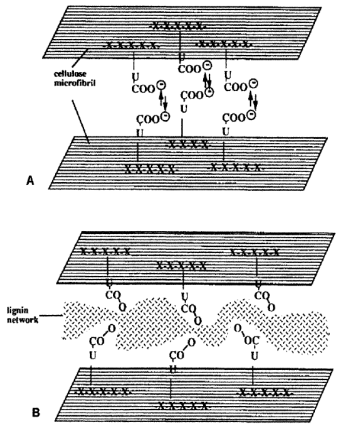


Fig. 1.18 Scheme illustrating the hypothesis of glucuronic acid charges from xylan positioning and stabilising the cellulose microfibrils in (A) and acting as a host for subsequent lignification in (B). Reproduced from Reis D. and Vian B. Helicoidal pattern in secondary cell walls and possible role of xylans in their construction. C. R. Biologies 2004; 327: 785-790. ©2004 Academie des sciences. Published by Elsevier SAS. All rights reserved [112].

1.4 Organisation of this thesis

In chapter 1 of this thesis, all the relevant background and literature were briefly introduced, comprising the concept of structural colour and various aspects about plant cell walls.

Then, chapter 2 comprises the introduction for the publication 'Structural colours in the frond of *Microsorum thailandicum*'. It was submitted to the journal Royal Society Interface Focus and it was accepted with minor revisions just before the thesis deadline. The fern *Microsorum thailandicum* is introduced in chapter 2. Furthermore, it puts the publication into context of the research field, and it outlines what my contributions were and what contributions were made by the co-authors.

As already mentioned, chapter 3 is the submitted publication, titled 'Structural colours in the frond of *Microsorum thailandicum*' [130]. I am the first author, and I mainly contributed to the paper, and co-authors are Yu Ogawa (who performed the anatomical investigations with electron microscopy), Villads Egede Johansen (who performed the simulations), and Clive Lundquist, Heather Whitney, and Silvia Vignolini (with whom I had discussed the results and structure of the manuscript). It is an in-depth study of the optical properties of the fern. The ultrastructure was elucidated and the helicoidal architecture of cellulose microfibrils was confirmed in the secondary cell wall of the first and upper side of the second layer of epidermal cells, both on the adaxial and the abaxial surface. The variation of the structure was characterised, both by optical and electron microscopy, and the disorder was investigated by modelling.

All the methods that were used for obtaining the results presented in the following chapter (that were not already included in chapter 3) are described in chapter 4. Standard techniques, like NMR and XRD spectroscopy, are only briefly listed, while more unusual or specialised techniques are explained in more detail. The NREL biomass quantification process was implemented in the scope of this thesis, and is thus explained in more detail. Furthermore, the isolation of cellulose microfibrils from the *M. nobilis* materials was adapted, and is thus explained in more detail as well. Finally, the analysis of xylan is very specialised and all the steps are explained thoroughly.

Finally, chapter 5 summarises all the results concerning *M. nobilis*. Firstly, it outlines the composition of both the leaves and the endocarp. Then it shows the isolation of cellulose microfibrils, again from both tissues, and all the chemical and morphological properties investigated, as well as assessing the microfibrils' crystallinity. Then, the molecular structure of xylan, again from both tissues, is elucidated. Briefly, SAXS analyses and coarse-grain modelling (carried out by Yu Ogawa) are summarised, and finally all results are discussed.

Chapter 2

Introduction to 'Structural colours in the frond of *Microsorum thailandicum*'

2.1 Background

Some species of *Microsorum* are known to exhibit bright blue coloration, thus we wanted to investigate whether we could identify any structurally coloured ones and characterise their optical response. In particular, *Microsorum thailandicum* [12], *Microsorum musifolium* [108] and *Microsorum whiteheadii* [129] were studied in the context of this thesis, the plants were provided by our collaborator Dr. Heather Whitney, University of Bristol, as shown in Figure 2.1 A-C. *Microsorum* is a big genus of ferns, and there is still ongoing debate about their classifications [109]. All species discussed here are in the *Microsorum punctatum* clade [70].

M. thailandicum looked very blue and shiny by eye, while the other two species were less distinctive. To understand the origin of this appearance, I carried out polarised optical microscopy on the fresh fronds of the live plants, and I confirmed that *M. thailandicum* is indeed structurally coloured, while the other two species, *M. musifolium* and *M. whiteheadii*, are not. *M. thailandicum* exhibits intense reflection only in the left-handed circular polarisation channel (LCP), while it is colourless in the right-handed circular polarisation channel (RCP), meaning that the structural coloration is caused by a helicoidal architecture.

Due to the strong reflection of *M. thailandicum*, we decided to investigate it further. (Reflection can be above unity when referenced to a white diffuser, if the reflection of the sample is more directional.) *M. thailandicum* was only first described in 2001 by Boonkerd and Nooteboom [12], see Figure 2.2. They reported the blue iridescence of its fronds (when alive, brown when dead), and that they grow in semi-shade, even though deep shade was

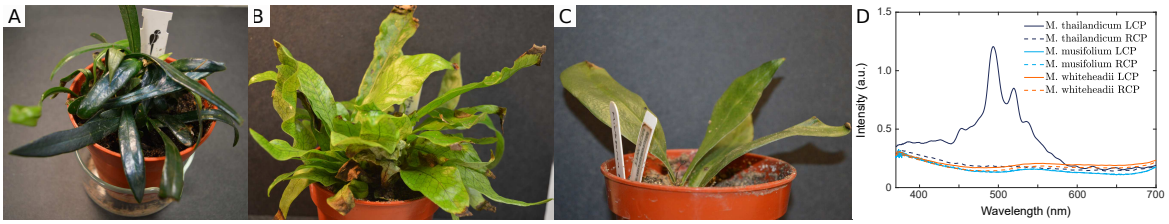


Fig. 2.1 Photos and spectra of three *Microsorum* species. Photo of (A) *M. thailandicum*, (B) *M. musifolium*, (C) *M. whiteheadii*, and (D) reflection spectra of left-handed and right-handed circular polarisation channel normalised to a white diffuser.

expected due to the coloration. The fern was introduced to the Sunday Market in Bangkok in early 1994, and is relatively easy to keep as a pot plant, thus having become very popular in Europe and the US.

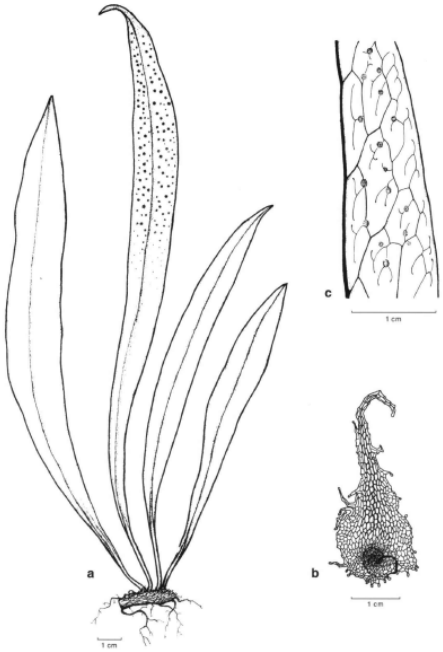


Fig. 2.2 Botanical sketch of *M. thailandicum*, from [12]. (A) Habit (characteristic form in which a given species of plant grows). (B) Rhizome (creeping rootstalk). (C) Detail of venation (pattern of frond veins).

In order to locate the area responsible for the optical appearance, I sectioned the fronds to test whether the cell wall of both the first and the second layer of the epidermal cells contribute to the structural coloration. I peeled the top layer off the frond with tweezers, quickly put water on both the peel and the main part of the frond to avoid dehydration, and I could indeed find LCP reflection in all cells of both the peeled-off layer, which is most likely the first layer of epidermal cells, and in the main part of the the frond, which most likely is

covered by the second layer of epidermal cells. When collecting POM images and spectra of the entire tissue, only signal from the top layer is collected, since the depth of focus of the objectives is small, but the second layer probably also contributes to the macroscopic appearance of the frond.

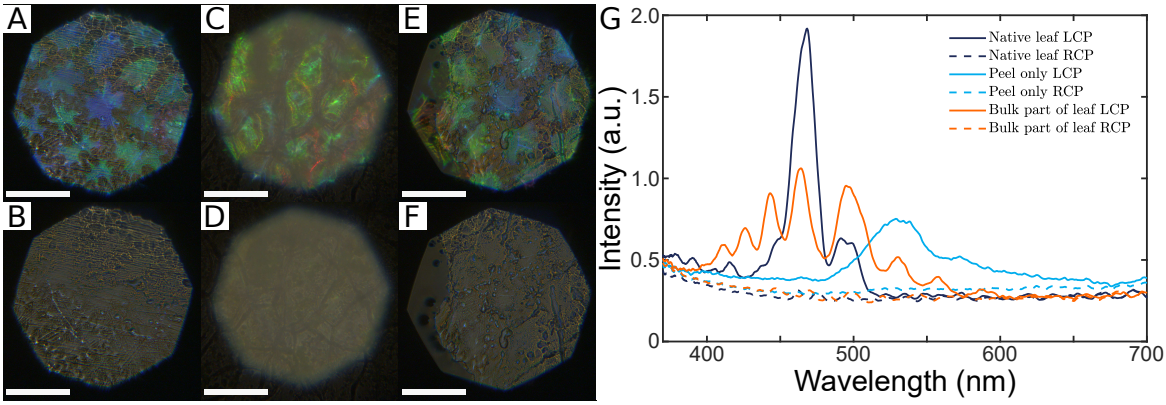


Fig. 2.3 Micrographs and spectra of native frond, the peel only, and the main part of the frond only. (A,B) Micrographs of the native frond, LCP in (A) and RCP in (B). (C,D) Micrographs of the peel only. (E,F) Micrographs of the main part of the frond only. Scale bar is 100 μm. (G) LCP and RCP spectra of A-F.

It is an unusual feature to have two layers of epidermal cells, as most leaves have only a single layer of epidermal cells [65]. In the frond of *M. thailandicum*, it could either be a multiple epidermis, or it could be a so-called hypodermis, which is a modified outer mesophyll layer. A multiple epidermis can be found in some plants, e.g. in *Peperomia*, succulents in the order of Piperales in the clade of Magnoliids (in the clade of basal angiosperms at 11 o'clock in green in Figure 1.2) [64]. Fronds or leaves with a hypodermis seem to be more common, at least among some ferns, and also in most palm trees [55, 138] (palms are monocots, depicted in blue at 10 o'clock in Figure 1.2). Furthermore, *Cyathea caracasana*, which is a tree fern, was found to have an epidermis and hypodermis that seem to be light-dependent: In ferns that grew in an open habitat with high sun exposure, both epidermis and hypodermis were much thicker than in ferns that grew in the shade under the canopy [4].

It is difficult to distinguish between a multiple epidermis and a hypodermis without studying the frond development, which was outside the scope of this project. For simplicity, the second layer of cells is referred to as epidermal cells in the publication, even though adding this explanation would be helpful.

Additionally to studying the frond development to understand whether the second layer of cells is a hypodermis or part of a multiple epidermis, it could be interesting to study

whether there are any differences in the helicoidal architecture in the first cell layer outer wall versus inner wall and the second cell layer outer wall. Potential differences could have significance for their biosynthesis, for example whether biosynthesis is differentiated in different areas of the wall or between the first and second cell layer. While it would be possible to investigate the different areas separately by electron microscopy, it is impossible optically (and also biochemically it would be very complicated). When taking spectra from the single cell via microphotospectrometry, only the outermost helicoidal wall is measured, since the depth of focus of the objective is very small. And while it is possible to section the frond, most likely between the first and second layer of cells, it is hard to know with certainty where the cut was made. Furthermore, the helicoidal structures would dry out fast when exposed to air, and while adding water to keep them hydrated for long enough means it can be shown qualitatively, the measured reflection wavelength would not be very meaningful - the helicoidal structure would shrink upon drying, or might expand when water is added. For these reasons, local differences were not investigated in this project.

2.2 Impact

In the following chapter, chapter 3, I present the manuscript that I submitted to the journal Royal Society Interface Focus, http://rsfs.royalsocietypublishing.org/. The manuscript is titled 'Structural colours in the frond of *Microsorum thailandicum*', and I submitted it as a Research Article for the '*Living Light Proceedings*' special issue in the journal Interface Focus. I carried out all the experimental work as part of my research degree into structural colour in plants.

The manuscript reports a thorough investigation of structural colours displayed by the fern *M. thailandicum*. Many measurements were carried out, polarised microphotospectrometry, electron microscopy, and numerical modelling, to understand the optical response and relate it to the ultrastructure of the fern. We confirmed that the helicoidal arrangement of cellulose microfibrils within the secondary cell walls of the epidermal cells gives rise to this structural colouration. It is the same type of photonic structure that will be discussed in more depth in chapter 5, for the fruit *Margaritaria nobilis*. The manuscript describes furthermore that whereas the adaxial cells have well-defined spectral reflection with little variation, the abaxial cells exhibit significantly more variation, spanning almost the entire visible range. These findings open many questions about the function of the reflection of the adaxial and abaxial epidermis, which is presented in the discussion section. In particular, it is interesting to find a system where the adaxial surface seems tightly controlled and the abaxial surface much

less so, which indicates different biosynthetic mechanisms and biological functions. To our knowledge, this phenomenon has not been reported before, and it is intriguing from a cell wall biosynthesis point of view.

Generally, the investigation of structural colour in plants is still underreported, and this work is therefore important for understanding and uncovering the diversity and biological functions of structural colours in plants.

2.3 Contributions

Detailed author contributions are stated in section 3.4.4. I was involved in planning and discussing all experiments and all results. I carried out all the photography, all the optical microscopy and spectroscopy including data processing, all the data processing of electron microscopy measurements, and integrating sphere measurements and data processing, and I wrote the entire manuscript. Additionally, Dr. Yu Ogawa performed electron microscopy, Dr. Villads Egede Johansen performed the numerical modelling of the optical response.

Chapter 3

Structural colours in the frond of *Microsorum thailandicum*

Lisa Maria Steiner[1], Yu Ogawa[1,2], Villads Egede Johansen[1], Clive Lundquist[3], Heather Whitney[3], Silvia Vignolini[1]

1. University of Cambridge, Department of Chemistry, Lensfield Road, Cambridge, CB2 1EW, United Kingdom
2. Universite Grenoble-Alps, CNRS, CERMAV, 38000 Grenoble, France
3. University of Bristol, School of Biological Sciences, Tyndall Avenue, Bristol, BS8 1TQ, United Kingdom

Steiner, L. M., Ogawa, Y., Johansen, V. E., Lundquist, C. R., Whitney, H., and Vignolini, S. (2018). Structural colours in the frond of *Microsorum thailandicum*. Journal of the Royal Society Interface Focus, 9(1):20180055.
url: https://royalsocietypublishing.org/doi/10.1098/rsfs.2018.0055,
doi: 10.1098/rsfs.2018.0055 [130]

Abstract

Blue and near-ultraviolet structural colours have often been reported in understorey plants living in high shade. While this intense blue coloration is very catchy to the eye of a human observer, there are cases in which structural colours can be hidden either by the scattered light interacting with pigments or because they are found in unexpected positions in the plants. Here we show that the fronds of *Microsorum thailandicum* produce structural coloration on

both the adaxial and abaxial epidermal surface. While cellulose helicoidal structures are responsible for this coloration in both epidermal layers, the reflected colours are consistently different: an intense blue reflection is found in the adaxial epidermis while red-shifted and less intense colours are observed in the abaxial epidermis, possibly suggesting photo-adaptation of the plant to the light environment. By comparing the optical properties of the fern with its anatomy we computed the theoretical reflection accounting for the presence of disorder in the cellulose helicoidal architecture.

3.1 Background

Structural colours are extremely widespread in nature [60, 77, 149, 106]. From flowers [153, 90] to fruits [147, 145, 76] to leaves [131, 50], such brilliant colorations are observed in several plant tissues with different biological function [135, 58], and using several morphologies [146, 60]. A common architecture to produce structural colour that is found in several plant tissues consists of assembling cellulose microfibrils into helicoidal architectures in the cell wall [158]. The inherent birefringence of cellulose microfibrils and their chiral spatial organisation provide a circularly polarised light reflection in a range of wavelengths which are determined by the dimensionality (referred to as pitch) of the helicoid [157]. Such colour-generating cellulose-based structures are found in the cell wall of many different tissues like leaves and fruits [147, 145, 92, 76, 115, 77, 49, 131]. However, in leaves, these helicoidal structures have so far only been observed in the adaxial epidermal cell walls and their colour is prevalently in the blue and near UV spectral region. Here, we studied the optical properties and anatomy of *M. thailandicum*, a member of the large genus of *Microsorum* ferns that belongs to the *Microsorum punctatum* complex [12, 109, 70]. *M. thailandicum* was described as iridescent by Boonkerd and Nooteboom [12], but the origin of the structural colour was still unclear. In our investigation, we observed that both the adaxial and abaxial epidermal cells contain helicoidal cell walls reflecting blue and green-to-red circularly polarised light, respectively. Finally, by performing electron microscopy and (micro-)spectrophotometry on the same area, we quantitatively correlated the measured reflectivity with the anatomy of the structures, providing an understanding on how the structural disorder affects the optical properties of the fern.

3.2 Methods

3.2.1 Plants

Plants were either grown in an office environment or in a growth cabinet set to 25 °C and lowest light option (approx. 2000 lux illuminance) during the day, for 16 hours, and 20 °C and darkness during the night (Panasonic Versatile Environment Test Chamber MLR-352-PE, Panasonic Healthcare Co., Ltd., Japan). They were watered from below once a week and misted daily.

3.2.2 Photography

Photos of the fern and its fronds were taken with a Nikon D3200 camera (18-55 VR II kit, AF-S DX Nikkor 18-55 mm f/3.5-5.6G VR II, Nikon, Japan), in macro mode, and automatic focussing, using a tripod. Generally, a polariser was added to reduce gloss from the cuticle (Hoya CIR-PL slim, Hoya corporation, Japan).

3.2.3 Optical microscopy

Cross-section of fronds A TEM block was made via high pressure freezing and freeze substitution (subsection 3.2.4) and sliced into semi-thin cross-sections with a Leica Ultracut E ultramicrotome (Leica microsystem GmbH, Austria). It was stained with Richardson's stain and observed in transmission microscopy on a Zeiss microscope and 5x objective to investigate the ultrastructure of the frond.

Thickness of fronds The thickness of the same fronds used for the gradient investigation (comparison of reflection intensity between fronds, see subsection 3.3.3) and integrating sphere measurements (for total transmission and reflection, see supporting material SI) was determined by cutting two thin slices in the same tip, middle and base area with a razor blade. Per cross-section, 6 thicknesses were measured, 3 on each side of the frond. They were measured at equal spacing across the area where the frond surface was horizontal (since spectra were also generally taken from that area), aligned with the shortest distance at this cross-sectional point. Images of freshly cut cross-sections of fronds to obtain their thickness were recorded with a Zeiss stereoscope and processed with ImageJ [120, 111], see subsection 3.3.3 and SI.

Polarised optical microscopy and microspectroscopy

Optical microscopy was carried out on a customised Zeiss microscope equipped with epi-illumination and a 5x, 20x and 50x objective. Different configurations were used. For polarised optical imaging, the sample was illuminated with an unpolarised halogen lamp, and a polariser and a quarter-waveplate were mounted into the optical path, and left and right channel configuration was obtained by independent motors. For images of cells, the 20x objective was used and spectra were collected via a 100 μm optical fibre. Investigation on the single-cell level was achieved with the 50x objective and the 100 μm optical fibre for collecting spectra. In-house software controlling a motorised stage was empoyed to scan a single cell with high spatial resolution.

Three fronds from one plant were investigated for the statistics of optical response study, to guarantee equivalent growth conditions.

Three fronds from another plant were investigated with respect to the gradient of structural colour observed, and whether both the adaxial and abaxial surface are always both coloured. Three spectra from the tip, the middle and the base of each frond, adaxial and abaxial surface, were taken.

All fronds were imaged while still on the plant, or alternatively the entire frond was cut off and imaged straight away without making any further cuts on it. Generally, upon being cut off and cut into pieces, the fronds will dry out, the coloration decreases and finally disappears within a few hours.

3.2.4 Electron microscopy

TEM

Sample embedding by high pressure freezing and freeze substitution 3 mm circular frond samples were cut, placed within brass specimen carriers and loaded into a Leica EM ICE high-pressure freezer, followed by immersion in liquid nitrogen and freeze substitution (Leica AFS2, Leica microsystems GmbH, Germany). The frozen samples still in their carriers were placed inside 2 ml capped cryovials with 1 ml of acetone at the surface of a bath of liquid nitrogen to avoid warming of the samples. The tips of the tweezers were similarly cooled before use. Samples were brought up to room temperature over four days after which they were transferred to 100 % ethanol, followed by a resin series (alcohol:medium grade LR White resin ratios of 3:1, 1:1, 1:3 and 100 % LR White resin). The resin was changed daily over four days after which the specimens were placed in gelatin capsules and polymerised in a Fistreem vacuum oven (digital, Fistreem International Limited, UK) at 60 °C and 440 mmHg for 22 hours.

Sample embedding by chemical fixation Small pieces of native, hydrated plant tissue were cut and entirely immersed in a buffered fixative solution containing glutaraldehyde (2 wt%) and formaldehyde (2 wt%) for 16 hours at 4 °C. The specimens were then rinsed with deionised water and fixed for 2 hours at 4 °C in a buffered OsO_4 solution. The specimens were rinsed again in deionised water and successively dehydrated in graded ethanol aqueous solutions (30-100 wt%) and then dry acetonitrile. They were incubated for 16 hours in a 50:50 (v/v) mixture of acetonitrile and Quetol 651 epoxy resin, and subsequently immersed in Quetol resin for 2 weeks, allowing the resin infiltrating into the specimens. The specimens were placed in a silicon mould with Quetol resin and cured for 48 hours at 65 °C. Finally, ultrathin sections were prepared using an ultramicrotome (Ultracut UCT, Leica microsystem GmbH, Austria) equipped with a 35° diamond knife (Diatom, USA) and mounted on continuous carbon coated copper grids. The sections were then post-stained with 1 wt% uranyl acetate aqueous solution and Reynolds lead citrate solution. TEM observations were carried out with a Philips CM-200 'Cryo' electron microscope operated at 200 kV (Thermo Fisher Scientific Inc., USA).

Cryo-SEM

Cryogenic scanning electron microscopy (cryo-SEM) observation was performed using a field-emission scanning electron microscope (Verios 460, Thermo-Fisher Scientific Inc., USA) equipped with a cryo-preparation system (PP3010T, Quorum, UK). The frond was cut into a small strip and mounted upright on a specimen holder using a colloidal graphite suspension. The specimen was quench-frozen in liquid ethane and transferred into the cryo-preparation chamber, where it was freeze-fractured, sublimed, and subsequently sputter-coated with platinum. SEM imaging was carried out at an acceleration voltage of 2 kV and a working distance of approx. 4 mm.

Before cryo-SEM, a few mm wide section of the frond was marked of and polarised optical microscopy carried out on it, so that the optical response could be correlated to the same section of the frond that the cryo-SEM measurements are from. This way, the modelling approach is based on the same small area of the frond.

Block-face SEM

A smooth surface of resin embedded specimen was prepared using an ultramicrotome (Ultracut UCT, Leica microsystem GmbH, Austria) for block-face SEM observation. SEM imaging was carried out using a concentric backscatter (CBS) detector on a field emission

scanning electron microscope (Quanta 250, Thermo-Fisher Scientific Inc, USA) operated at 4 kV with a working distance of 7 mm.

3.2.5 Data and spectra processing

Matlab was used for all spectra and data processing. Spectra were always referenced to a white diffuser (USRS-99-010, Labsphere, USA), except for the ones used for modelling, which were referenced to a silver mirror (PF10-03-P01, Thorlabs, USA). When referencing to a white diffuser, it is not unusual to obtain reflection intensities higher than unity, in the case that the reflection of the sample is strong and more directional than the white diffuser. For the gradient analysis, the spectra of three cells at the tip, the middle and the base of each frond were averaged, for both adaxial and abaxial surface, and a high, a medium and a low structurally coloured frond, respectively. To obtain histograms of the peak maxima, the inbuilt matlab maxvalue function was used. To obtain the full width at half maximum (FWHM), we used the matlab function findpeaks. The distribution of peak positions was then fitted with a Gaussian distribution, and the distribution of peak widths (FWHM) with a log-normal distribution. For obtaining the average spectra, all spectra of the same surface were averaged. For the analysis of pitches from the cryo-SEM and block-face SEM images, the imageJ greyscale function was used to count layers. Three lines were measured per cell (corresponding to one SEM image), on the left, middle and right, and these values were averaged, to obtain the average and standard deviation (plotted as error bars) per cell. Finally, individual spectra are plotted for the modelling section, to demonstrate how much heterogeneity and how many different spectral features are found for different cells.

3.2.6 Modelling

The freely available Python implementation of Berreman 4x4 (Aug 21, 2016 on github [22]) was used for all simulations, using Python 3.6. We tried out different combinations of parameter fitting of twist defects, pitches, and normal and extraordinary refractive indices in order to match the experimental spectra. The total number of half pitches was fixed to 80 to resemble a total cell wall height of approx. 10 to 15 μm. We used the discrete sum of absolute differences between spectrum and simulation of the main peak as the fitting objective (also called the L1 norm). Since the optimisation problem was expected to have many local minima, we tried the different global optimisation algorithms implemented in the Python library SciPy 0.19. We found that basinhopping worked best for this case and gave most consistent results throughout multiple runs. Convergent results were normally observed within a few minutes, and the optimiser was therefore manually terminated. We found that

refractive index parameters of $n_o = 1.528 + 0.0075i$ and $n_e = 1.474 + 0.0075i$ gave a peak height and width that most closely fitted the shape of the main peak of the measurements. This is close to reported parameters for other helicoidal systems [34, 132]. In trying to fit to more complex spectral shapes, we did not obtain any better fit (with realistic parameters) that resembled the shape of the recorded spectra.

3.3 Results

3.3.1 Optical response of frond

The optical response of the fronds of *M. thailandicum* was studied using a customised microscopy setup which allows to simultaneously image the epidermis of the frond in different polarisation configurations and to collect spectra in the corresponding imaged area.

Figure 3.1 A shows a photo of the plant, there are several fronds with an intense blue coloration. Figure 3.1 B and C depict epi-illumation microscope images of the same area of the adaxial epidermis of a blue frond in the left and right circular polarisation channels (LCP and RCP), respectively. The blue colour, corresponding to the position of the cell in the epidermal layer, is visible only in the LCP, but not in the RCP (Figure 3.1 B and C), indicating the presence of a helicoidal structure in the cell wall of the outer epidermal cell. Interestingly, the abaxial epidermis displays similar properties: a clear LCP reflection but no RCP reflection is observed from the cells in the green and red spectral region, see Figure 3.1 E and F, even though the abaxial surface appears non-iridescent by naked eye, as seen on the encircled area in Figure 3.1 D.

3.3.2 Anatomy of frond

Due to the interesting optical appearance, the ultrastructure of the frond was investigated to locate and confirm the helicoidal architecture. Figure 3.2 A depicts an optical transmission image of a semi-thin section of an embedded cross-section. Both on the adaxial and abaxial surface, the upper or lower epidermis are clearly visible, the cell walls of these cells look thickened, and the epidermis is covered with the cuticle. The mesophyll consists of the palisade tissue towards the adaxial surface, and the spongy mesophyll towards the abaxial surface.

Zooms of the adaxial and abaxial epidermal cells via cryo-SEM are shown in Figure 3.2 B-C and E-F, respectively. By increasing the magnification, it is possible to observe a thickened cell wall with a layered structure for the outermost layer of cells in the two epidermises. Furthermore, the same thickening with layered structure can also be observed

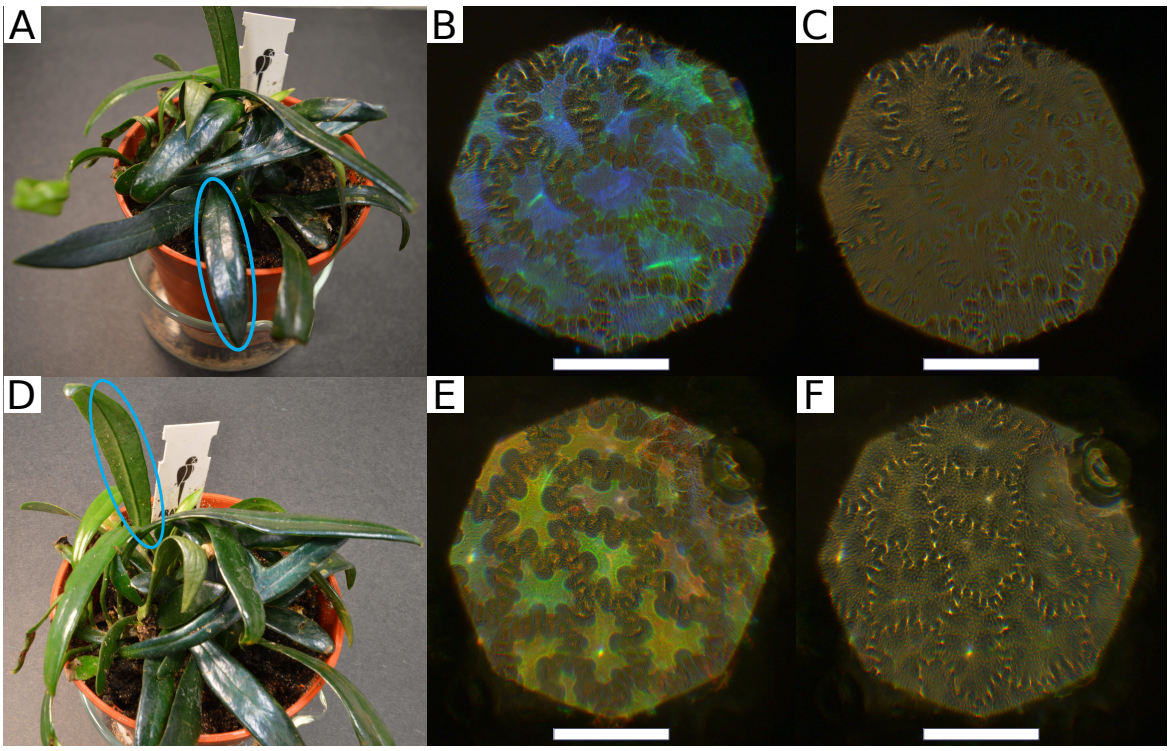


Fig. 3.1 (A) Photo of *M. thailandicum*, adaxial surface encircled in blue was studied. (B,C) Optical micrograph of reflection in left-handed (LCP) and right-handed circular polarised light channel (RCP) of adaxial surface. (D) Photo of *M. thailandicum*, abaxial surface encircled in blue was studied. (E,F) Optical micrograph of reflection in LCP and RCP of abaxial surface. Scale bar is 100 μm.

in the surface-facing side of the cell walls of the second epidermal layer, see red boxes (B-F). TEM was used to further investigate these regions, see Figure 3.2 D and G, where the Bouligand arcs characteristic for the helicoidal arrangement of cellulose microfibrils in the cell wall are observed [13, 85].

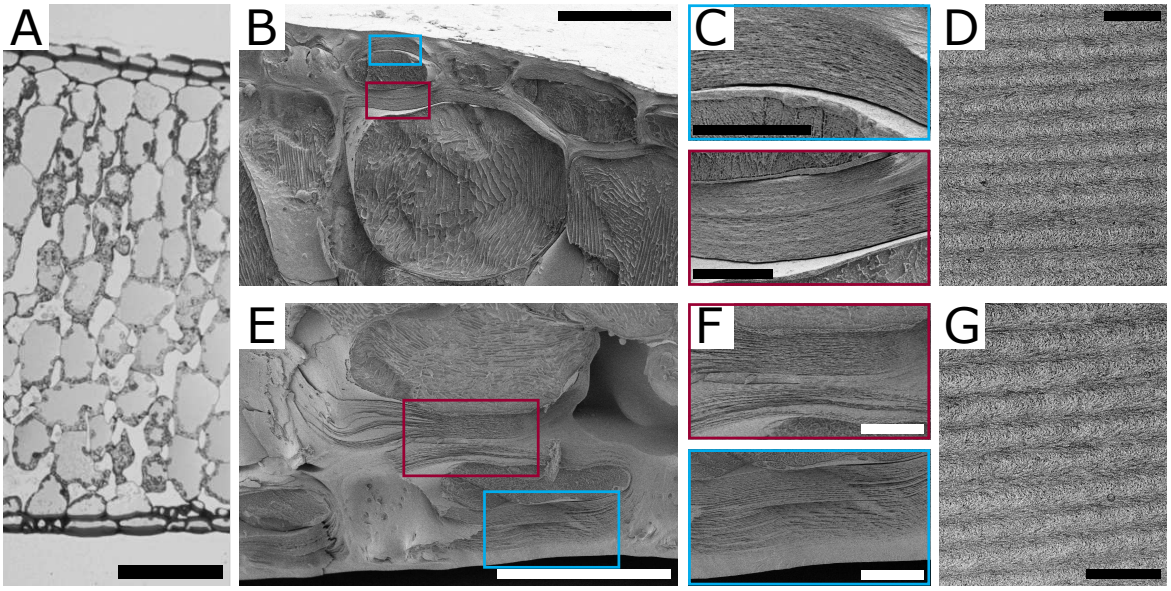


Fig. 3.2 (A) Optical transmission micrograph of semi-thin section of embedded frond. Scale bar is 200 μm. (B,E) Cryo-SEM image of adaxial and abaxial epidermal cells of the frond, respectively. Scale bar is 50 μm. (C,F) Zoom of epidermal cell walls in the blue/red boxes for adaxial and abaxial epidermis, respectively. Scale bar is 10 μm. (D,G) TEM images of adaxial and abaxial cell wall, respectively, showing the Bouligand arcs characteristic for the helicoidal arrangement of cellulose microfibrils. Scale bar is 500 nm.

3.3.3 Variation between fronds

When looking at a number of plants (Figure 3.3 A), we noticed that, even though the macroscopic appearance of a single frond is fairly homogeneous, there is a large amount of variation in intensity of structural coloration between different plants, and even between different fronds on the same plant. Whether and how intensely fronds develop structural coloration probably depends on a variety of factors, like temperature, light and humidity, which we were not able to fully control over a long enough period of time. Therefore, to account for this variation, we studied the intensity and spectral variation of the structural colour in representative fronds from the same plant. The fronds were chosen to have one frond show very intense blue structural colouration on the adaxial surface, one that almost did not show any, and one in between the two extremes. To account for the age of the frond,

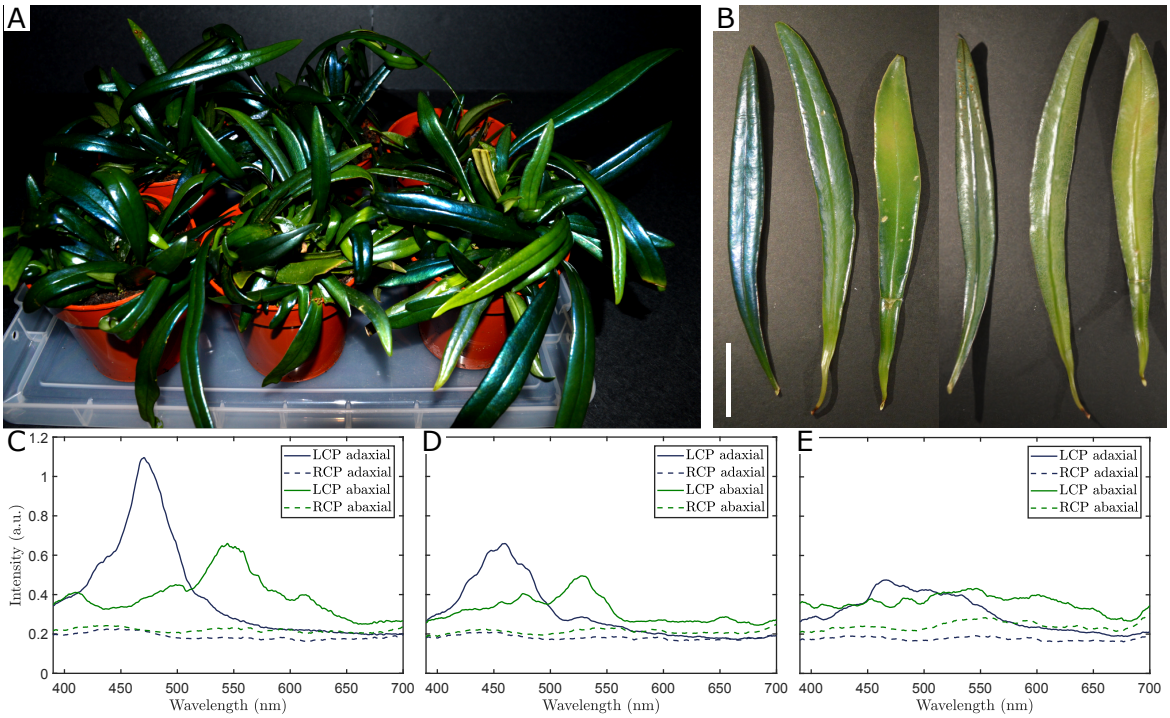


Fig. 3.3 (A) Photo of a variety of specimens. (B) Photo of the adaxial (left) and abaxial (right) surface of three fronds showing high, medium and low structural coloration, from left to right. Scale bar is 2 cm. (C-E) Average reflection spectrum in LCP and RCP of frond, adaxial and abaxial surface: (C) Spectra of the most intensely structurally coloured frond. (D) Spectra of the medium structurally coloured frond. (E) Spectra of the low structurally coloured frond. All spectra are averaged from measurements of nine cells each across the frond.

we collected fronds with comparable stiffness, length and thickness. In more detail, the collected fronds had a thickness of (0.92 ± 0.20) mm for the intensely structural coloured one, the medium one of (1.00 ± 0.20) mm, and the low coloured one (1.07 ± 0.27) mm.

In particular, we investigated the gradient of structural colour observed, and whether both the adaxial and abaxial surface are always both coloured with the same intensity, see Figure 3.3 B. As visible in the photograph, the spectra of the most intensely coloured frond showed the highest intensity, the medium coloured frond medium intensity, and the low coloured frond the least intensity of reflection. This trend was observed for both the adaxial and the abaxial surface and structural colouration for both epidermises correlate in intensity, see Figure 3.3 C-E.

3.3.4 Variation of the optical response within the same frond

Additionally to the variation of macroscopic appearance of the fronds, variation between the individual cells on each frond is revealed by optical microscopy, see Figure 3.1 B and E. To account for this variation, we statistically analysed the optical response of 100 cells of the adaxial surface from the strongly structurally coloured frond shown in Figure 3.1 A, and 74 cells of the abaxial surface from the strongly structurally coloured frond shown in Figure 3.1 D. Firstly, the maximum of each LCP reflection was determined for both adaxial and abaxial epidermal cells, and these values were plotted in a histogram, shown in Figure 3.4 A and D, respectively . The distribution of the maximum reflection wavelength of the peaks is well approximated with a Gaussian distribution, while the distribution of the full width half maximum (FWHM) of the peaks shown in Figure 3.4 B and E is approximated with a log-normal distribution.

Interestingly, the variation of reflected colours is much narrower for the adaxial surface than for the abaxial surface. For the adaxial surface, all reflection maxima are found between 400-550 nm. The Gaussian distribution gives a mean value of 460 nm, and a standard deviation of $\pm$ 26 nm. This corresponds very closely with the averaged spectrum of all 100 cells, which has a maximum reflection at 461 nm, see the supplementary material (SI). Furthermore, the log-normal fitting of the FWHM histogram yields the values 18.2 nm and 1.4 nm for the mean μ and for σ, giving a standard deviation of $\pm$ 6.7 nm. On the other hand, the reflection of the abaxial epidermis varies much more, all the reflection maxima are found between 400-650 nm, so almost over the entire visible spectrum. When sorting the reflection maxima into a histogram, the approximated Gaussian distribution yields a mean value of 524 nm and a standard deviation of $\pm$ 56 nm. Correspondingly, the shape of the average abaxial reflection spectrum is very wide, and the reflection maximum lies at 511 nm, see SI. Moreover, μ, σ and the standard deviation for the log-normal distribution of peak widths are 21.5, 1.7 and $\pm$ 13.2 nm, respectively, also showing greater variation than for the adaxial epidermis.

In order to compare these observations to the anatomy of the fronds, we investigated the helicoidal architecture on an individual cell level by electron microscopy. A representative block-face SEM image of the adaxial or abaxial outermost thickened epidermal cell wall used to measure the pitch p are shown in the SI. The pitch p is the height of the helicoidal axis within which the cellulose microfibrils complete a 180° rotation.

The pitch p lies between 150 to 220 nm for the adaxial outermost thickened epidermal cell wall, with most values between 160 and 200 nm, with the standard deviation within each cell below $\pm$ 22 nm, typically below $\pm$ 10 nm, see Figure 3.4 C (from 24 cells). For the abaxial outermost thickened epidermal cell wall, the pitch p is spread out over a much bigger

range, between 120 to 290 nm with most values between 150 and 250 nm, and the standard deviation within each cell is bigger as well, up to $\pm$ 42 nm (from 25 cells), as depicted in Figure 3.4 F.

The pitch p is related to the reflection maximum λ via $\lambda = 2 \cdot n \cdot p$, where n is the average refractive index of the medium, which can be approximated to 1.50, by considering that the typical value of the refractive index for pure crystalline cellulose is 1.55 [157], and an average value of 1.45 from the other cell wall components, namely cellulose microfibrils with disordered surface, hemicellulose, lignin, water, small amounts of protein, etc. [16].

Even if the the calculated reflection maxima from the pitch data are higher for both adaxial and abaxial cell walls, which is unusual since one may expect shrinkage of structures during the chemical TEM sample processing [29, 87], we observed the same trends for both surfaces, using both techniques (cryo-SEM and block-face SEM). We can account for this discrepancy by considering that we are overestimating the pitch by approximately 10%. We believe that the main reason that leads to the measurement of a larger pitch, both with cryo-SEM and block-face SEM, is due to the not perfectly perpendicular cuts with respect to the direction of the helicoidal axis and due to the not perfectly perpendicular measurement direction when processing the images. If the measurement direction is perfectly aligned with the helicoidal axis, the value of the pitch is correct, but any deviation from perfect alignment will always lead to an overestimation of the pitch [42]. While utmost care was taken during measurements and during image processing in imageJ, contributions from this issue cannot be fully excluded, also since the axis of the cell and therefore the helicoids is different from cell to cell and can also change slightly within the cell.

Additionally, integrating sphere measurements of areas of a few millimetre from very strongly and very low structurally coloured fronds were also performed to estimate the total transmission through and reflection of the fronds and the adaxial and abaxial epidermis separately, see SI.

3.3.5 Variation within single cell

To investigate the disorder in the distribution of the pitch, a single cell of the adaxial epidermis was scanned with larger spatial resolution. Figure 3.5 A and B report the statistical analysis of the measured reflection spectra (22 in total) as processed in the same way as in subsection 3.3.4 and in the SI, reporting the reflection maxima in a histogram and the average of all spectra. The average of all 22 spectra of this cell gives a maximum of 506 nm, again corresponding closely with the histogram. Furthermore, all the obtained reflection spectra fall well within the range observed for the previous measurement of 100 adaxial cells across

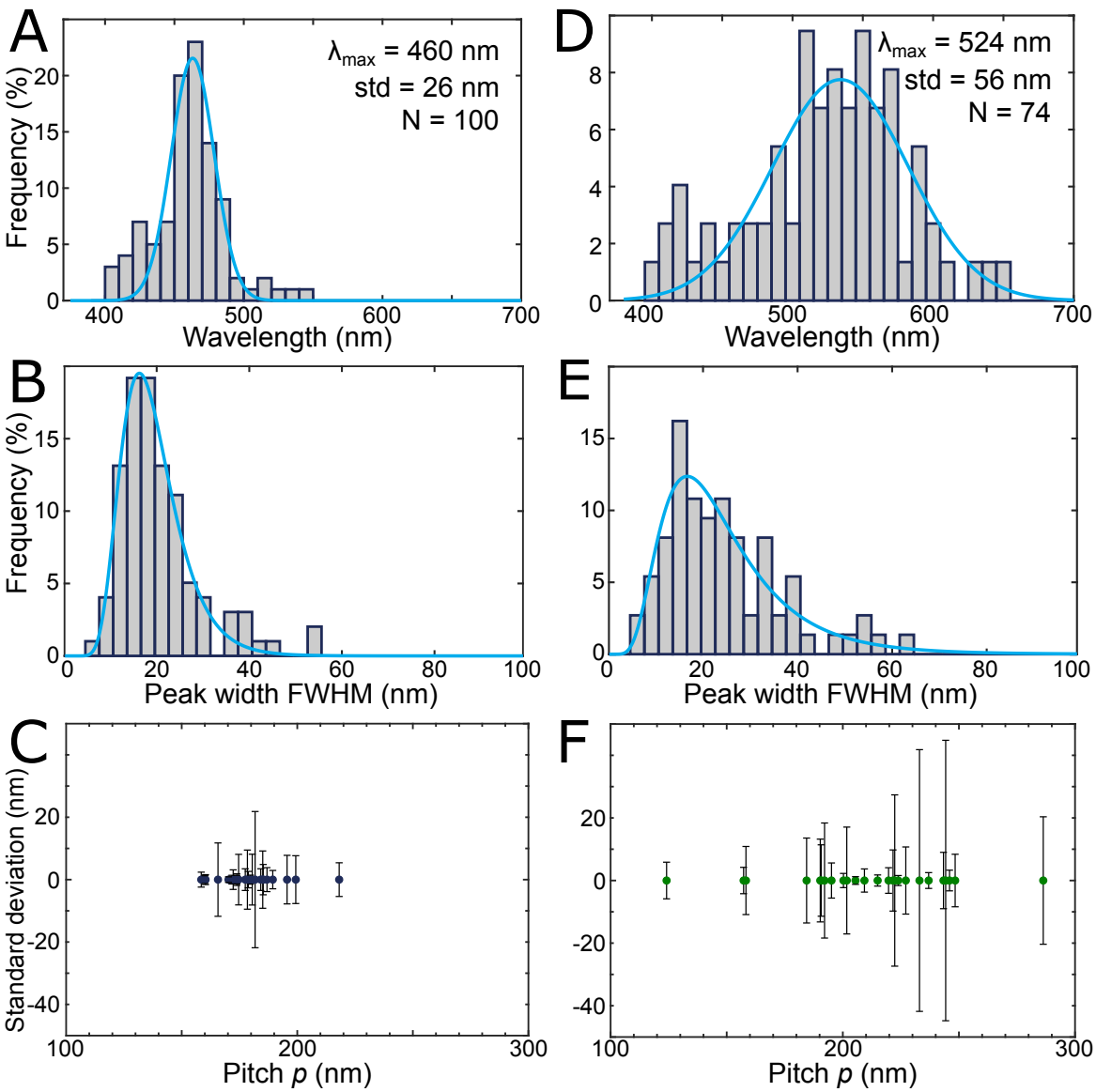


Fig. 3.4 (A-C) Adaxial surface. (A) Distribution of peak wavelengths from structurally coloured cells of frond. (B) Distribution of peak widths as determined by finding the full width at half maximum (FWHM). (C) Distribution of pitches p and their standard deviation measured from the layering visible in the uppermost helicoidal cell wall measured from block-face SEM and cryo-SEM images. (D-F) Abaxial surface. Same as for adaxial surface.

a frond. This means that the variation is the same on a frond, between different cells, as it is within a single cell.

3.3.6 Modelling the optical response

Next, a very small area of a frond was marked off and subjected both to polarised optical microscopy and then cryo-SEM analyses. A representative optical micrograph of the LCP is shown in Figure 3.5 D, and example spectra from that area are shown in Figure 3.5 F, while a cryo-SEM image of the layered cell wall of the same area is shown in Figure 3.5 E. The variation of the pitch p measured from this frond area was then utilised to model the optical response (Figure 3.5 G).

We modelled the circularly polarised spectral reflection using an open source Python implementation of Berreman's 4x4 matrix method that simulates stratified (layered) anisotropic media [22, 9]. This is the most common approach for simulating helicoidally arranged cellulose microfibrils [34]. In Figure 3.5 G, an ideal helicoidal arrangement of cellulose microfibrils with 40 half pitches of 159.9 nm and refractive indices of $n_o = 1.528 + 0.0075i$ and $n_e = 1.474 + 0.0075i$ was simulated. The figure shows that the main reflection peak position, height and width can be roughly captured, but also that the recorded spectra are much more complex. This is probably due to significant deviations in the helicoidal twist of the cellulose layers from an ideal helicoidal structure (twist defects, varying pitch, etc.) and other geometrical artifacts (curvature of cells, non-planar cellulose layers,etc.). To obtain more information on the arrangement of the cellulose stack, we furthermore tried to use the classifications proposed by Carter et al. for a non-ideal helicoidal reflector in beetles. The best matching classification is 'Spectra with diminishing oscillations', but such spectra are not well described by a few local defects or pitch changes, as described in their SI [21]. Furthermore, looking at individual spectra from several cells (Figure 3.5 F), we also found that no single classification fitted them all. We therefore conclude that the spectral features indicate a large degree of disorder distributed throughout the cell wall and are not localised to a few defect sites or abrupt pitch changes. This inference corresponds well with the observations made from electron microscopy imaging, see Figure 3.5 E for an example.

3.4 Discussion and conclusions

3.4.1 Colour variation and plant cell wall biosynthesis

Our systematic statistical investigations allow us to conclude that, despite the variation in the reflection response from different fronds in the plant, the reflection maxima of the

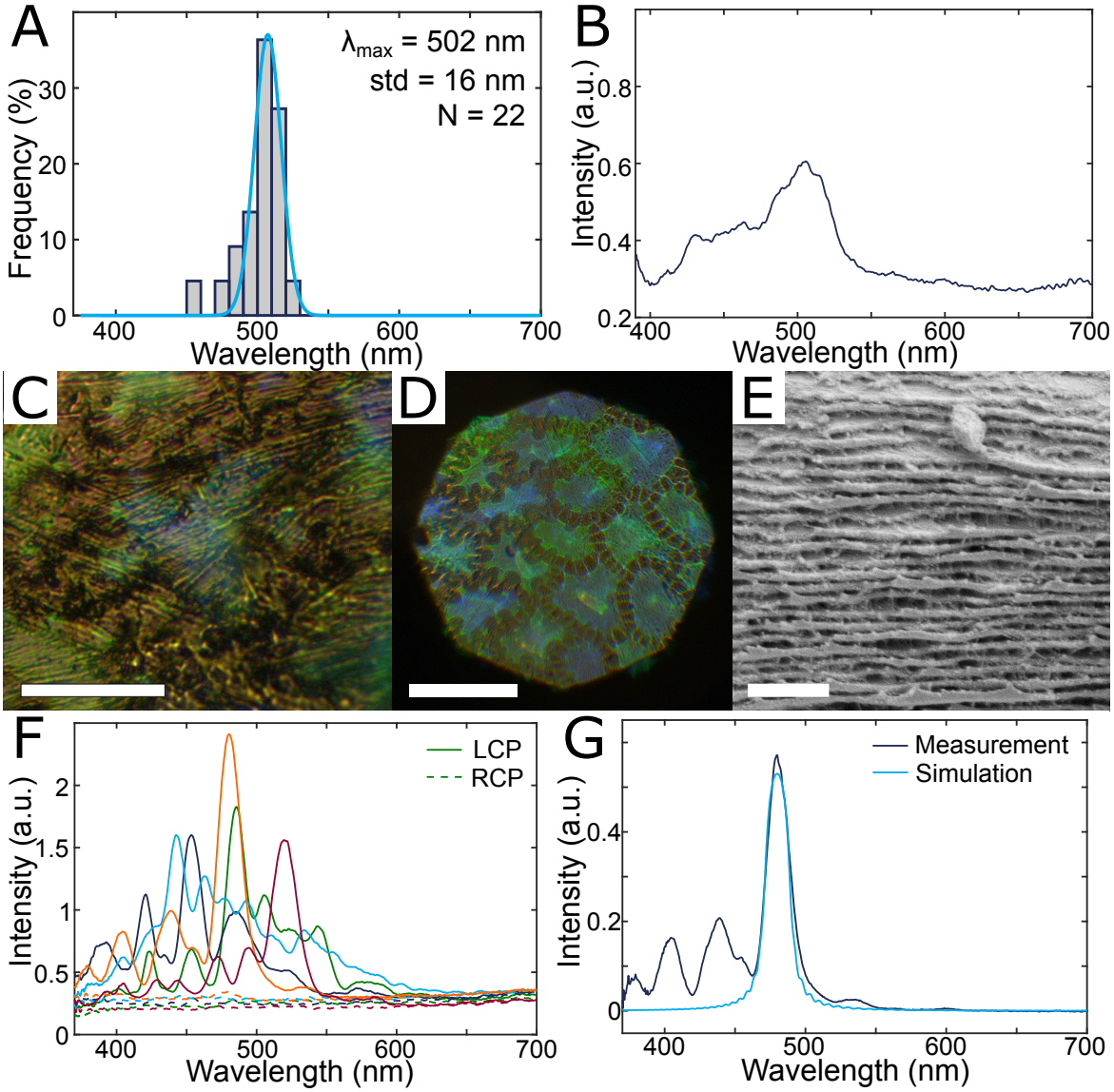


Fig. 3.5 (A) Reflection maxima distribution of single cell. (B) Average spectrum of all 22 LCP spectra from single cell. (C) Optical micrograph of reflection in LCP of single cell of adaxial surface. Scale bar is 20 μm. (D) Optical micrograph of reflection in LCP of adaxial surface, specific part of frond for cryo-SEM. Scale bar is 100 μm. (E) Cryo-SEM micrograph zoom of the helicoidal layering, adaxial surface. Scale bar is 1 μm. (F) Individual reflection spectra in LCP and RCP of 5 cells of a specific part of frond taken for cryo-SEM, adaxial. (G) Measurement taken from a single cell and simulation based on the Berreman 4x4 method (referenced to silver mirror).

adaxial epidermis are much narrower with less variation between cells and within cells, compared to the abaxial epidermis. These observations have interesting implications for the biosynthesis of the adaxial versus the abaxial epidermal cell walls. For the abaxial cells, there is much bigger variation between the different cells, and the range of reflected colours is larger, hence the biosynthesis of the plant cell wall is presumably less orderly regulated than for the adaxial surface. Furthermore, when the reflection wavelength lies at the green or red end of the visible spectrum, it means that the cellulose microfibrils in the helicoidal cell wall architecture have to be spaced further apart than when reflecting in the blue range. We do not know what exactly is used as a spacer between the layers of parallel cellulose microfibrils, but there are at least three different options: *(i) hemicellulose and lignin content.* We think it is most likely that hemicellulose or lignin or both act as a spacer between the cellulose microfibrils. This would mean that in green and red cells, more material, like hemicellulose or lignin is deposited. Quantifying the hemicellulose and lignin content (ideally at a single cell level, but at least for adaxial and abaxial epidermis separately) could give some insights on whether they act as a spacer, in which case their biosynthetic pathways should be investigated further. *(ii) rotation angle.* Another option is that the rotation angle between the cellulose microfibrils is smaller in the abaxial epidermal cell walls. In this case, more layers of cellulose microfibrils would be layered up and would thus increase the pitch, resulting in a shift of reflection to longer wavelengths. If there is a smaller rotation angle between cellulose microfibrils, a higher cellulose content in the abaxial epidermis could be observed. To investigate this hypothesis, the cellulose content should be investigated for the two epidermises separately. *(iii) water content.* The third option is a difference in water content. The abaxial epidermis does seem more hydrated than the adaxial when preparing TEM specimens, and dehydration does affect the structural colouration of the fronds, making higher water content in the abaxial cell walls an option.

3.4.2 Influence of disorder on optical response

The model adopted so far to systematically analyse the optical response by helicoidal structures does not encompass all factors contributing to the spectral response. By trying to fit our spectra to the model developed to take into account defects and irregularities in spectra of helicoidal beetle cuticles[21], we were not able to reproduce all the measured spectral features. Therefore, we conclude that a few discrete defect sites do not dominate the reflection spectrum of the plant cell walls. Rather, we expect the reflection to be caused by a more complex and distributed disorder in the cell walls. Looking at spectra from several cells, we also found that no single classification from Carter et al. fitted them all, suggesting that the spectral features indicate a large degree of disorder distributed throughout the cell

wall and not localised to a few defect sites or abrupt pitch changes. This conclusion is also supported by our TEM imaging, where a lot of small irregularities in the layering can be observed. Furthermore, in our statistical analysis of the variation of reflections on the same frond, we found that the peak widths (FWHM) follow a log-normal distribution, rather than a Gaussian. This hints to the concept that many small defects are found increasing the peak width, and they add up in a logarithmic way.

3.4.3 Photosynthesis and light harvesting

We speculate that the transmitted light through the adaxial epidermis and mesophyll (including chlorophyll) could then be reflected back into the mesophyll by the abaxial epidermis. So while the mainly blue reflection of the adaxial epidermis could protect the frond from photo-damage in high light conditions, the abaxial epidermis could increase light harvesting in low light conditions, by reflecting parts of the light back into the mesophyll, that would have otherwise just been transmitted and lost for photosynthesis. However, investigating this hypothesis experimentally has proved difficult for a number of reasons: *(i) separating the different layers.* Unfortunately it is not possible to remove either of the epidermal layers without damaging the mesophyll. Ideally, we would have measured total transmission through the native frond, and then removed the abaxial epidermis and measured transmission through the adaxial epidermis plus mesophyll. This issue could be circumnavigated by just removing the respective other epidermis, taking spectra thereof and then subtracting them from the spectra of the native frond, but there are additional complications. Even though it is possible to remove either epidermis intact and carefully scrape off remaining mesophyll tissue with a razor blade, even after rinsing, there is always a small amount of chlorophyll left on the epidermis, which is impossible to remove fully without destroying the epidermis, and which is impossible to quantify. This small amount of chlorophyll will always influence measurements in a non-controllable way. *(ii) different amounts of chloroplasts.* Fronds will have different amounts of chlorophyll, and possibly varying ratios of the different types like chlorophyll a and b [24, 160]. *(iii) different thicknesses of fronds.* We tried to only pick mature fronds of similar length for analysis to keep results comparable (younger fronds are less stiff and still more flexible), but the plants grow really slowly, and the fronds possibly thicken with age. *(iv) curvature of fronds.* The curvature of fronds varies considerably, and with it the surface area which is horizontal. This surface property will especially influence reflection properties, unless it is possible to decrease the spot size enough, in which case a large amount of measurements is necessary to obtain meaningful statistics. While the issue of the curvature of fronds *(iv)* could be circumvented by investing a large amount of working hours, tackling the issues of different amounts of chloroplasts *(ii)* and the different thicknesses of

fronds *(iii)* are more challenging. Carrying out the integrating sphere measurements requires separating the three layers, and the mesophyll is always destroyed in the process. Hence, the thickness and chlorophyll content could not be determined on the same area of frond either way. Again, carrying out these analyses on a large amount of fronds to obtain meaningful statistics to then relate to any part of frond might be an option.

From our transmission data of the native fronds, we observed that there is very little transmission for most part of the visible spectrum, except for a small spectral area peaking at 550 nm, approximately from 510 to 590 nm. We found this interesting, since the area more or less coincides with the main reflection range of the abaxial epidermis. At the same time, however, there is also an absorption minimum for chlorophyll in this range [24, 160], suggesting that there possibly is no optical function of the abaxial reflection.

3.4.4 Conclusions

In conclusion, we observed that the cell walls of *Microsorum thailandicum* produce structural coloration on both the adaxial and abaxial epidermal surface by a helicoidal architecture of cellulose microfibrils. Whilst there is a large variation in the optical response of the fronds, we find significant trends in the response of the adaxial and abaxial epidermis: the adaxial cell walls cause a much more well-defined reflection than the abaxial cell walls. While the biosynthesis of the plant cell wall and the biological significance of the differences between adaxial and abaxial epidermis are still far from being understood, we speculate that there might be a function in such an optical response, and suggest that there is still a lot to do to understand the strategies that plants use to manage light transport in their tissues.

Author contributions

LMS, HW, SV designed research, LMS, YO, VEJ, CL, SV performed research, LMS, YO, VEJ, SV analysed data, LMS, SV led the writing of the manuscript, YO, VEJ, CL, HW contributed to discussions. LMS planned experiments, performed optical microscopy and photography, analysed and processed data. YO performed electron microscopy. VEJ performed simulations, automated data acquisition, data processing and data analysis. LMS, SV performed integrating sphere measurements.

Acknowledgements

We thank Bonan Zhu and Gen Kamita for development of software for hyperspectral imaging of the single cell. YO thanks the NanoBio-ICMG platform (FR 2607) for granting access to the electron microscopy facility.

3.5 Supplementary information

3.5.1 Original micrographs of figure 2

Figure 3.6 depicts the original, uncropped SEM and TEM images used in figure 2 of the main text.

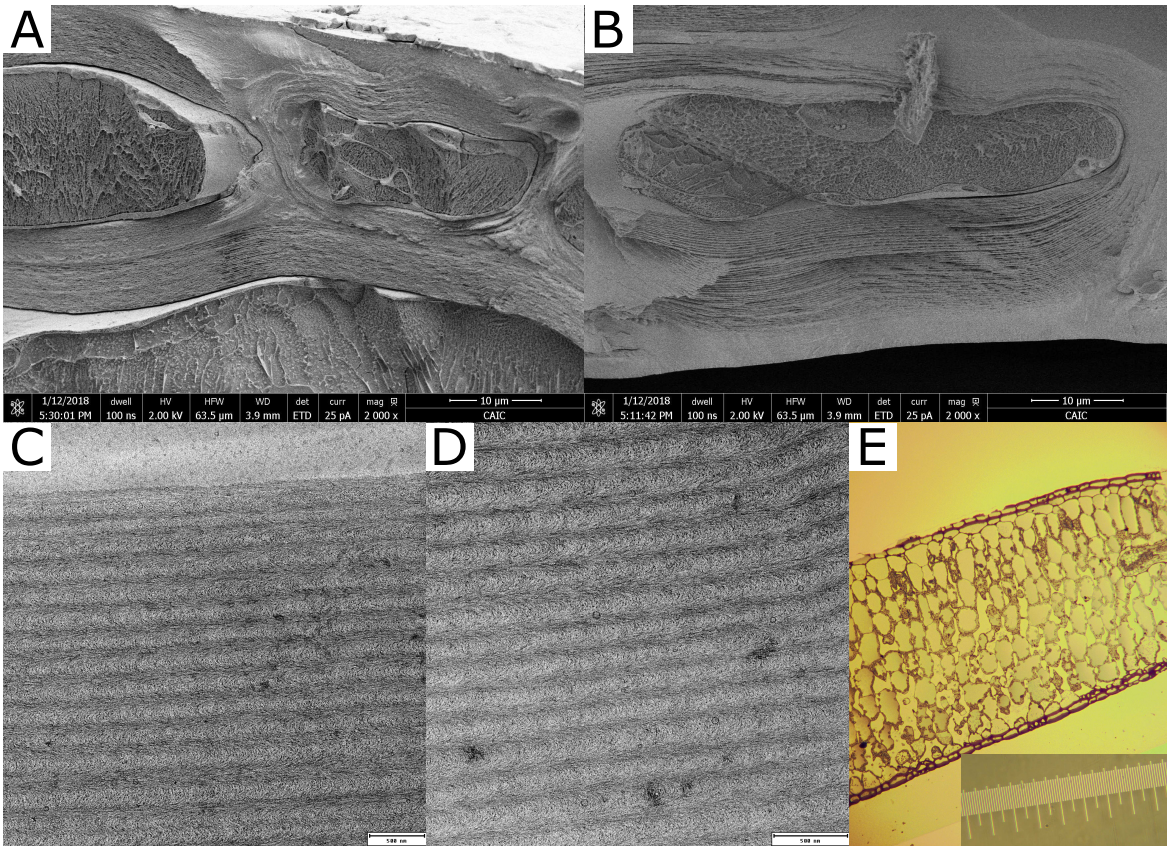


Fig. 3.6 Original micrographs of figure 2. (A,B) Zooms of cryo-SEM images. (C,D) TEM images. (E) Optical transmission image.

3.5.2 Block-face SEM for single cell statistics

Small pieces of native, hydrated plant tissue were cut out from a frond of *M. thailandicum*, and entirely immersed in a buffered fixative solution containing glutaraldehyde (2 wt%) and formaldehyde (2 wt%) for 16 hours at 4 °C. The specimens were then rinsed with deionised water and fixed for 2 hours at 4 °C in a buffered OsO_4 solution. The specimens were rinsed again in deionised water and successively dehydrated in graded ethanol aqueous solutions (30-100 wt%) and then dry acetonitrile. They were incubated for 16 hours in a 50:50 (v/v) mixture of acetonitrile and Quetol 651 epoxy resin, and subsequently immersed in Quetol resin for 2 weeks, allowing the resin infiltrating into the specimens. The specimens were placed in a silicon mould with Quetol resin and cured for 48 hours at 65 °C. A smooth surface of resin embedded specimen was prepared using an ultramicrotome (Ultracut UCT, Leica microsystem GmbH, Austria) for block-face SEM observation. SEM imaging was carried out using a concentric backscatter (CBS) detector on a field emission scanning electron microscope (Quanta 250, Thermo-Fisher Scientific Inc, U.S.A) operated at 4 kV with a working distance of 7 mm.

Figure 3.7 shows SEM micrographs of the outermost thickened cell wall of the adaxial and abaxial epidermis, respectively. The layering is clearly visible, and imageJ was used to count layers via their greyscale function. Three lines were measured per image, on the left, middle and the right of the image, paying attention to aligning the line parallel to the helicoidal axis as much as possible. Then the average and standard deviation per cell were calculated from the obtained pitch values.

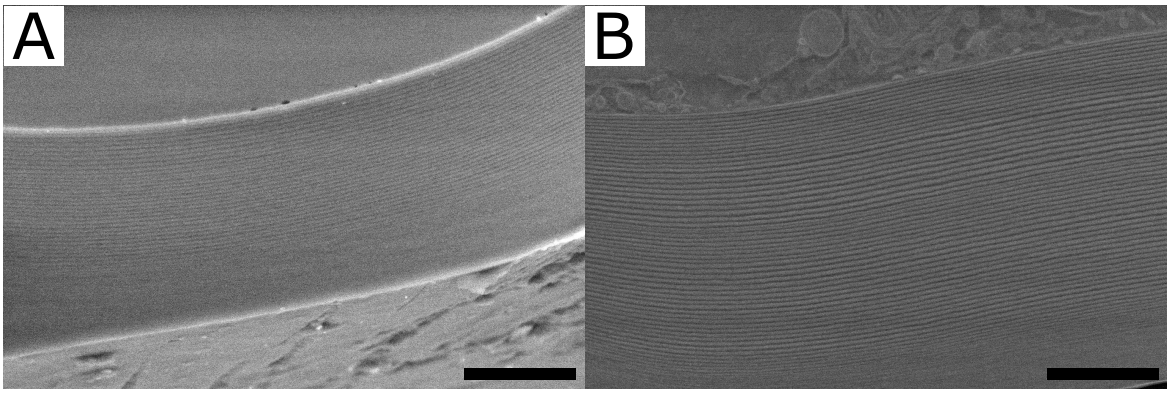


Fig. 3.7 Block-face SEM micrographs. (A) Zoom of the outermost thickened cell wall of the adaxial epidermis. (B) Zoom of the outermost thickened cell wall of the abaxial epidermis. Scale bar is 5 μm.

Variation of the optical response within the same frond and within a single cell

100 cells of the adaxial surface of a strongly structurally coloured frond were analysed, and 74 cells of the abaxial surface from another strongly structurally coloured frond. Additionally to the peak position and peak width histograms (figure 4 of the main text), the average of all obtained LCP spectra was calculated, see Figure 3.8 A and B. For both surfaces, the shape of the peak position histogram and the shape of the average spectrum and their ranges correspond very closely. The obtained averaged maximum for the adaxial cells lies at 461 nm, while for the abaxial cells it lies at 524 nm, both very close to the maxima of the Gaussian distribution fitted to the peak maxima.

To obtain the standard variation based on the fitting results for the log-normal distribution, to following equation was used:

$$s = e^{\mu+0.5\cdot\sigma^2} \cdot \sqrt{e^{\sigma^2} - 1} \tag{3.1}$$

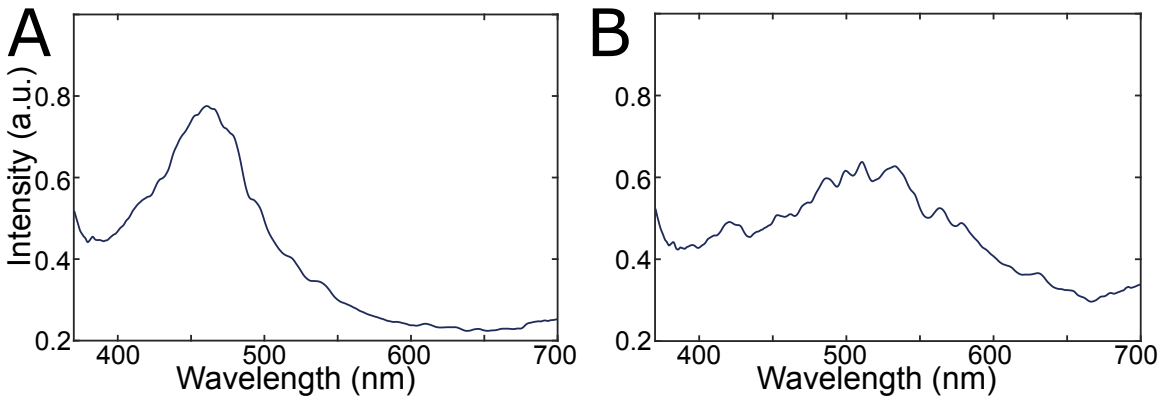


Fig. 3.8 Average of all LCP spectra. (A) Average of 100 cells of the adaxial surface of a strongly structurally coloured frond. (B) Average of 74 cells of the abaxial surface from another strongly structurally coloured frond.

3.5.3 Thickness of fronds

The thickness of the same fronds used for the gradient investigation and integrating sphere measurements was determined by cutting two thin slices in the same tip, middle and base area with a razor blade. Per cross-section, 6 thicknesses were measured, 3 on each side of the frond. They were measured in equal distance across the area where the frond surface was horizontal (since spectra were also generally taken from that area), and aligned with the shortest distance at this cross-sectional point. Images of freshly-cut cross-sections of fronds to obtain their thickness were recorded with a Zeiss stereoscope.

Figure 3.9 depicts the cross-section of the middle section of a strongly structurally coloured frond, with the 6 lines superimposed. The insets show a cross-section through the base section, and through the tip section of the same frond, respectively. It is obvious that the thickness varies across the cross-section, as well as along the frond the curvature of and where its surface is horizontal.

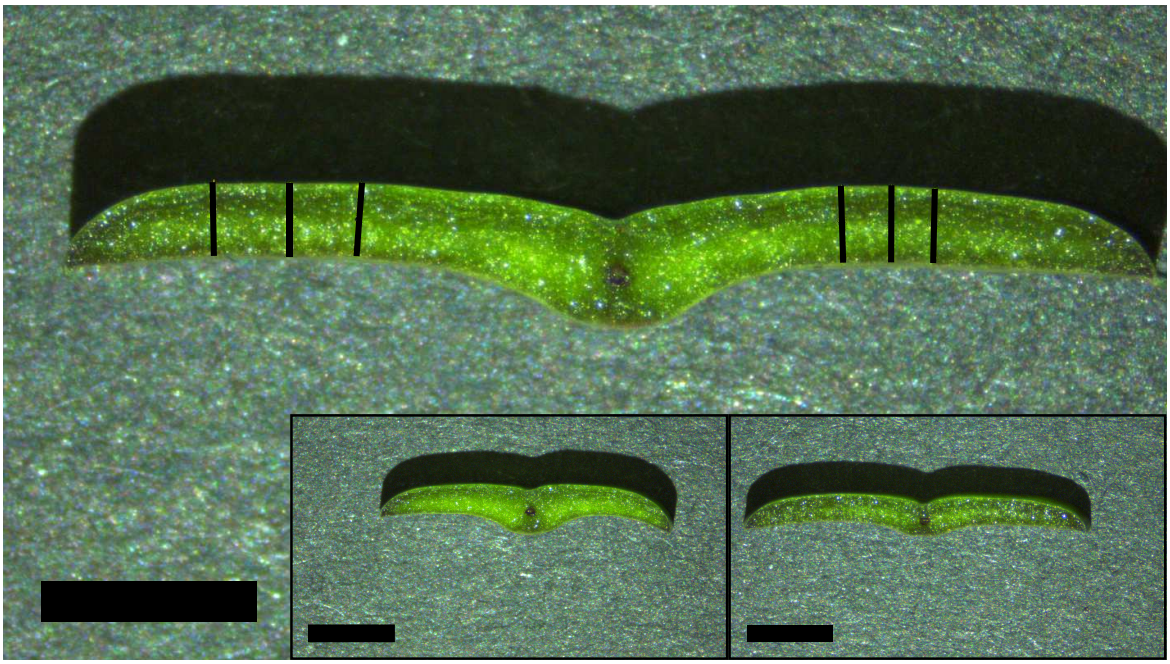


Fig. 3.9 Micrographs of cross-sections through a strongly structurally coloured frond. The 6 lines used for thickness measurement of this middle section are superimposed. Insets: cross-section of base section and tip section, respectively. Scale bar is 3 mm.

3.5.4 Integrating sphere measurements

Total transmission and total reflection were measured using a Labsphere (Halma group, USA), a xenon lamp (Ocean Optics Inc. Halma PLC, USA), and 600 μm optical fibres for illumination and collection. Native fronds or only the peeled abaxial or adaxial epidermis were mounted. The epidermis was peeled by slicing the frond horizontally with a razor blade, and carefully scrapping off all mesophyll tissue, occasionally rinsing with milliQ, until only an almost transparent thin layer was left.

All measurements were carried out on blue and green fronds, but no obvious difference was observed between the two. First, the native frond was measured, and then in the same area, it was sliced apart and the abaxial and adaxial cuticle isolated and measured separately within a few minutes to avoid drying out. This way, all measurements of transmission

or reflection were carried out in the same area, to minimise effects of frond curvature. Figure 3.10 shows representative results for a strongly structurally coloured frond.

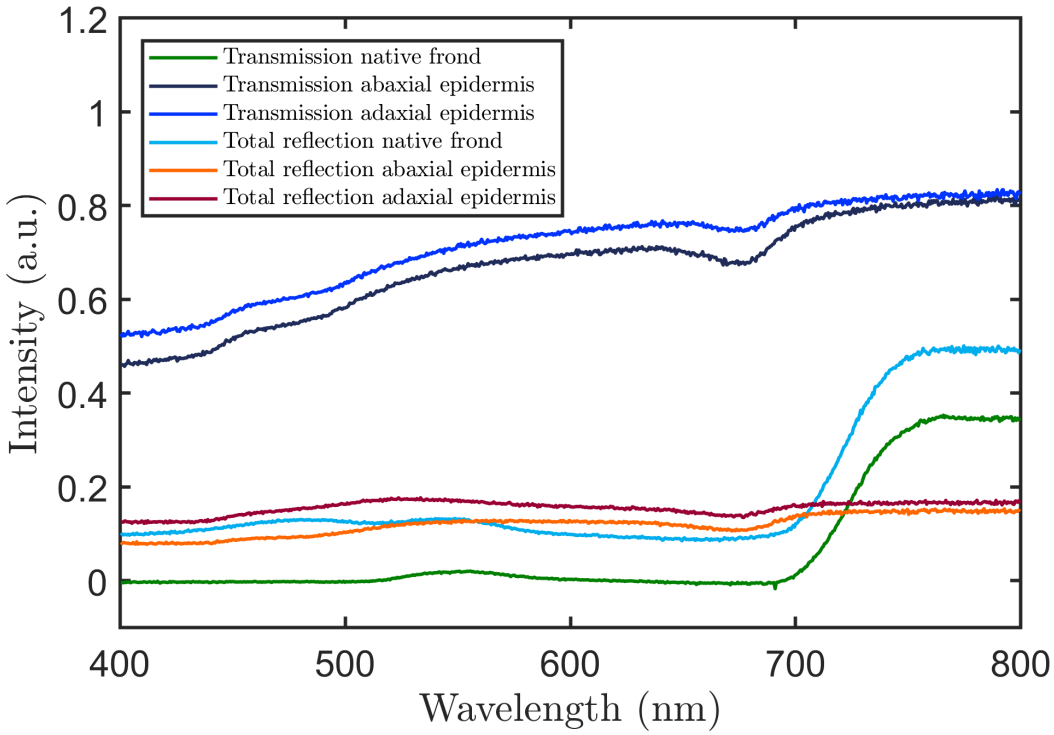


Fig. 3.10 Integrating sphere measurements.

3.5.5 Data fitting to Berreman 4x4 simulations

In Figure 3.11 some of the outcomes from trying to fit simple defect types to a reflection spectrum obtained from a cell on the adaxial surface of the frond is shown. As it can be seen, we do not manage to capture more than the effect of the main peak with this approach.

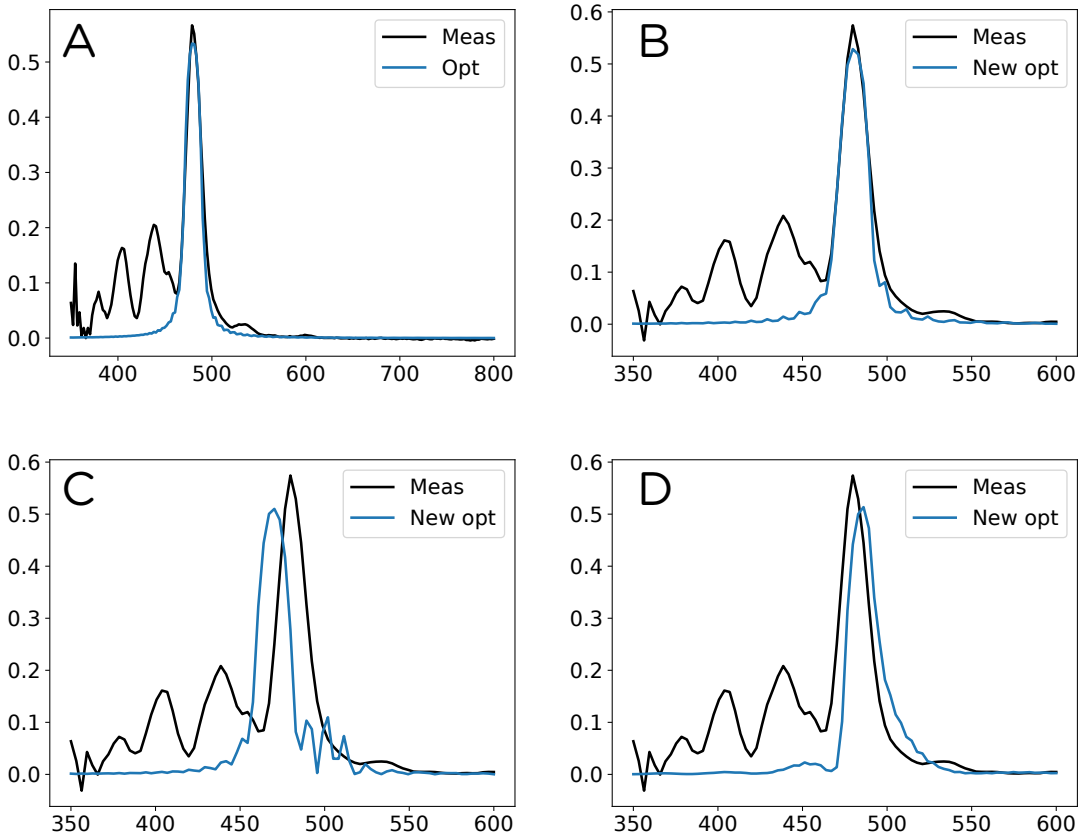


Fig. 3.11 Different results obtained during parameter fitting. A) Fitting of refractive index for a perfect helicoid 80 pitches thick. B) Fitting of a twist defect after 40 pitches and with fixed refractive indices. C) Fit with fixed refractive indices and same number of pitches with a twist defect in the middle and a different pitch length on each side. D) Same as in C, with the same total number of pitches, but freedom for the twist defect site to vary. None of these efforts captured any other effects than the main reflection peak.

Chapter 4

Margaritaria nobilis - materials and methods

4.1 Biomass composition

The different compounds in *Margaritaria nobilis* endocarp and leaves were determined according to the NREL protocols [128, 52, 123, 53, 124, 126, 125]. The overall scheme, adapted to the endocarp and leaves, is shown in Figure 4.1. The following reactions or operations were performed consecutively: aqueous extraction, ethanol extraction, acid hydrolysis and filtration, the reactions are depicted in the orange boxes. All the obtained products are depicted in round dark red ellipses, they comprise the extractives and the extracted tissue, the filtrate and the filter cake (which corresponds to the acid insoluble lignin). Furthermore, everything in square light blue boxes constitutes additional quantifications, like total solids content, ash content, and protein content, as well as carbohydrates quantification and acid soluble lignin quantification. For the total solids, ash and protein content, in the first instance (labelled 1), they were used to calculate the ash and protein content of the tissue, and their second and third measurements (labelled 2 or 3) were used to correct for their interference with later products, hence the double-pointed arrows. Finally, everything shaded in light blue is one of the final quantities obtained.

Both endocarp and (dried) leaves were ground to smaller than 2 mm and have less than 10 wt% moisture, so no additional preparation was needed as described in [52]. Protein content was obtained by determining the amount of nitrogen in the sample (departmental service). According to the respective NREL protocol [53], an appropriate protein conversion factor should be calculated to be more precise than multiplying it by 6.25. However, since the amount of N is very small in the leaves and negligible in the endocarp, this step was omitted

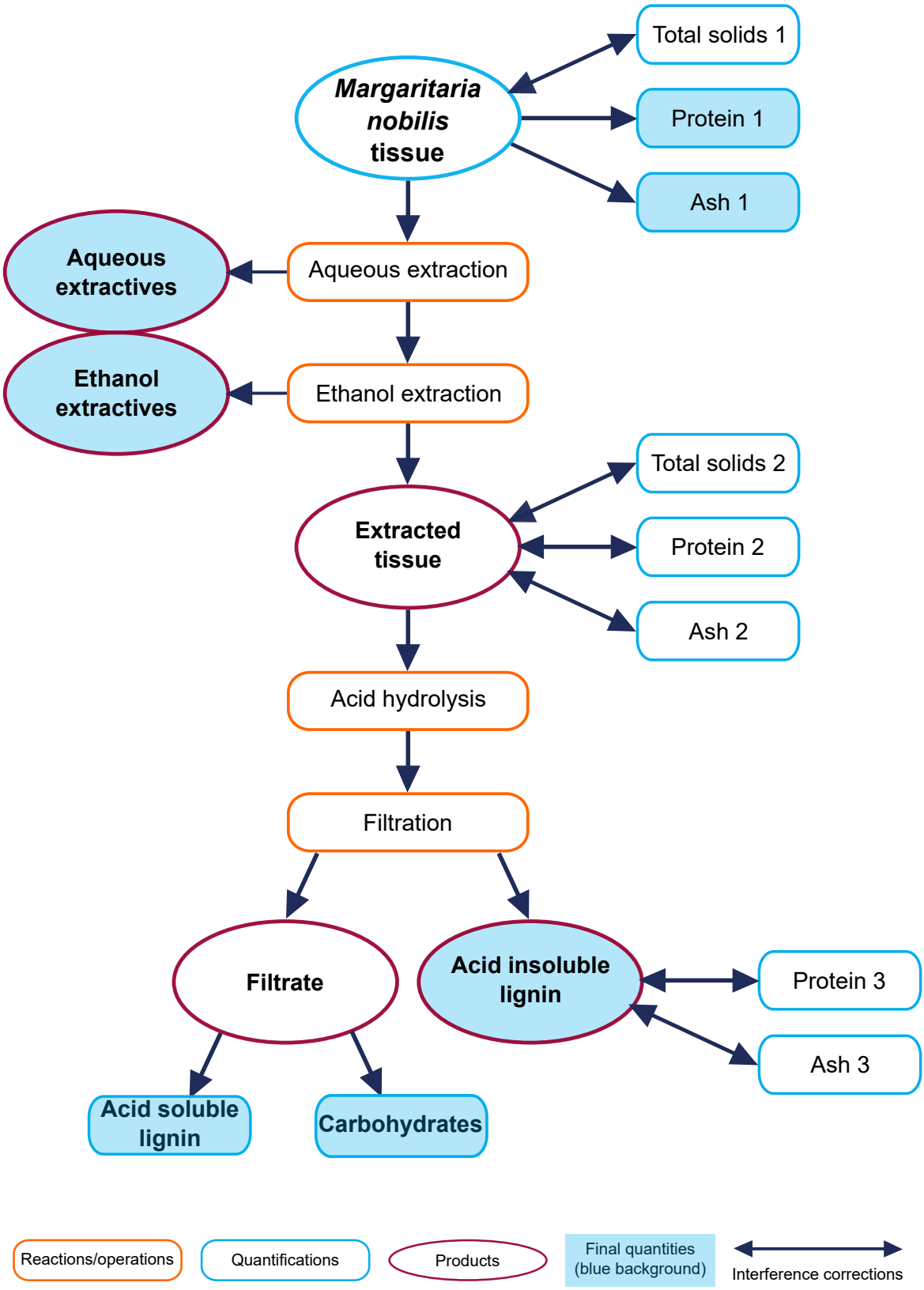


Fig. 4.1 Adapted scheme of quantification procedure of different compounds in *M. nobilis* endocarp and leaves according to NREL.

and the factor 6.25 was used instead. Ash content was also determined via the departmental microanalysis service [124].

Total solids content was determined by placing around 100 mg of sample in the oven at 105 °C for 4 hours, followed by cooling in the desiccator, another hour in the oven, cooling in the desiccator again, until constant weight was reached [123]. It is recommended to use 0.5 - 2 g for solid samples, however both leaves and endocarp are very homogeneous tissues, so we are confident in the obtained results, and no more endocarp sample was available.

Aqueous and ethanol extraction were carried out for both *M. nobilis* tissue samples [126]. 800 mg of sample were extracted with 190 ± 5 ml of water (milliQ) and subsequently ethanol in a Soxhlet extractor setup for 24 hours. (2 - 10 g of sample are recommended in the protocol.) Next, the solvent was removed on a rotary evaporator at 40 ± 5 °C, followed by placing the flask in the vacuum oven at 40 ± 5 °C for 24 hours, then cooling in a desiccator and weighing.

The acid hydrolysis consisted of a two step process, first 72 wt% sulphuric acid at 30 ± 3 °C for one hour, then diluting to 4 wt% acid concentration and hydrolysing in pressure tubes at 121 °C for another hour. 300 ± 10 mg extracted sample were used, and 3.00 ± 0.01 mL 72 wt% sulphuric acid, and for the dilution, 84.00 ± 0.04 mL milliQ were added. These quantities were adjusted if the sample amount differed. In the first hydrolysis step, the pressure tubes were left open and the reaction mixture continuously stirred. For the second hydrolysis step, a sugar recovery standard (SRS) was added in another tube, to correct the final results for degradation of monosaccharides.

After the hydrolysis, the sample was filtered. The filter cake was dried at 105 ± 3 °C for 4 hours, to constant weight, to determine the amount of acid insoluble lignin (AIL). The filtrate was used to determine the concentration of acid soluble lignin (ASL) via UV-Vis spectroscopy, and the concentration of various monosaccharides via HPLC and a differential refractive index detector.

For the ASL determination, the absorptivity was determined at 240 nm, where there is least interference from degraded monosaccharides, the filtrate was diluted to obtain an absorptivity value between 0.7 and 1, and 23 L/g cm was used as the absorptivity coefficient. An HPLC set-up with a Shodex sugar SP0810 column and guard column and a differential refractive index detector was used to determine the concentrations of sugars. The set-up was calibrated with D-glucose, D-xylose, L-arabinose, D-mannose, D-galactose and β-1,4-D-cellobiose. If significant amounts of cellobiose are discovered in the hydrolysed product, the hydrolysis is not extensive enough and the reaction time needs to be extended. At the same time, if too much of the SRS is degraded, hydrolysis time should be reduced. For our samples, neither was the case.

4.2 Cellulose

4.2.1 Isolation of native cellulose microfibrils

Cellulose microfibrils were isolated from the fruit endocarp. The isolation reactions were chosen based on a thorough literature search. The focus was on keeping the native dimensions and crystal structure of the cellulose microfibrils intact, while removing as much of the matrix as possible. Different procedures and reactions were tried out and evaluated, with the chosen optimised protocol listed in Table 4.1 below.

Since the CMF length was surprisingly short, I also isolated them from the leaves of *M. nobilis* with the same purification procedure, as a comparison to assure that none of the purification steps performed were disrupting the native morphology of the fibres. The following reactions were carried out:

Table 4.1 Purification steps to isolate native cellulose microfibrils

Purification step	**Abbreviation**	**Compounds removed**
Organic extraction	1OE	Lipids, waxes
Aqueous extraction	2AE	Water solubles, e.g. proteins
Acidic bleaching	3AB	Polyphenolics
Base treatment	4BT	Heteropolysaccharides
Acidic bleaching	5AB	Polyphenolics
Base treatment	6BT	Heteropolysaccharides
Acidic bleaching	7AB	Polyphenolics
Base treatment	8BT	Heteropolysaccharides
DMSO extraction	9DE	Heteropolysaccharides
Base treatment	10BT	Heteropolysaccharides
(DMSO extraction)	11DE	Heteropolysaccharides
Tip-sonication	12TS	Breaking up aggregates
Base treatment 4 wt%	13BT	Heteropolysaccharides

250 ml reaction mixture were used for both the endocarp and the leaves. The amount of starting material for the endocarp was (0.49 g for the first batch,) 5.06 g for the second batch, and for the leaves 10.71 g. Both tissues were not dried before starting, however the fruit were collected in Panama several weeks before the experiments, and they are a partially dehydrated tissue anyway, having a low water content upon maturation. The fruit endocarp was broken into small pieces with a mortar and pestle, and passed through a tea strainer with about 1 mm pore size. Ball-milling or any other harsher form of breaking down the tissue was omitted in order to not possibly reduce the crystallinity of the cellulose. The leaves were obtained fresh from the tree in the greenhouse in the Cambridge University Botanic

Garden, and chopped into small pieces with scissors. They contain much more water than the endocarp, so the amount of dry starting material should not differ too much between the two tissues, and ratios of reactants to plant material should be comparable.

All the reaction conditions are listed in Table 4.2.

Table 4.2 Reaction conditions of cellulose isolation procedure

Reaction	**Chemicals**	**Conc.** in wt%	**Temp.** in °C	**Time** in h	**Comment**
1OE	$C_6H_5CH_3 : EtOH$	NA	reflux	48	2:1 v:v
2AE	H_2O	NA	reflux	48	
3AB, 5AB, 7AB	$NaClO_2$	0.3	75	3	to pH 3 with CH_3COOH
4BT, 6BT, 8BT, 10BT	$NaOH$	2	80	2	
9DE	$DMSO$	NA	RT	48	freeze-dried, anhydrous, 5 wt% $LiCl$
11DE	$DMSO$	NA	RT	168	freeze-dried, anhydrous, 5 wt% $LiCl$
12TS	NA	1:10 dil.	RT		40 kJ/g
13BT	$NaOH$	4	80	2	

For the first two extractions, 1OE and 2AE, the crushed or chopped plant material was placed in a thimble and extracted in a Soxhlet extractor set-up. After the second extraction, the solid plant material was transferred to a round bottom flask for the first acidic bleaching step. After the reaction time of 3 hours, the heating was stopped and the reaction mixture flushed with N_2 to remove residual ClO_2 (which was trapped in a basic solution in a gas-washing set-up for safe disposal).

Then, centrifugations were performed to wash out the bleaching agent. The reaction mixture was transferred to 250 ml centrifuge bottles, centrifuged at 25,000 g for 30 min at 20 °C, then the supernatant was poured into the next centrifuge bottle, and the bottle with the solid plant material was filled up with milliQ and shaken. Then both bottles were centrifuged again at 25,000 g for 30 min at 20 °C, the supernatant of the second bottle was poured into a third bottle, while the bottom few milliliters of liquid and a bit of solid was combined with the solid plant material from bottle 1 again. The supernatant of the first bottle was again poured into the second bottle, and the combined solids from bottle 1 and 2 were mixed with fresh milliQ again. This procedure was repeated with 6 centrifuge bottles, so that the supernatant would always get centrifuged 6 times before disposal. Care was taken to not dispose of supernatant still containing cellulose, since this would shift the size distribution of cellulose

microfibrils to longer lengths and thus introduce a systematic error. The centrifuge washes were repeated until the supernatant of bottle 1 reached neutral pH, when all sodium chlorite and acetic acid were washed out. For the acidic bleaching reaction, it usually required about 12-15 washes to reach neutral pH.

Finally, the solid residue was transferred to a round bottom flask again, and without drying, the subsequent base treatment was carried out. Washing out base took significantly longer and required around 20-25 washes. I also tried to remove the base by dialysing against gradually weaker base or milliQ water, but by observing the tubes, NMR and TEM measurements, it seems that the base diffuses through the dialysis membrane so fast that the xylan dissolved in the basic solution does not have time to diffuse with it fast enough. Then the pH in the tubes sinks and xylan seems to come out of solution and regenerates in aggregates or recrystallises onto the cellulose bundles and microfibrils.

The small fragments of sample tissue started to disintegrate within the first acidic bleaching step, and became fully decomposed into small powders by the second round of reactions. After centrifuging, it was easy to mix the solids with fresh milliQ again, and no or only very few lumps much smaller than a millimetre were observed. Furthermore, I observed that the sample was plain white after the three cycles of bleaching.

Extensive NMR, XRD and TEM analyses were carried out after three cycles of acidic bleaching and base treatments (of the 8BT sample), since I expected to have removed all other compounds from the cellulose microfibrils, but there was still a significant amount of xylan present. I decided to carry out a DMSO extraction, for which the solid plant material was freeze-dried. Water was removed as much as possible from the solid plant material via centrifugation, and then the round bottom flask was immersed and swirled in an acetone-dry ice mixture to freeze it as fast as possible in a thin film, and freeze-dried, usually overnight. As opposed to air-drying or oven-drying, where the material sticks together in very hard lumps that almost cannot be broken again, freeze-drying produces a delicate film that can very easily be crumbled or chopped into smaller parts with a razor blade, and will disintegrate into a very fine powder very easily, for example when stuffing it into the ssNMR rotor. DMSO was either bought fresh or dried over a molecular sieve overnight, *LiCl* was dried in the vacuum oven overnight, the freeze-dried plant material was added, stirred vigorously and the container sealed tight. After 48 hours of extraction, the extraction mixture had turned into a translucent slurry. The DMSO was washed out via centrifugation as before, and after about 15 washes, the sample was put into dialysis against milliQ overnight, to remove the last traces of DMSO until there was no more odor.

Then, another base treatment was carried out. Xylan was already removed from the leaves sample at this point, the microfibrils were slightly aggregated, so I proceeded with tip-sonication. However, the endocarp sample still seemed to have a high content of xylan, as observed by STEM-SEM, so I decided to do another DMSO extraction for a week, which was omitted for the leaves due to time constraints and since the previous DMSO extraction had not shown any negative effect on the sample anyway.

For the tip-sonication, I had to dilute the endocarp sample ten times, since undiluted, it looked like it flocculated into floating lumps at shorter tip-sonication times, and turned gelly at very prolonged times. This effect was not observed for the leaves sample, so the dilution (and the subsequently necessary and time-consuming partial removal of water on the rotary evaporator at 50 °C and around 35 mbar) was omitted. The tip-sonication time, and hence the energy input, was determined for both samples individually, since it depends on the extent of aggregation of the microfibrils, by testing different sonication times on a small part of the sample and checking the morphology with TEM or STEM-SEM. Eventually, the longest possible time was chosen before any kinks were observed, so that as many aggregates could be broken up as possible without damaging the cellulose microfibrils. The energy input E (in J/g dry cellulose) is correlated to the tip-sonication time, amount and concentration of sample, and power of the tip-sonicator via:

$$E = \frac{P \cdot t}{m_{cell} \cdot w} \tag{4.1}$$

with P being the power of the tip-sonicator in W (set to 150 W), t the sonication time in s, m_{cell} the mass of the cellulose-water mixture, and w the cellulose concentration in weight%. For the leaves, an energy input of around 24 kJ/g was chosen, and around 40 kJ/g for the endocarp sample. These energies are estimates since the concentration of the mixture was not determined gravimetrically to save sample material. Finally, the last base treatment at 4 wt% NaOH concentration was carried out.

4.2.2 Chemical analyses

NMR measurements

Solid state magic angle spinning cross polarisation NMR (ss MAS CP NMR) was carried out on the solid plant samples. Sample preparation always consisted of cutting the dried starting materials or freeze-dried sample material into a fine powder with a razor blade or grinding it with mortar and pestle. Too vigorous ball-milling was avoided to not decrease the cellulose microfibrils' crystallinity. Then, the sample was dried in the desiccator overnight and finally tightly packed into a 4 mm outer diameter zirconia rotor (Bruker, Karlsruhe, Germany). The

peak of the α-carbon of glycine is set to 43.1 ppm. For the plant samples, a spinning rate of 14 kHz was chosen.

The assignments of compounds in the NMR spectrum are shown in Figure 4.2. Basically, hemicellulose and cellulose have peaks in the same range of chemical shifts, between 60 and 110 ppm, for the carbons in the sugar ring, but hemicellulose has additional peaks at 20 ppm for the methyl group of a ketone or ester, and a peak for the carboxylic acid carbon at 170 ppm. Both ketone/ester and carboxylic acid functional groups only occur in hemicellulose, but not in cellulose, for example as acetyl groups substituting xylose on the C-2 or C-3 ring carbon in the xylan backbone, or as the carboxylic acid of glucuronic acid substituting the xylan backbone. Furthermore, various peaks between 110 and 160 ppm can be observed, which are caused by the carbons in the aromatic ring. These peaks are indicative for compounds such as lignin and tannins, as well as the peak between 50 and 60 ppm, which is the carbon of a methyl group of a phenyl ether, a typical motive in lignin. On the right, the peak assignments of the carbons of the glucose ring of cellulose are depicted in detail.

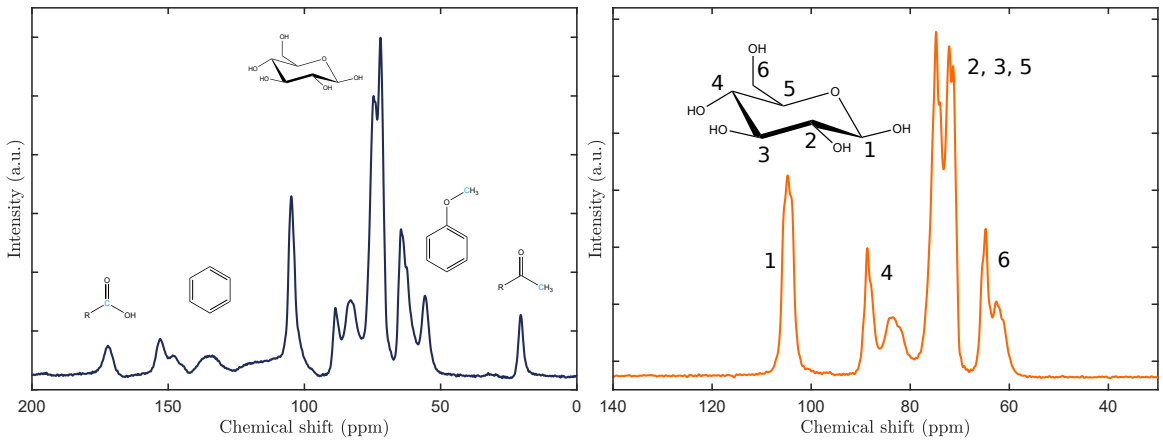


Fig. 4.2 NMR spectra of *M. nobilis* endocarp with peak assignments of plant tissue components in the native endocarp on the left, and peak assignments of the carbons of the glucose ring of cellulose on the right.

A ^{1}H - ^{13}C cross polarisation (CP) contact time suitable for plant samples had to be found, so a series of experiments with increasing CP times was run. 2500 μs was determined as the most suitable CP time: At shorter CP times, see Figure 4.3 on the left, all signals are weak, at 1000 μs, the cellulose signals are already good, but most other signals are low. At longer CP times, see Figure 4.3 on the right, cellulose signals start to decrease, while the rest stays at the same intensity.

Furthermore, experiments were carried out with 2 s relaxation time, and at least approx. 30000 scans were accumulated per spectrum (overnight).

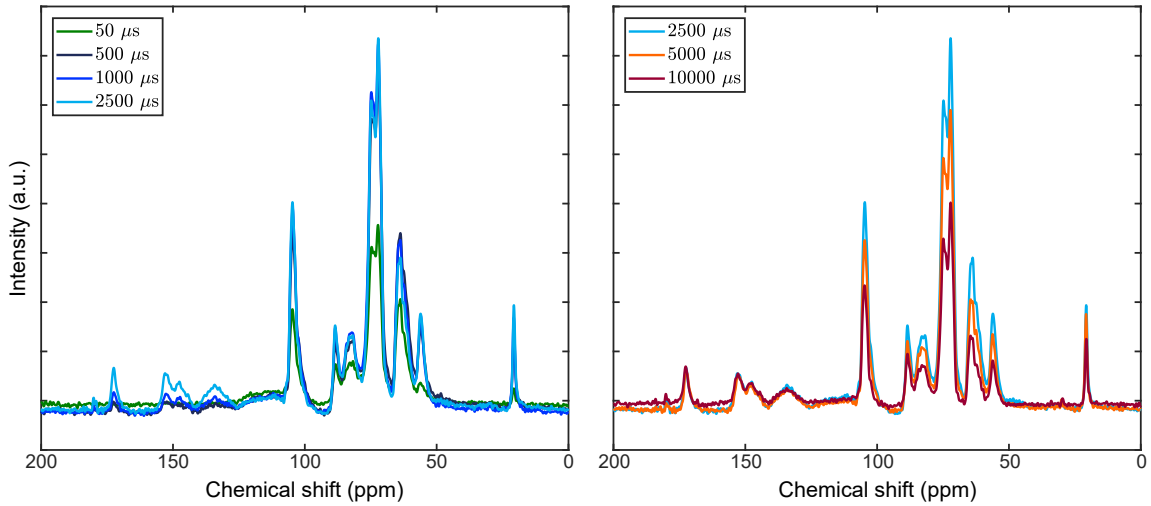


Fig. 4.3 NMR spectra of *M. nobilis* endocarp at different CP times.

XRD measurements

Powder X-ray diffraction (XRD) measurements were carried out on a PANanalytical XRD. In powder diffraction, it is assumed that each crystal orientation is represented equally, i.e. no preferential orientation is observed. In order to meet this assumption, the sample stage was rotated at a revolution time of 1 s (in reality, there is preferential orientation). Sample preparation was the same as for NMR spectroscopy, drying overnight in the desiccator, then samples were placed on the sample holder, compressing the material and smoothing the surface. The range of 2θ was chosen between 5° and 40°, scan step size was 0.01° and time per step was 40 s. The operating voltage was set to 40 kV and the filament current to 40 mA.

4.2.3 Microscopy

Most microscopy techniques are described in chapter 3. Embedding and transmission electron microscopy were carried out by a collaborator, Dr. Yu Ogawa.

SEM measurements

SEM measurements were carried out on a TESCAN MIRA3 FEG-SEM system with an SE detector (Everhart-Thornley Type), and in-Beam SE detector (TESCAN ORSAY HOLDING, a.s., Czech Republic). Images were taken in top view or in cross section at an angle of 90° with respect to the electron beam. The samples, usually small fragments of dried plant material, were placed on an aluminium stub with conductive carbon tape and sputter-coated with a thin layer (5 nm) of platinum, with a Q150T ES Turbo-Pumped Sputter Coater/Carbon

Coater (Quorum Technologies, UK). An accelerating voltage of 0.8 kV to 1.4 kV was used, and the working distance was typically 3 mm.

STEM measurements

Scanning transmission electron microscopy (STEM) was performed on the same SEM set-up, with a retractable STEM detector in bright field mode. Sample preparation was the same as for TEM imaging. Continuous carbon copper grids were glow-discharged with a GloQube Dual Chamber Glow Discharge System (Quorum Technologies, UK). Then, a drop of diluted sample (approximately 0.01 wt% concentration) was deposited on the grid and blotted with filter paper after two minutes. Subsequently, a drop of staining solution (1 wt% aqueous uranyl acetate) was deposited, and again blotted after two minutes, and the sample was left to air-dry. The samples were imaged at 30 kV accelerating voltage and approximately 5 mm working distance.

4.3 Xylan

The main method to analyse xylan used in this project is PACE - polysaccharide analysis by carbohydrate gel electrophoresis [48]. In the following section the experimental steps performed for the xylan analysis are explained in detail. In short: (i) xylan is made accessible in the biomass by breaking up the tissue and removing other compounds, (ii) then extracted via base extraction, or other compounds removed via depectination and delignification, (iii) enzyme hydrolysis is carried out depending on the specific question, to obtain xylose and xylan oligosaccharide fragments, (iv) the obtained fragments are labelled with a fluorophore on the reducing end to enable their detection, and (v) finally, the fragments are sorted by size and charge during gel electrophoresis and then imaged and interpreted.
Furthermore, additional approaches were used, like so-called 'sequencing experiments' and mass spectrometry (MALDI-ToF MS).
The xylan analysis was carried out in collaboration with the Paul Dupree group, Department of Biochemistry, University of Cambridge. All the procedures and protocols were developed or tested in their lab over decades of dedicated research. They are specialised on xylan analysis, where different aspects are important compared to the CMF isolation procedure I developed for this project. For example, base concentration was kept low for CMF isolation, since at high base concentration, the Iβ allomorph can be converted to cellulose II (mercerisation), and it was essential to keep the native crystalline structure. At low base concentrations, hemicellulose is not as soluble, which is the reason why the base treatment had to be repeated many times. For the xylan analysis, on the other hand, mercerisation of cellulose is irrelevant,

so that higher base concentrations for xylan extraction can be used without any problem. While for the CMF isolation, NaOH concentration was 4 wt% maximum, it was 4 M for the xylan extraction, which corresponds to 16 wt%, and would most likely lead to mercerisation. At the same time, it is much more effective at extracting hemicellulose, and it is assumed that all xylan is extracted this way. Furthermore, the CMF isolation procedure was carried out at the gram scale, to obtain enough material for e.g. NMR and XRD analysis, while it was possible to work in the milligram range for the xylan analysis, which also influenced the choice of methods and equipment. In this section, quantities or concentrations reported in % are always by volume.

4.3.1 Alcohol insoluble residue

Alcohol insoluble residue (AIR) from the plant tissue was always prepared as starting material for any xylan analysis: (i) the material was first heated to render any enzymes that it contains inactive, (ii) then it was ground in a ball mill and (iii) subjected to a series of solvent extractions. In these solvent extractions, soluble compounds like fats and lipids from the cuticle are removed. This procedure decreases interference from non-polysaccharide components and makes the xylan more accessible by breaking down the tissue.
In more detail, after heating to 70 °C in 96 % ethanol for 30 min to inactivate any enzymes, the tissue-ethanol mix was homogenised in a ball mill (Mixer Mill MM 200, Retsch GmbH, Germany). Then, the sample was centrifuged at maximum speed for 15 min and the supernatant pipetted off. This step was repeated after every subsequent extraction, as listed in Table 4.3. Thereafter, the AIR was air-dried, diluted to the required concentration, aliquotted and dried again [14].

Table 4.3 Overview of all extraction steps in the preparation of AIR

Step	**1**	**2**	**3**	**4**
Solvent	EtOH	EtOH	MeOH:$CHCl_3$	MeOH:$CHCl_3$
Concentration	96 %	100 %	2:3 v/v	2:3 v/v
Time	15 min	15 min	o/n	1 h
Step	**5**	**6**	**7**	**8**
Solvent	EtOH	EtOH	EtOH	EtOH
Concentration	100 %	65 %	80 %	100 %
Time	15 min	15 min	15 min	15 min

4.3.2 Xylan extraction

Once the AIR is prepared, xylan is extracted. Two different methods were used, depending on the experimental question.

Extraction with base

The first method is a simple extraction by base: 4 M NaOH was added to the AIR and xylan was extracted at room temperature (RT) for 30 min, and subsequently neutralised with 1 M HCl. This approach is fast and simple, however it also removes any possible acetylation of the xylan backbone. Therefore, this method is suitable when an undecorated chain is required when using certain enzymes (like for example GH10), but obviously, it does not provide any information about the acetylation pattern.

Isolation of acetylated heteroxylan

Using this method allows to maintain the acetylation of xylan intact. The first step is depectination and is optional. Even though pectin was not expected to be found in the secondary plant cell wall in large amounts (and indeed rhamnose concentration was found to be negligible, see section A.4), it was still carried out on all samples. 0.5 % ammonium oxalate ($(NH_4)_2C_2O_4$) was added to AIR, and shaken at 85 °C for 2 h, followed by centrifugation and washing with water. The second step was delignifaction, where the AIR was shaken with 11 % peracetic acid (CH_3COOOH) at 85 °C for 25 to 30 min, then cooled down on ice and washed with water [18, 46].

4.3.3 Enzyme digestion

Ammonium acetate buffer (AmAc, pH5.5, 50 mM or 100 mM) and the respective enzyme were added to the extracted xylan for digestion. Generally, 1 mg AIR was processed, in 500 μL buffer. A range of enzymes was used to digest xylan, depending on the specific question:

- xylanases GH10, GH11 [107, 141]
- glucuronoxylanase GH30 (BO and CE) [139]
- arabinosidase GH62, glucuronidase GH115 [20, 156, 114]
- xylanase GH3, glucuronidase GH67 [71, 114]

Each of these enzymes has a very specific mode of action and will only (ideally) cleave the xylan backbone or the decorations at very specific locations, thus enabling identification of the fragments constituting the chain.
To illustrate the specific mode of action of each enzyme, a short example of a hypothetical xylan chain was drawn, see Figure 4.4. The molecular structure is shown, with the respective abbreviations with configurations and bond conformations, and a simplified scheme, that will be used in all the subsequent cartoons. The systematic nomenclature proposed for xylo-oligosaccharides by Faure et al. is followed [37].

Figure 4.5 depicts the simplified scheme of this short example xylan chain in (A), and the blue arrows indicate the position where the specific enzyme, GH10 in this case, will cleave the molecule. Figure 4.5 B reports the generated fragments that the enzyme cleaved from the short example xylan chain in A. The obtained PACE gel is depicted in (C), gel interpretation is discussed below. Cartoons depicting the mode of action, fragments and PACE gels for the other enzymes are shown in the appendix section A.4.

The incubation conditions for each enzyme are listed in Table 4.4.

Table 4.4 Incubation conditions for all the enzymes

Incubation	Time	Temperature
		°C
GH10	overnight	37
GH11	overnight	45
GH30	30 min	RT
GH62	overnight	RT
GH115	overnight	57
GH3 xylanase	overnight	37
GH67	overnight	37

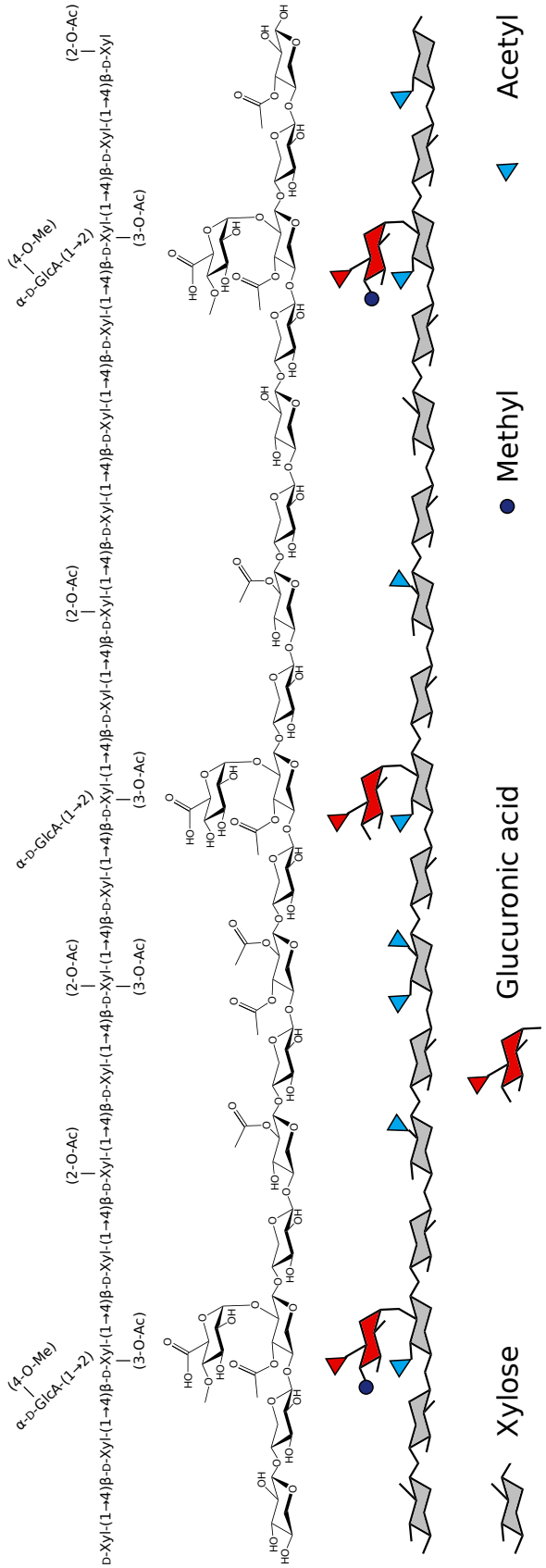


Fig. 4.4 Abbreviations with configurations and bond conformations, molecular structure and simplified scheme of a hypothetical short example xylan chain.

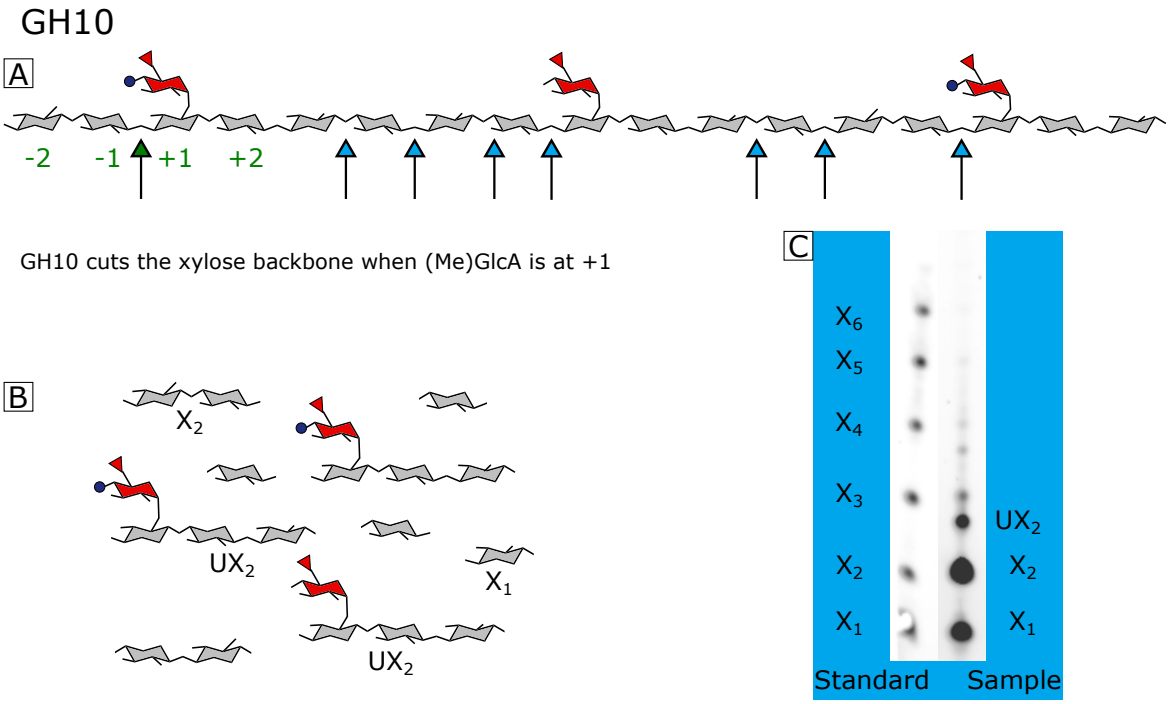


Fig. 4.5 (A) Short example xylan chain. Arrows show the specific site of cleavage for GH10. (B) Fragments obtained from digesting the short example xylan chain with GH10. (C) PACE gel showing all the characteristic fragments of GH10 digestion.

4.3.4 Labelling

In order to visualise the different fragments of monosaccharides and oligosaccharides created by the various enzyme digestions, the reducing ends were derivatised with a fluorophore, 8-aminonaphthalene-1,3,6-trisulfonic acid (ANTS) [48]. To that end, a 2:1:1 labelling master mix was created from a DMSO buffer, the ANTS solution and a 2-PB solution.

- DMSO buffer: 30 μl acetic acid, 170 μl H_2O and 200 μl DMSO
- ANTS solution: 0.2 M ANTS (molecular probes, A350) in H_2O:acetic acid 17:3
- 2-PB solution: 0.2 M 2-picoline-borane (2-PB) in DMSO

The ANTS labelling master mix was added to the digested sample (20 μl per 1 mg AIR) and incubated at 37 °C overnight, then dried in the speedvac. The mechanism is shown in Figure 4.6, F stands for fluorophore, ANTS in this case [48]. Instead of sodium cyanoborohydride, $NaCNBH_3$, picoline-borane was used as the reducing agent.

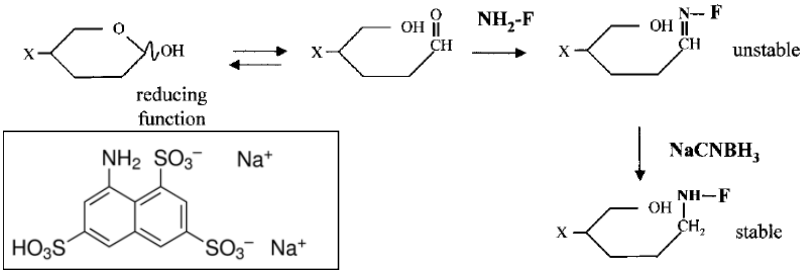


Fig. 4.6 Scheme depicting labelling of xylan fragments with a fluorophore (F), ANTS is used in this case, shown in the inset. Monosaccharide - X: OH, oligosaccharide - X: sugars. Reproduced with permission from [48].

4.3.5 Running a PACE gel

Finally, electrophoresis in a polyacrylamide gel was carried out. The gel was prepared by mixing the reagents in Table 4.5 and quickly pouring them between two glass plates assembled in a rig, inserting a comb at the top to make the wells where the samples will be deposited. The gel was then left to settle, usually in the cold room (4 °C) overnight.

Table 4.5 Reagents to make up a PACE gel

Reagent	Amount
H_2O	20 ml
1 M Tris-Base Borate (pH 8.2)	3 ml
40 % polyacrylamide (29:1; acrylamide/bis)	8.4 ml
TEMED (tetramethylethylenediamine)	24 μl
10 % (w/v) APS (ammonium persulfate)	180 μl

To run the gel, 3 M or 6 M urea was added to the dried, labelled samples to resuspend the xylan fragments. The wells of the gel were filled with 0.1 M tris-borate buffer, and the samples were loaded into the wells. The gel was run for 40 min at 1000 V. Then, the gel was visualised using a Syngene G-box, fitted with long wave UV tubes (emitting at 365 nm) and a short pass detection filter (500–600 nm).

4.3.6 Assignment of bands

A typical PACE gel for digesting xylan with GH10 is shown in Figure 4.5 C. In the left column, a standard of xylose to xylohexaose was run to be able to assign the bands observed in the actual sample. Xylose, X_1, is the smallest and was transported the furthest, it is thus visible at the very bottom of the gel, while xylohexaose, X_6, is the largest, and was found at the very top. Some of these fragments, X_1 and X_2, were also observed in this sample.

Additionally, the fragments are separated by charge: if a fragment carries one glucuronic acid, GlcA or U, that can provide a negative charge from the carboxylic acid group on the GlcA C_6, it will run further than a fragment without a GlcA. Similarly, fragments with more GlcA will run further than fragments with fewer GlcA. In this sample, there was a band for UX_2 visible, underneath the band for X_3 of the standard. Both UX_2 and X_3 have the same amount of xylose monomers, but UX_2 carries one negative charge, and ran thus further down on the gel.

To sum up the assignment of bands, positions of fragments are compared to a standard consisting of the (oligo-)saccharides xylose to xylohexaose, X_1 to X_6. The smaller a fragment, the further it will travel and thus the further down it will be on the gel. Secondly, the more charged a fragment is, the further down it will travel, so for example UX_2 will be below X_3, but above X_2 on the gel.

4.3.7 Quantification of glucuronic acid content

The glucuronic acid content was determined via two methods: quantification on the PACE gel, and monosaccharide analysis. For the first method, the ratios of different fragments to each other can be quantified using the Fiji plug-in in the ImageJ software [120, 119]. It is important that the exposure time of the detector is chosen so that the bands of interest are not saturated while still giving enough signal. Then, the intensity of greyscale and area are calculated, and different bands on the same gel can thus be compared to each other. The procedure and calculation is shown in more detail in the appendix, section A.4.

For the second method, mild acid hydrolysis with trifluoroacetic acid (TFA) and high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) was carried out by Henry Temple, to determine the monosaccharide sugar composition of hemicelluloses contained in the sample, and especially the amount of glucuronic acid. The hydrolysis step is not harsh enough to decompose cellulose [14]. More details are found in the appendix, section A.4.

4.3.8 Sequencing experiments

This type of experiment is a very new approach developed by the group members in the Dupree group at the Department of Biochemistry, University of Cambridge, and has yet to be extensively tested and published. So far, I used only one or two steps of enzyme digestions per sample, and information about how the observed fragments are arranged along the xylan molecule were lost. With this type of sequencing experiment, subsequent digestions of the same sample were carried out. Rather than labelling the fragments after the enzyme digestion,

the intact xylan molecule was labelled before digestion, and all surplus labelling agent was washed out. Washing out is achieved by quenching surplus fluorophore with glucose at 37 °C for 1 hour, then precipitating xylan in 96 % EtOH, at -20 °C for 20 min, and then centrifuging at 4 °C and removing the supernatant that still contains glucose with surplus fluorophore in solution. Then, 96% EtOH was added again, followed by xylan precipitation, centrifuging and pipetting off the supernatant. This procedure was repeated until no more fluorophore was observed in the supernatant with the Syngene G-box (at least 8 times). This way, only the long main xylan chain will be visible on the PACE gel, but not the hydrolysed fragments, since there was no surplus labelling agent still present or added that could label freshly cut oligosaccharides at their reducing end.

20 mg of AIR were treated this way, and after the final EtOH removal and drying, 1200 μL of ammonium acetate buffer were added. Alternating digestions with GH3 (1 μL) and GH67 (4 μL) were completed, then an aliquot corresponding to 1 mg AIR (60 μL) was taken, and the sample was heated to 100 °C for 20 min after each run to inactivate the previous run's enzyme, before adding the new enzyme. All the aliquots were stored at -20 °C, then dried in the speed-vac, resuspended in 20 μL 6 M urea, and 5 μL each were brought on the gel. The gel was run as usual, 1000 V for 40 min. Exposure time when imaging it had to be increased, since signal intensity was lower than usual.

4.3.9 MALDI-ToF MS

For the MALDI-ToF MS experiment, 10 mg xylan AIR was subjected to the isolation of acetylated heteroxylan procedure and digested with GH10. Then, the supernatant was filtered with nanosep 10K cut-off filters (Pall Corporation, USA) at maximum centrifuge speed for 20 min. After drying in the speed-vac, the sample was resuspended in 50 μL milliQ and diluted by 1:5. 2,5-dihydroxybenzoic acid (DHB) matrix (10 mg mg/mL dissolved in 50 % MeOH) was added in a 1:1 ratio, 1 μL spotted on the plate and quickly crystallised in a desiccator to obtain a small crystal size for better evaporation. The samples were analysed via MALDI-ToF MS with a 4700 Proteomics Analyser (Applied Biosystems, USA) equipped with a 200 Hz frequency triple Nd-YAG laser operating at a wavelength of 355 nm (with support from Theodora Tryfona) [83].

4.4 Small-angle X-ray scattering

X-ray scattering experiments were carried out by Dr. Yu Ogawa at the beamline D2AM at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. X-rays of 18 keV (λ = 0.6888 Å) with a spot size of 100 μm were used. The X-ray beam was incident tangential to the endocarp surface to obtain the structural information from the cross section of cells. The two-dimensional scattering pattern was recorded on a photon-counting pixel detector at a camera distance of about 800 mm. The equatorial anisotropic intensity components were extracted according to the method by Nishiyama et al. [99].

4.5 Coarse grain molecular dynamics modelling of cellulose-xylan assembly

Coarse grain (CG) molecular dynamics modelling with Martini force field was employed to simulate the assembly of cellulose microfibrils and xylan molecules, since the system size is too big for all-atom (AA) simulation. Force field parameters for cellulose and other saccharide molecules are available in the literature [80, 159], while they had to be developed for xylan: they were optimised against the structure simulated by AA simulation. The conformational parameters had to be set by defining the CG beads and by adjusting the angular and dihedral parameters, see Figure 4.7 A.

All the atoms in the xylose molecule were 'coarsened' into larger portions, the coarse grain beads B1, B2 and B3, as depicted by the black outlines. Then, the bonded interactions of these beads, defined as angular and dihedral parameters, were optimised. While the angular parameters were straightforwardly determined based on molecular geometries in AA simulations, the dihedral parameters required substantial optimisation efforts. In total, six dihedral angles were calculated for the coarse grains, and adjusted to the results from the AA simulation. The dihedral angles closely related to chain conformation, B1-B2-B1'-B2' and B1-B2-B1'-B3', are depicted below. The B1-B2-B1'-B2' angle describes the angle of the glycosidic bond, and hence the orientation of two xylose monomers to each other, which defines in which conformation the chain is. The dihedral angle is found to be around 100° from the AA simulation (carried out in water), which roughly corresponds to a three-fold helix (which would be 120° [18]), and the CG angle was adjusted accordingly. The B1-B2-B1'-B3' angle describes the orientation of the O2 and O3 hydroxymethyl groups with respect to the xylan backbone. The other four dihedral angles are shown in the appendix, section A.5. (The dihedral angle plots are histograms: the xylan molecule in the CG description was solvated and equilibrated for 20 ns in a CG water box, and each dihedral angle distribution

was extracted from the last 10 ns of the molecular dynamics trajectory.) The CG dihedral angle distributions were compared to those from the AA simulations, and the parameters were optimised to minimise the difference between the two distributions.

Once all the CG model parameters were adjusted so that the dihedral angles corresponded to the ones calculated from the AA simulations, the non-bonded interaction parameters between cellulose and xylan were determined. Again, this was first carried out for the AA model (Figure 4.7 B), and then the CG results were scaled to it (Figure 4.7 C). This was done via potential of mean force (PMF) curves (pull code + umbrella sampling): a xylan chain was brought in close proximity to a cellulose chain and equibrilated (again in water), where the xylan chain absorbed onto the cellulose surface. Then, the xylan chain was pulled off at one end (schematic shown in B on the top), and the free energy change caused by the manipulation, depending on the distance between the cellulose microfibril surface and the xylose monomer where the force was applied, was monitored, depicted in Figure 4.7 B on the bottom. (In more detail, the distance was defined as the distance between the centre of mass of the topmost glucose chain to the centre of mass of the encircled xylose monomer, with the absorbed equilibrium state set to zero, and the orientation of the cellulose crystal fixed perpendicular to the pulling direction.) For the AA simulations, this experiment was carried out for the different cellulose surfaces 100 (for a two-fold and three-fold xylan helix, 2_1 and 3_1), 110 and $1\bar{1}0$. Then, the same experiment was carried out for the CG model (Figure 4.7 C, red line), and then scaled to the AA result (C, blue line). It was computed for the 110 surface, since this is expected to be one of the most exposed surfaces of the cellulose microfibril, together with $1\bar{1}0$, sharing the hydrophilic nature [59]. Finally, the CG parameters of xylan were defined, and the actual simulation of the arrangement of the helicoidal architecture could be started.

All of this modelling work was carried out by Dr. Yu Ogawa.

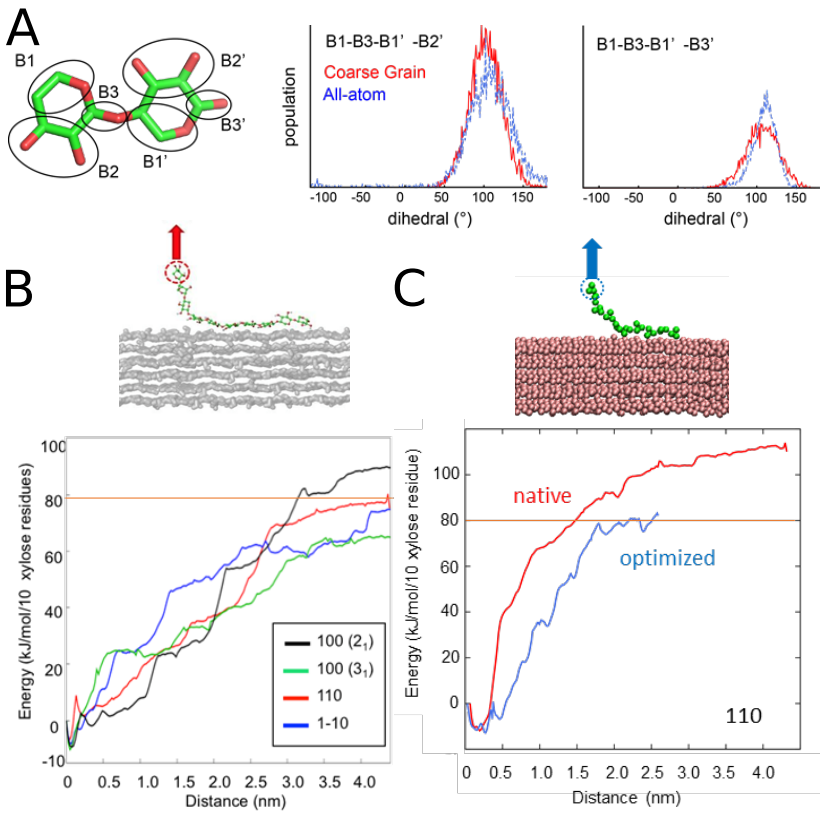


Fig. 4.7 Martini coarse grain (CG) force field modelling: (A) Definition of coarse grain beads and adjustment of angular and dihedral parameters. (B) Optimisation of interactions based on PMF curves via AA simulations. (C) Scaling of CG interaction terms to AA results.

Chapter 5

Margaritaria nobilis - results, discussion and conclusion

Margaritaria nobilis is a rosid eudicot tree growing in Central and South America. It grows in secondary forests, in tropical and seasonal forests, as well as evergreen forests. They grow fruit of approximately 10.9 mm diameter and 6.14 mm length and with a total fruit mass of 0.49 g [23, 145]. The fruit are covered with a green exocarp, which dehisces at maturity, over the structurally coloured endocarp, see Figure 5.1 A. The fruit contain four to six segments, most commonly five, each consisting of a single dark seed. Upon drying, the fruit lose the macroscopic blue coloration, looking whitish pearlescent, but the endocarp still exhibits the same structural colour under the microscope. Figure 5.1 B shows photos of fruit subjected to increasing dehydration from left to right. This effect is caused by the seed shrinking upon dehydration, thus introducing a small air gap between the seed and the endocarp, as depicted in the scheme in Figure 5.1 C. In the hydrated state (top), there is no interface between the endocarp and the seed, and the dark pigments in the seed absorb all the transmitted light. When dry (bottom), light is scattered back at the endocarp-air interface, which leads to all wavelengths being reflected, thus resulting in an overall whitish macroscopic appearance.

5.1 Optical properties and ultrastructure

The optical properties and ultrastructure of *M. nobilis* were thoroughly characterised in Vignolini et al. [145].

Figure 5.2 shows a photo of a mature, hydrated fruit, exhibiting strong blue structural coloration. This appearance can be observed when subjecting the endocarp to reflection

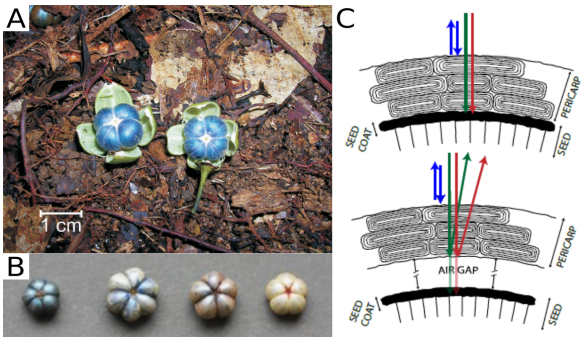


Fig. 5.1 *Margaritaria nobilis*. (A) Photo of fruit clearly exhibiting the green exocarp and the structurally coloured endocarp. From [23]. (B) Photo of fruits under increasing dehydration from left to right. (C) Scheme of hydrated (top) and dried (bottom) fruit endocarp and seed. B-C from [145].

microscopy. Intense colour was only visible in the left-handed circularly polarised (LCP) channel. The micrograph of the right-handed circularly polarised (RCP) channel of the same area only looks slightly brownish, the dry endocarp probably containing some tannin and lignin, but no structural colour was observed. Spectra can be collected from a small spot from these micrographs, by coupling a spectrometer via an optical fibre through the microscope set-up. The spectrum of the LCP reflection shows a strong peak with a maximum reflection wavelength of approximately 450 nm, as observed in the LCP micrograph, but is almost flat for the RCP reflection.

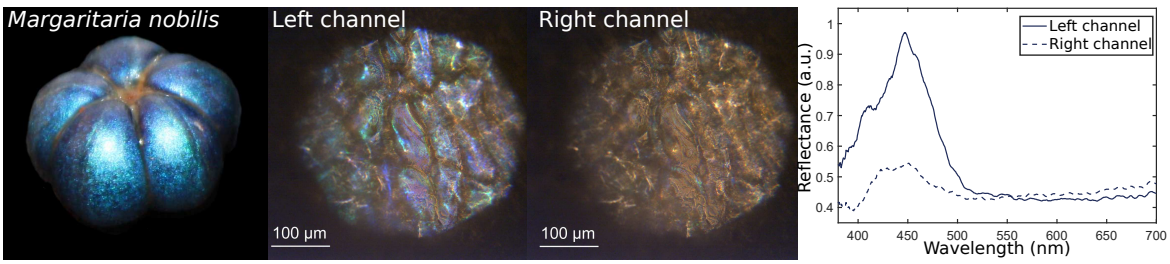


Fig. 5.2 Optical properties of *M. nobilis*: Photo of fruit. LCP and RCP micrographs of endocarp. Reflection spectra of a small spot of the LCP and RCP micrographs.

This structural colour arises from the helicoidal arrangement of cellulose microfibrils in the secondary cell walls of the endocarp.

The ultrastructure of the endocarp is shown in Figure 5.3. It is made up of the cuticle and three to four layers of these helicoidal cells. When zooming in further, one can see that the cells consist entirely of the thickened cell wall and there is no cell content left. These secondary cell walls, upon zooming in further, reveal to consist of a layered structure, and the

layers clearly consist of fibres, more or less parallel aligned, and even the rotation between the layers can be seen. Thus, a helicoidal architecture with its helicoidal axis perpendicular to the fruit surface is responsible for the structural colour.

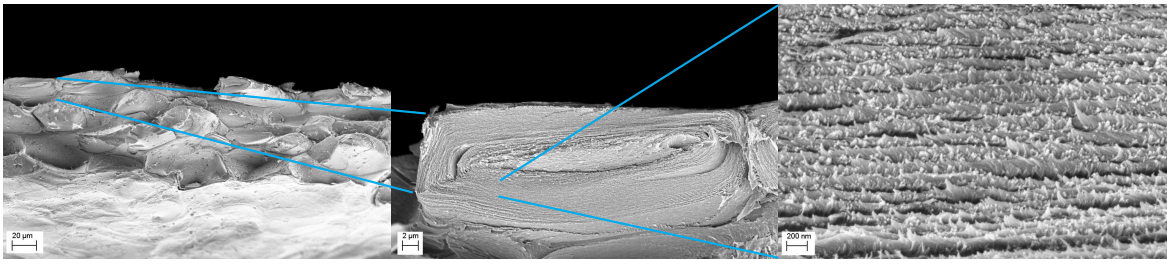


Fig. 5.3 SEM images revealing the ultrastructure of the *M. nobilis* endocarp with higher magnification from left to right.

In this chapter, I present all the investigations and results about what building blocks this structure is made of, and what their properties are.

I also carried out the same investigations for the leaves of the same plant. The leaves exhibit no structural colour at all, and are thus not interesting from an optical point of view. However, all plant material and its components is subject to biological variation, and properties like e.g. cellulose microfibril dimensions vary significantly between species. It is hard to define absolute values for these dimensions, and there are also many different approaches to assess them, which can have an influence on the results. Furthermore, some quantities like e.g. the crystallinity index, are not absolute values, but should rather be considered a way of comparing different materials to each other, meaningful only if the same method was used to assess it. These factors make it difficult to compare obtained results with the literature, which is why I decided to use the leaves as a comparison. Furthermore, some results for the endocarp were quite surprising, hence carrying out the same procedures on the leaves also served as a validation that I did not alter or influence the measured sample properties.

5.2 Biomass composition

In order to understand what components the endocarp contains, the biomass composition was analysed.

The biomass composition of the endocarp and the leaves was determined via the NREL protocols [128, 52, 123, 53, 124, 126, 125]. These protocols work well for woody and

herbaceous materials, so were considered appropriate for the endocarp, which is a lignified secondary cell wall, and the leaves, which can be considered a herbaceous material.

The analysis flow was adapted according to the protocols, e.g. herbaceous material was submitted to an aqueous extraction as recommended. Furthermore, leaves are expected to contain a high amount of protein, which can later interfere with the determination of acid insoluble lignin, so it was mathematically corrected for. The protein conversion factor was used as 6.25, rather than determining it depending on the precise amino acid composition, to reduce time and effort, and since protein content was very low for the endocarp anyway. High moisture or ash content are not an expected problem for either of the samples, which would interfere with the acid hydrolysis step. For the determination of acid soluble lignin, an extinction coefficient of 23 L/g cm was chosen by averaging the suggested values for different types of biomass. Even though acetyl content should have been determined, no appropriate HPLC column was accessible and this step was therefore omitted. However, it was considered in the more detailed xylan analysis. Starch, on the other hand, was not expected to be present in the endocarp. It is likely to be present in the leaves, depending on the time of day that the leaves were harvested: generally, leaves produce starch during the day while photosynthesising, and consume it at night [161, 79]. Starch content used to be determined according to the respective NREL protocol based on the megazyme Total Starch Assay Procedure [127]. However it is no longer available from the NREL website and thus no longer recommended [101]. Either way, starch content will contribute to the glucose value [110, 47]. It means that the cellulose content might be lower if some of the glucose is derived from starch, but it will not influence the overall mass balance closure. Finally, during the HPLC separation, the potential peak for cellobiose was checked, to rule out incomplete hydrolysis. For consistency, the analysis flow was decided to be kept identical for both tissues, as depicted in Figure 4.1.

The detailed formulae and calculation steps are summarised in the appendix section A.2, as well as the HPLC calibration curves, and all the detailed results. Two runs were carried out per sample, and after the initial extractions, the samples were split in two for hydrolysis, meaning that lignin and monosaccharide determination were performed four times per sample. (The monosaccharide composition of the endocarp was only determined thrice as explained in section A.2). The root mean square (RMS) and standard deviation (stdev) of all runs were calculated. An overview of the composition is given in Figure 5.4, with the endocarp on the left and the leaves on the right, and the respective values are listed in Table 5.1.

The mass balance for the endocarp closes almost perfectly to 100.37 wt%, hence we can conclude that the analysis flow was ideally optimised for this type of tissue. The leaves

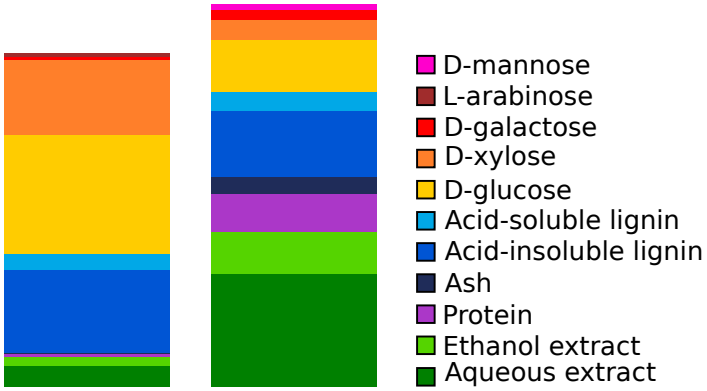


Fig. 5.4 Overview of the biomass composition in wt% of the endocarp (left) and the leaves (right).

Table 5.1 Biomass composition of the endocarp and the leaves. RMS - root mean square, STDEV - standard deviation.

	Endocarp		**Leaves**	
	RMS	STDEV	RMS	STDEV
	wt%	wt%	wt%	wt%
Aqueous extractives	6.25	0.51	34.14	0.75
EtOH extractives	2.55	0.28	12.47	1.68
Protein	0.89	0.68	11.27	0.03
Ash	0.44	0.13	5.26	0.00
Acid insoluble lignin	25.10	3.29	20.17	4.47
Acid soluble lignin	4.87	1.45	5.73	1.68
D-glucose	35.80	0.22	15.49	0.54
D-xylose	22.57	0.45	6.15	1.30
D-galactose	0.89	0.82	2.80	0.28
L-arabinose	n.d.	n.d.	1.66	0.42
D-mannose	1.00	0.48	n.d.	n.d.

sample, however, measures 114.11 wt%. Most likely, the error stems from problems with the lignin determination, and possibly also (ethanol) extraction, since these values show a higher standard deviation between runs. This is not surprising as the lignin content of leaves is known to be particularily difficult to determine, as leaf lignin can be even more heterogeneous, for example with incorporation of more unusual monomers, and there can thus be large fluctuations [63]. Furthermore, interference from monolignols and tannins found in *M. nobilis* leaves is possible [32]. Another potential issue is the protein correction factor of 6.25, which could be too high for the leaves. However since there is very little protein in the endocarp, and the leaves are not the main sample of interest, the detailed determination of the protein conversion factor is omitted. Furthermore, interference from ash during acid hydrolysis could be possible for the leaf sample, but ash content was also determined for extractives and acid insoluble lignin and corrected for. Another possible issue is that the extinction coefficient of 23 L/g cm for acid insoluble lignin is less appropriate for the leaves, but it worked well for the endocarp and was thus kept constant for consistency.

5.3 Cellulose

5.3.1 Isolation of cellulose microfibrils

In this section, the product of 8BT is always called intermediate, and the product of 13BT is called final product, see section 4.2 for abbreviations. 8BT included organic and aqueous extraction, and three alternating repetitions of acidic bleaching and base treatment each, while 13BT further included DMSO extraction, base treatment, DMSO extraction (only for the endocarp), tip-sonication and a final base treatment. NMR and XRD spectroscopy, as well as electron microscopy, were used to follow the purification process.

For the endocarp, after the first few purification steps, at the intermediate 8BT, I assumed to have been able to purify the cellulose microfibrils, since this already comprised an extensive purification protocol. However when imaging with TEM, shown in Figure 5.5 A, I found that most material was still in larger entities, aggregated in bundles. The fibrous structure of the material was visible, but it was aligned in bundles, and with some undefined lumps on the structure. The TEM images of the leaves intermediate, Figure 5.5 B, on the other hand, already showed well-separated microfibrils. Furthermore, even though the endocarp-derived microfibrils were still in bundles, they appeared unusually short - presumably on the scale of 1 μm, which is very unusual. Cellulose microfibrils from e.g. cotton or wood tend to be in the order of at least several micrometers, but as their beginnings and ends are hard to see

with TEM imaging, their length is rarely determined. For the outer periclinal epidermal wall of onion, they are at least several tens of micrometres [162]. Meanwhile, the leaves-derived microfibrils were longer, as usually observed in most plant tissues. It was not possible to statistically assess their length distributions but they seem to be much longer than a few μm.

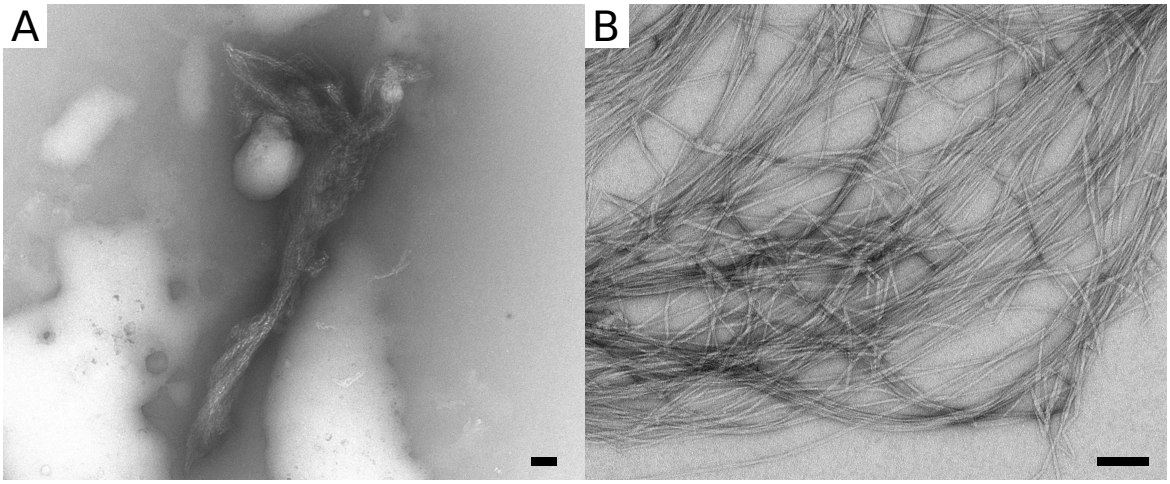


Fig. 5.5 TEM images of the intermediate 8BT: (A) endocarp, (B) leaf. Scale bars are 100 nm.

(The white areas in Figure 5.5 (A) are probably not artifacts from TEM imaging, but some sort of impurity in the sample. For example, they could be leftover traces of lipids, in case that the organic extraction was not complete. Even though the amount of lipids in the endocarp is low, 2.55 wt%, they might still be difficult to remove e.g. if they are very long chains. Similar white spots were also seen in Figure 5.7 A and Figure 5.8 A, and were probably found for the same reasons.)

In the NMR spectra reported in Figure 5.6, it is possible to observe that signals from many compounds in 8BT (light blue line) have already declined compared to the starting material (dark blue line). For both the endocarp (top) and the leaves (bottom), aliphatic carbon signals were observed in the starting material, between approximately 5-45 ppm chemical shift. These signals derive from aliphatic carbon atoms in lipids and waxes, which are found in the cuticle covering the entire outside of the tissue. These compounds should be dissolved in the organic extraction, and indeed these peaks were no longer present in the intermediate 8BT spectrum, indicating successful removal of the cuticle for both the endocarp and the leaves.

Secondly, aromatic carbon signals were present in the starting material for both the endocarp and the leaves. They show up between approximately 110-160 ppm, and derive from polyphenolic compounds comprising lignin and tannin. Another peak at approximately

55 ppm was assigned to the carbon of a methyl group of a phenyl ether, which is also a common motif in lignin. The polyphenolic compounds were oxidised during the acidic bleaching steps, and were thus successfully removed in the intermediate 8BT for both tissues.

Finally, the NMR spectrum of the starting material contains peaks for cellulose and hemicellulose, and these are more difficult to distinguish, since the two compounds are chemically very similar to each other. The sugar ring carbons exhibit peaks between approximately 60-110 ppm for both cellulose and hemicellulose. For hemicellulose, there is an additional peak at around 175 ppm, which stems from the carboxylic acid carbon that can be found e.g. on glucuronic acid decorations on xylan, but is entirely absent from cellulose. Furthermore, hemicellulose yields a peak at 20 ppm, from a methyl carbon of a ketone, which can e.g. be from the acetyl groups on xylan. When comparing the NMR spectra of the starting material and the intermediate 8BT, the carboxylic acid peak at approximately 175 ppm and the methyl from acetyl peak at 20 pmm have definitely disappeared, indicating removal of hemicellulose. However, for the endocarp, in the overlapping area of 60-110 ppm, some of the cellulose peaks are not as distinct and sharp as they should be for purified cellulose. Hence, there are still small amounts of hemicellulose present in the endocarp 8BT sample. (The reason why the carboxylic acid carbon peak and the methyl carbon from an acetyl group peak were entirely gone but there is still some hemicellulose present in the sample is that hemicellulose may only contain small amounts of carboxylic acid groups, depending on the plant and tissue, and the acetyl groups were removed by the base treatments. Furthermore, NMR is not sensitive enough to detect traces of compounds, generally about one percent of hemicellulose could easily go undetected.)

Therefore, the amorphous-looking material observed in the endocarp 8BT TEM images was identified as xylan, which seems to interact very tightly with cellulose and seems difficult to remove.

Since the fibre length observed for the endocarp was unusually short, I repeated all the isolation procedure on a second batch. I observed identical NMR and XRD results, so only the NMR and XRD results of the second batch are displayed in this thesis for better readability. Figure 5.7 shows a TEM image of the batch2 material at the same intermediate stage 8BT, and the same morphology was found: bundles of fibres that are probably held together by xylan still remaining in the sample. Identical to the first batch, the fibrous structure of the bundles is visible, and they seem to be shorter than 1 μm.

Additionally to these bundles of fibres, we also found another population of cellulose microfibrils in the endocarp sample in both batches. They are depicted in Figure 5.8, and are

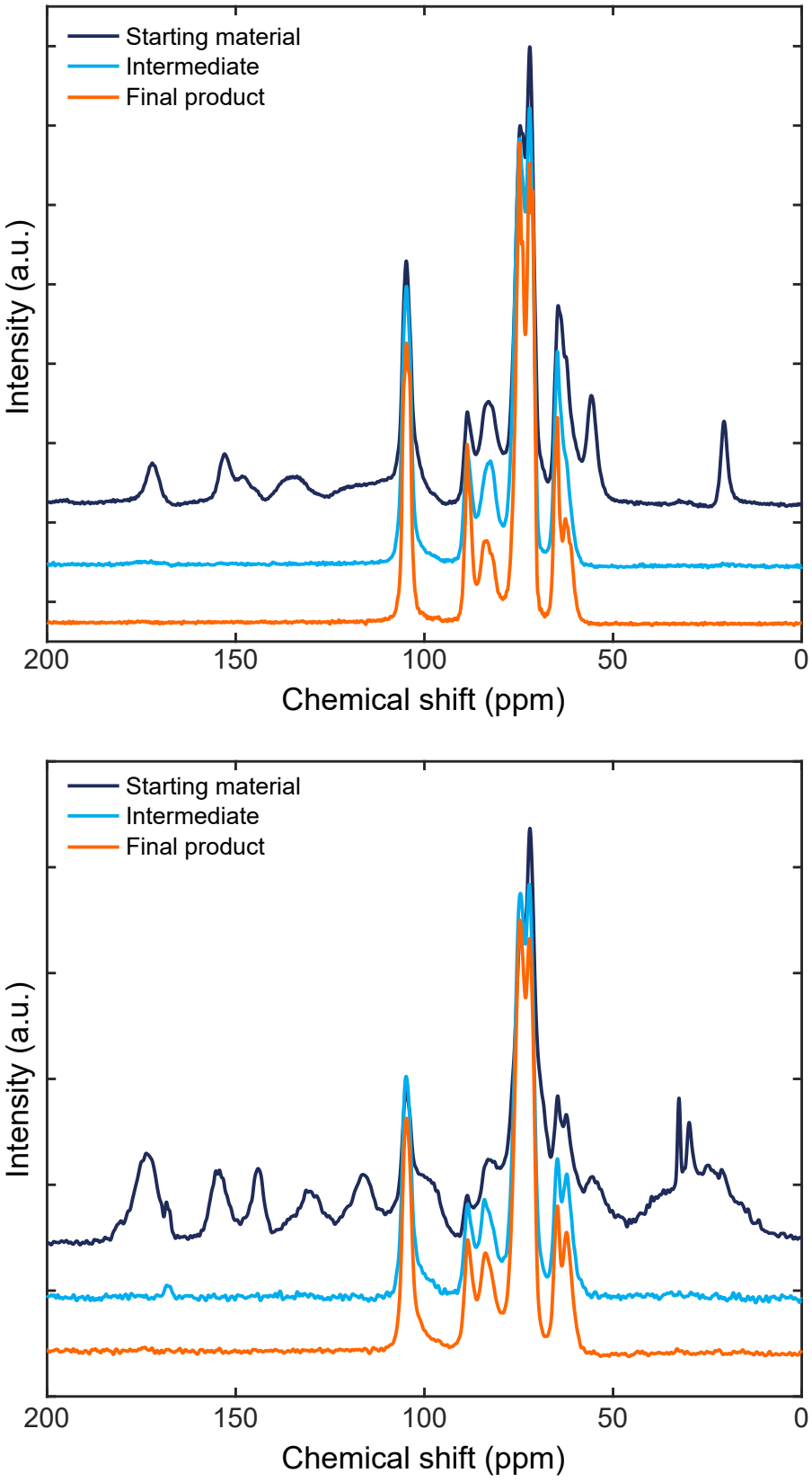


Fig. 5.6 NMR spectra of the starting material, intermediate 8BT and final product 13BT. Top: endocarp. Bottom: leaves.

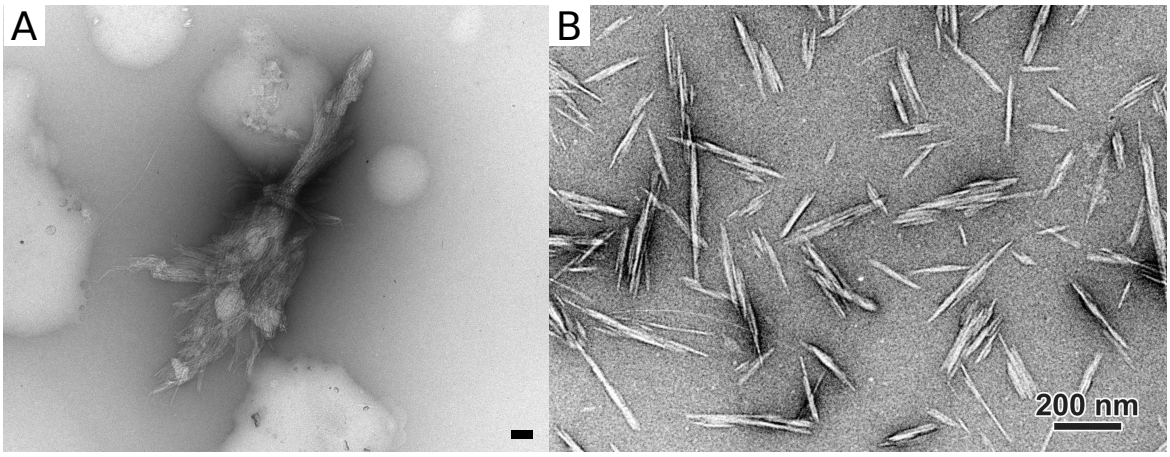


Fig. 5.7 (A) TEM image of the intermediate 8BT of the second batch of endocarp isolation. Scale bar is 100 nm. (B) TEM image of cotton CNC, from Dr. Yu Ogawa. Scale bar is 200 nm.

much shorter, around 100 - 200 nm, and have a bigger thickness, so overall a much smaller aspect ratio.

As it was very surprising to find this population of very short microfibrils, I made sure I did not introduce hydrolysis or impurities during the purification series. First, I checked that none of the purification reactions cleaved the microfibrils, and after thorough checking of all the procedures and also not finding any short microfibrils in the leaves samples, I was confident to exclude that possibility. Secondly, I considered possible cross-contamination from products that other lab members work with. One possible source of contamination present in the lab area is chitin nanocrystals - but they would be detected during NMR spectroscopy. The second population of microfibrils is present in the sample in considerable amounts, and if it were chitin nanocrystals, the NMR spectrum would display the peaks for the acetyl amine group. The last option would be cross-contamination from cellulose nanocrystals (CNC). The only type of CNC in use in the lab are CNC derived from cotton. While they are chemically indistinguishable, they have a different morphology and their aspect ratio is very different from the ones found in the endocarp sample (shown in Figure 5.7 B). The average width of CNC prepared from cotton at different hydrolysis temperatures was found to be around 16 nm, thicker than the endocarp second population [35]. Finally, the entire endocarp purification procedure was repeated, with possible cross-contamination in mind and thoroughly cleaning all glassware with concentrated H_2SO_4 before use, and the population of short microfibrils was still found in TEM imaging in this second batch. Thus this second population of short cellulose microfibrils has to be intrinsic to the endocarp tissue.

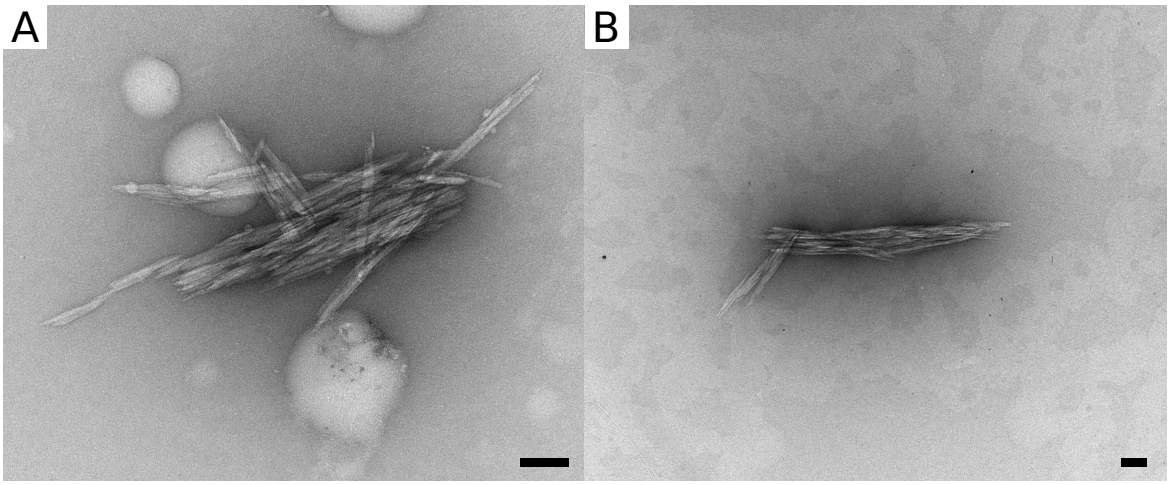


Fig. 5.8 TEM images of the intermediate 8BT, of the population of short cellulose microfibrils. (A) First batch of endocarp isolation. (B) Second batch of endocarp isolation. Scale bars are 100 nm.

It is hard to quantitatively assess in what ratio these two populations exist in the sample. They are only visible in TEM imaging when zoomed in significantly, requiring a very large number of images to survey a statistically relevant area. Another uncertainty is that the two different populations might not deposit in the same way on the TEM grid during preparation, e.g. one might aggregate more than the other, and possibly precipitate into too large chunks.

Since for the endocarp sample, the majority of the first population of cellulose microfibrils, of intermediate length and small width, was still observed to be aggregated in bundles, and since there was still a small amount of xylan observed in the NMR spectrum, I decided to carry out further purification treatments. I replicated the same reactions for the leaves as well, for consistency between samples.

In order to remove the residual xylan without altering the native cellulose structure, I decided to carry out DMSO extractions. Another possible option was to increase the base concentration to make the base treatments more efficient, but cellulose mercerisation might occur at higher base concentrations. Depending on the type of cellulose, its onset was observed at base concentrations as low as 9 wt% for primary cell wall cellulose isolated from sugar beet pulp [30] and 8 wt% for secondary cell wall prickly pear seed pericarp cellulose [51]. To be able to exclude this potential issue, I decided to not increase the base concentration higher than 4 wt%, and check the material with XRD.

Figure 5.9 shows TEM images of the second population of cellulose microfibrils after one more purification step, the first DMSO extraction 9DE. This second population of microfibrils, short and with thick diameter, was observed as isolated particles in (A), and

in the same frame as the first population of intermediate, thin microfibrils, in (B). It is obvious that the two different types seem to be distinct populations. Even though some of the intermediate microfibrils are still in bundles, individual thin fibrils are now visible, while the second population looks clearly thicker. It does not seem possible that one of the populations somehow transitions into the other during the treatments, since they would have to either become thicker or longer, which is impossible.

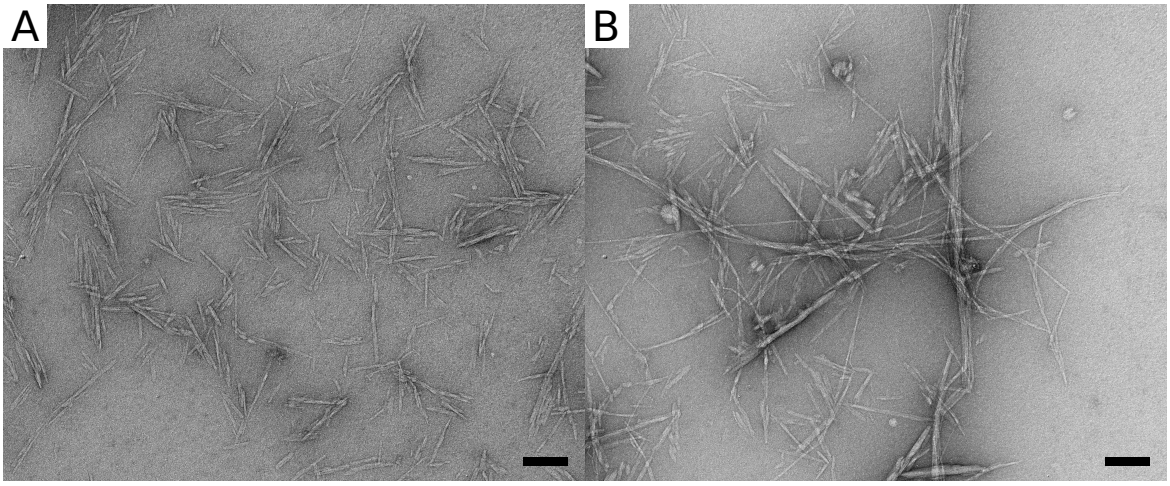


Fig. 5.9 TEM images of the endocarp cellulose microfibrils after the first DMSO extraction 9DE. (A) Second population of short, thick microfibrils. (B) Both populations in the same image. Scale bars are 100 nm.

Furthermore, after observations via electron microscopy revealed that the material was still slightly aggregated, I decided to tip-sonicate the sample, to help break up the aggregates and make the surface available for further base treatment. The NMR spectra of the final product are presented in Figure 5.6, orange lines. The endocarp product 13BT contained a small amount of regenerated xylan, which was subtracted from the spectrum for easier interpretation, the original data are shown in the appendix section A.3.

Electron micrographs of the final product 13BT of the endocarp first batch (A) and the leaves (B) are depicted in Figure 5.10, and Figure 5.11 (A) for the endocarp second batch. The first population of microfibrils observed for the endocarp is still of intermediate length, less than 1 or 2 μm and a thin diameter. They were now more separated than the microfibril bundles observed at the intermediate stage 8BT, indicating the effectiveness of the further purification treatments. The microfibrils derived from the leaf tissue are still of very long length, in the order of several μm, and it is impossible to see their beginnings and ends.

There is essentially no difference between the leaves at the intermediate and final stage, 8BT and 13BT respectively. The cellulose microfibrils were already well isolated at the

intermediate stage, and the further purification reactions did not cause any damage. The images in Figure 5.5 B and Figure 5.10 B only look different because they were taken at different magnifications, and 13BT is of a much larger observation area, to emphasise the long length.

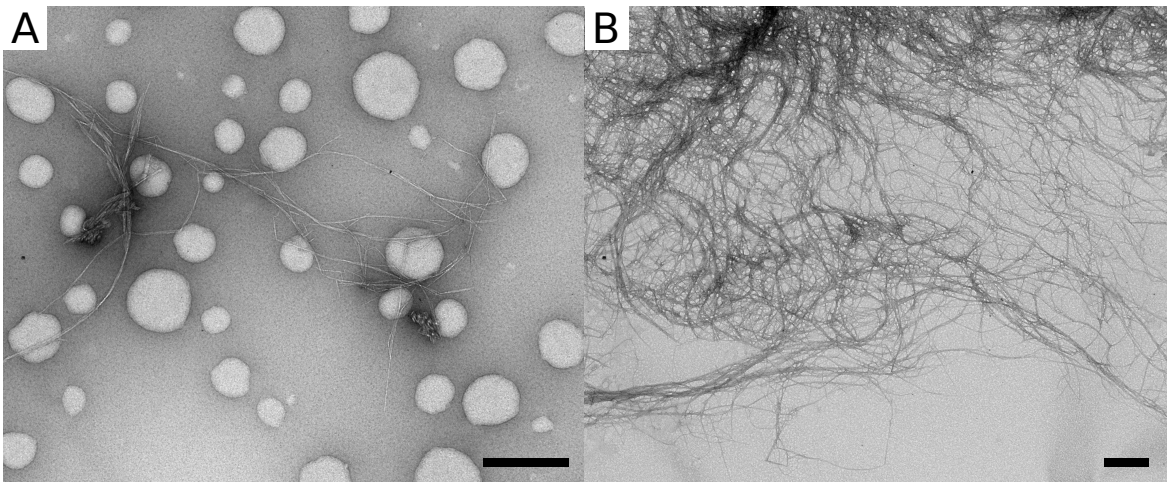


Fig. 5.10 TEM images of the final product 13BT: (A) endocarp, (B) leaf. Scale bars are 500 nm.

(Again, white impurities were observed in Figure 5.10 A, as well as Figure 5.11 B. They could be leftover lipids, or another option at this point could be residual DMSO that was not fully washed out after the two DMSO extractions, 9DE and 11DE.)

Figure 5.11 B shows another TEM image of the endocarp first batch final product 13BT, and again also containing the second population of cellulose microfibrils: they are short with only a few hundred nm length, and their thickness is bigger than those of the first population. In the second batch of the endocarp purification, we were not able to image the second population. It might be that during the centrifugation washes after every reaction, this population was washed out more due to their small size - they might not settle as well as the first population and be removed with the supernatant. Another possibility is that they aggregate more and are thus too big to be imaged in transmission electron microscopy. (Aggregation generally becomes a problem the more purified the microfibrils are, since amorphous xylan can also act as a coating introducing some steric hindrance for the microfibrils to align.) This explanation seems more likely for two reasons: firstly since they are so short and straight, they might align very well, enabling aggregation. Secondly, according to X-ray diffraction, which will be discussed more in the following section, the crystallite size (of the 200 crystal plane [98]) was found to be 4.02 nm. It is hard to tell from the TEM images (because of the negative staining and the small size), but it seems that if the sample only contained the first

population, the crystallite size would have to be smaller than 4.02 nm, since they look very thin. Since the sample used for XRD spectroscopy is freeze-dried, it should contain a good representation of the average of the entire sample, as different behaviour upon drying will not play a role. The hypothesis is hence that the crystallite size is 4.02 nm because of the contribution from the second population of microfibrils.

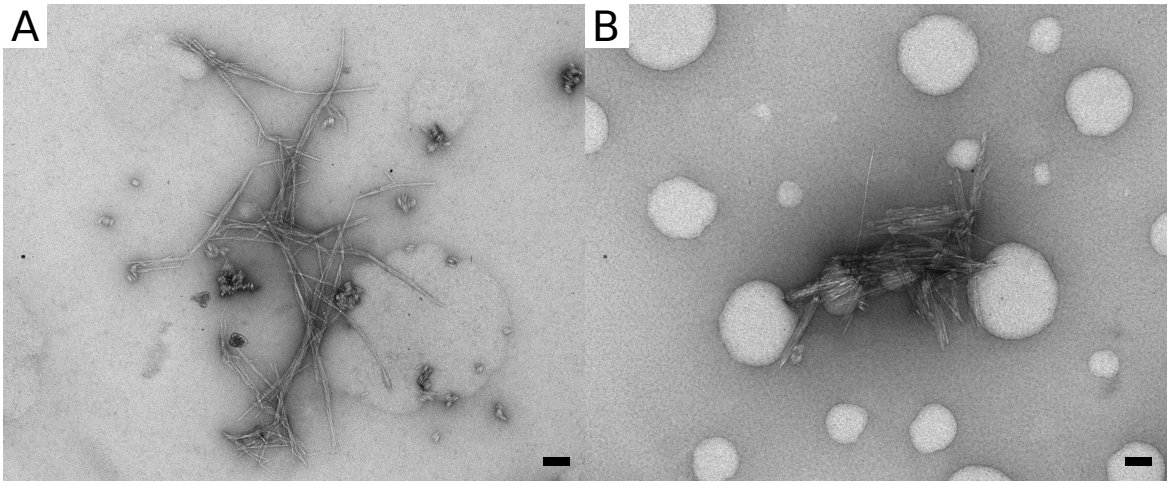


Fig. 5.11 TEM images of the final product 13BT: (A) Endocarp second batch, intermediate thin microfibrils. (B) Endocarp first batch, second population of short microfibrils with larger diameter. Scale bars are 100 nm.

The observation of these two populations of microfibrils was surprising. Both populations of cellulose microfibrils are shorter than 1 or 2 μm. Furthermore, they both seem to coexist in the endocarp tissue, and there is no way how they could transform into one another. At this point, it is not possible to infer where these two different populations are found in the endocarp tissue. One idea could be that the first population, the intermediate thin microfibrils, are deposited in the primary cell wall, while the second population, the shorter, thicker microfibrils, could make up the secondary cell wall and thus the helicoidal architecture. If this scenario was the case, the second population should be much more abundant, but possible different behaviours during the isolation and TEM sample preparation, like aggregation and settling, cannot be excluded. Furthermore, many more TEM images would have to be taken to conduct a statistically relevant survey of the ratios of the first to second population. According to literature, it would be possible to produce different populations of cellulose microfibrils in the different walls though, since different CESA make up the CSC in the PCW and SCW, as explained in subsection 1.3.2 [84].

Interestingly, preliminary experiments on the epicarp tissue of the monocot *Pollia condensata* seem to indicate very similar fibre populations [147]. Only a small amount of microfibril isolation steps were carried out: organic and aqueous extraction 1OE and 2AE, and one set of acidic bleaching and base treatment, 3AB and 4BT. It was only carried out on a small amount of fragments in an Eppendorf vial since only a small amount of material was available at the time. It was expected that most of the material stayed in large chunks, and imaging was carried out on the edges of chunks, where a bit of tissue had started to disintegrate.

Figure 5.12 depicts the different types of cellulose microfibrils found in the *P. condensata* epicarp: (A) bundles of fibrils, (B) microfibrils very similar to the first population of *M. nobilis* endocarp, with intermediate length and thin width, and (C) microfibrils very similar to the second population, with short length and bigger thickness. Overall, all the same types of fibrils were found as in the *M. nobilis* endocarp, and they all seem to have comparable morphology. It might be an indication that the helicoidal architecture of cellulose microfibrils in the secondary cell wall in different species relies on the same types of building blocks.

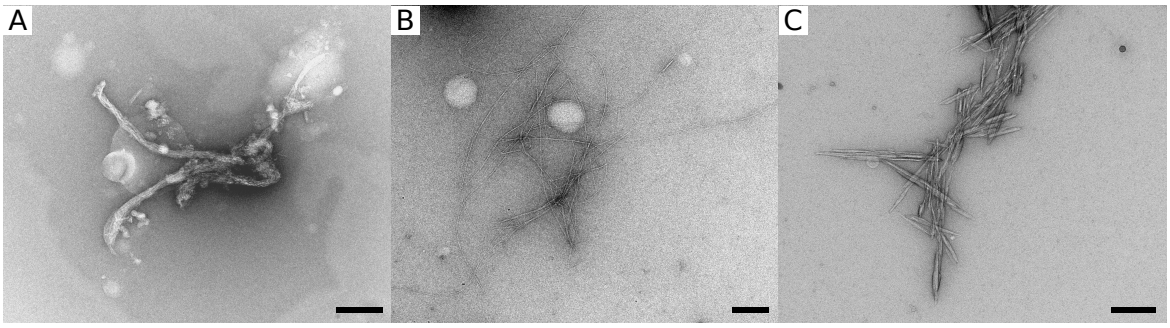


Fig. 5.12 TEM images of fibrils isolated from *Pollia condensata*: (A) bundles of fibrils, (B) microfibrils with intermediate length and thin width, and (C) microfibrils with short length and bigger thickness. Scale bars are 200 nm.

5.3.2 Crystallinity and crystallite size of cellulose microfibrils

XRD analyses

I monitored the crystalline structure of the plant materials throughout the isolation procedure, and investigated the crystallinity of the final products 13BT. Figure 5.13 shows the XRD spectra of the starting material, intermediate 8BT and final product 13BT of the endocarp on the left, and the leaves on the right. The observed peaks correspond to the $1\bar{1}0$, 110, 102, 200 and 004 crystal planes [97]. They are generally more distinct and defined for the endocarp sample, indicating more order. Furthermore, peaks became more distinct and defined the more purification treatments were carried out, showing that more amorphous

material was removed with each step. The peak positions correspond to the crystal structure of the cellulose Iβ allomorph as the major crystal structure, as expected from higher plants ($(1\bar{1}0)$ at 14.88° 2θ, (110) at 16.68° 2θ, and (200) at 23° 2θ) [98, 40]. Finally, it is obvious that the crystal structure is maintained throughout the purification steps, no conversion to other allomorphs is observed.

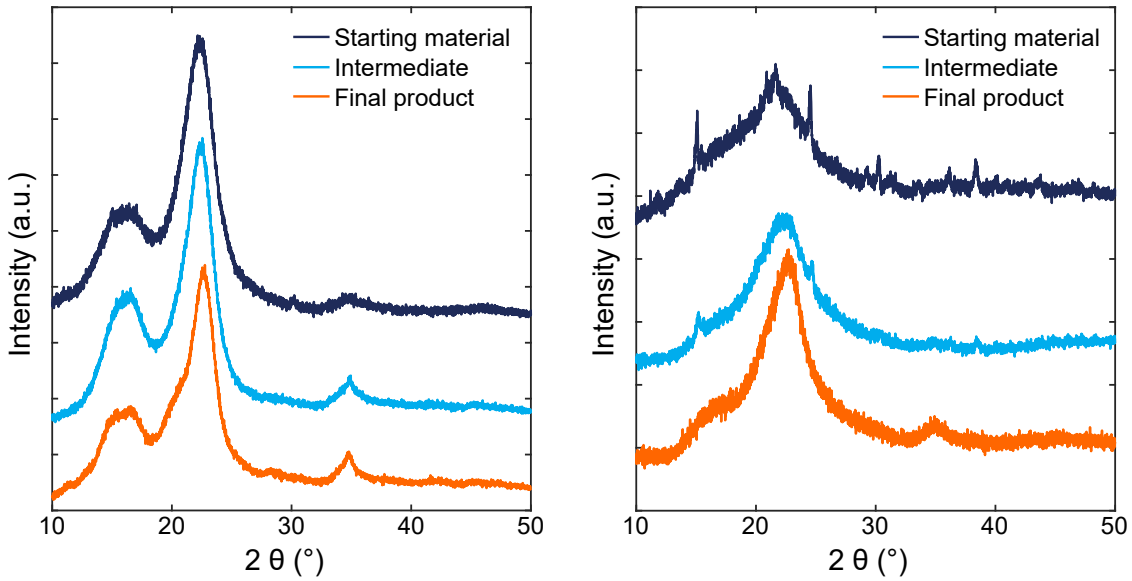


Fig. 5.13 XRD spectra of the starting material, intermediate 8BT and final product 13BT of the endocarp (left), and the leaves (right).

Next, I deconvoluted the obtained XRD spectra to approximate Gaussian fits to each of the diffraction peaks, depicted in Figure 5.14, again with the endocarp on the left and the leaves on the right. The dark blue line shows the original measurement data, the light blue lines are Gaussian fits of the crystalline peaks, the dark red line is a Gaussian fit of the amorphous contribution, and the middle blue line is a correction for the X-ray lamp degradation, while the orange line is the sum of all the fits. Unfortunately, the X-ray source used within the duration of the project was degrading and introduced some effects that could not be quantified or corrected for. This issue was worse for the endocarp sample, the leaf sample was measured 5 months prior. Some effect was already seen in the leaf spectrum, but it was stronger in the endocarp sample. For this reason, in order to still obtain a fit, I had to introduce this additional curve, the middle blue line, to the endocarp spectrum, to take the lamp degradation into account. I also used this additional curve for the leaf sample, for consistency between the two.

These Gaussian fits were used to calculate a crystallinity index (*CI*) and the Scherrer crystallite size. It is generally tricky to calculate the *CI*, since it is still not even well

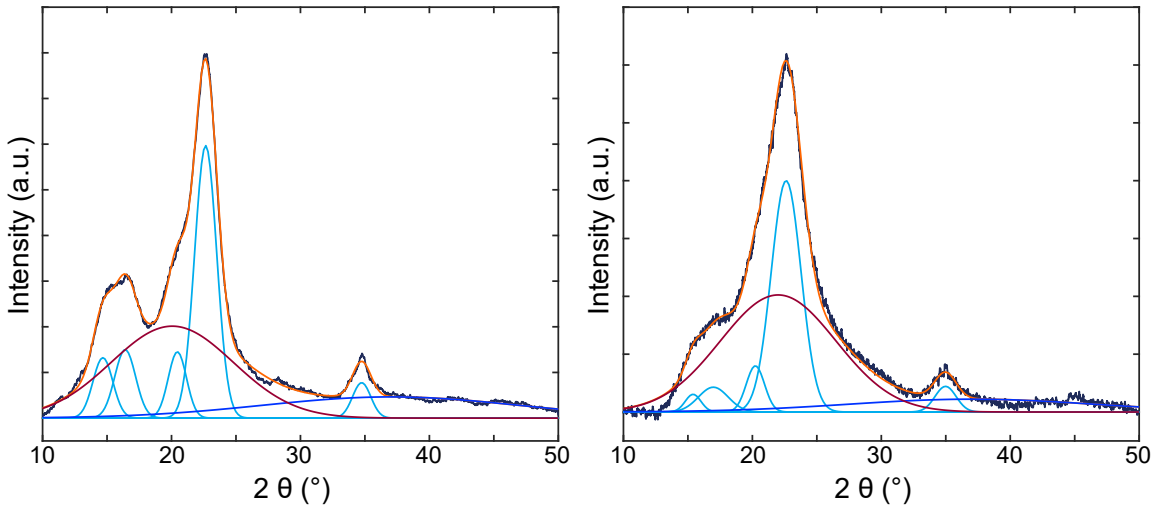


Fig. 5.14 Fitting of the 13BT XRD spectra of the endocarp (left) and leaves (right) with Gaussian distributions. Dark blue line: measurement data, light blue lines: Gaussian fits of the crystalline peaks, dark red line: Gaussian fit of the amorphous contribution, middle blue line: correction for the X-ray lamp degradation, orange line: sum of all the fits.

understood and defined what is meant by crystallinity (see subsection 1.3.1 for a short overview). Furthermore, many different methods are used for calculating the *CI*, mainly based on XRD or NMR measurements, comprising the so-called XRD-based peak height method, peak deconvolution method, and amorphous subtraction method, and the NMR-based peak separation method, and peak deconvolution methods. Significant deviations are observed when calculating the *CI* for the same sample with different methods [105, 39, 75, 155, 74]. Due to these large variations, comparing obtained values to the literature has to be done with caution. Therefore, I decided to use the XRD-based peak deconvolution method and the NMR-based peak deconvolution method here, comparing the endocarp and leaf final products to each other, keeping in mind that the *CI* (as opposed to the Scherrer crystallite size) should not be considered an absolute value.

Hence, for the XRD data, I calculated a crystallinity index *CI* by dividing the area of all the crystalline peaks (light blue lines) by the total area: the results are summarised in Table 5.3. (I defined all the crystalline and amorphous peaks as total area, and left the lamp correction middle blue line out, since it is intrinsic to the measurement and not the sample.) I obtained a *CI* of 48.75 % for the endocarp tissue, and 40.79 % for the leaves. It should be kept in mind that both values are probably underestimated, with the endocarp even more so than the leaf. I tried to compare these values to values found in the literature calculated with the same method, and it seems that the crystallinity is average or above average for a purified secondary cell wall, for example 33 % for bleached Kraft eucalyptus wood pulp [39].

$$CI_{XRD} = \frac{A_{crystalline}}{A_{total}} \cdot 100 \tag{5.1}$$

Next, the Scherrer crystallite size perpendicular to the *hkl* diffracting planes was calculated according to:

$$D_{hkl} = \frac{K \cdot \lambda}{\beta_{1/2} \cdot \cos\theta} \tag{5.2}$$

with K as a correction factor, usually 0.9 for cellulose, with λ as the X-ray wavelength (1.542512 Å), θ is the diffraction angle and $\beta_{1/2}$ is the peak width at half maximum intensity, both derived from the Gaussian fits [39]. It is based on the fact that very small crystallite sizes lead to line broadening of the diffraction peaks. However, other effects, like lattice strains, also cause line broadening, and the Scherrer crystallite size should therefore be considered to be the minimum size [69].

Table 5.2 Scherrer crystallite size of the different crystal planes for the final product 13BT of endocarp and leaves.

		hkl	$1\bar{1}0$	110	102	200	004
Endocarp	Crystallite size	nm	4.42	4.22	4.85	4.02	5.01
Leaf	Crystallite size	nm	5.60	2.97	4.85	2.98	4.33

These values for the endocarp are bigger than average lignified secondary cell wall in woody tissues (e.g. approximately 3 nm or just below for the 200 peak in many wood species, unpublished data from Kuribayashi et al., in preparation, and 3.2 $\pm$ 0.1 nm for Norway spruce (*Picea abies*) [3], and 3.29 $\pm$ 0.13 nm for Sitka spruce (*Picea sitchensis*) [38], and 3.84 $\pm$ 0.13 nm for Tonkin cane bamboo (*Pseudosasa amabilis*) [137]), with the endocarp generally having bigger crystallite sizes and thus more order than the leaves. The values for some of the leaves peaks should be considered with caution, especially $1\bar{1}0$, 110, and 102, since these peaks are not very defined in the spectra, and are mostly fitted for consistency with the endocarp data processing.

NMR analyses

Additionally to the XRD crystallinity analyses, I deconvoluted the C-4 peak of the final product 13BT spectra of both the endocarp and the leaf, to estimate the crystallinity index *CI* based on NMR measurements [75, 155, 74]. Furthermore, it is possible to deconvolute the C-1 and the C-6 peaks in the same way, but since there is an overlap with the free xylan

spectrum in these areas, which is difficult to precisely quantify, it was omitted (see the explanation in the appendix section A.3). Fitting was carried out with a combination of Gaussian (for amorphous contributions, continuous lines) and Lorentzian (for crystalline contributions, dashed lines) functions. At around 79.5 ppm, the C-4 peak started to overlap with the large peaks for the C-2, C-3 and C-5 carbons. This area could also be fitted, and then subtracted from the C-4 area so that the overall fitted curve of the spectrum would reach the baseline, but this correction is expected to be minimal and is therefore omitted, cutting off the C-4 peak at 79.5 ppm instead. The spectra and fitting curves are depicted in Figure 5.15, with the golden line being the original data, and the black line as the sum of all the fits.

A *CI* based on the NMR spectrum was calculated, equivalent to Equation 5.1, according to:

$$CI_{NMR} = \frac{A_{crystalline}}{A_{total}} \cdot 100 \tag{5.3}$$

The crystalline area comprised the peaks for I α, I $(\alpha+\beta)$, para-crystalline, and I β contributions, and the total area consisted of the crystalline peaks plus accessible and inaccessible peaks, but not the xylan peak at 82 ppm. Values of 47.98 % for the endocarp and 35.77 % for the leaves were obtained. This approach to calculate the *CI* is not very widespread in the literature (access to ssNMR facilities might be restricted and the curve fitting is more sophisticated), so cannot be compared to any *CI* from different cellulose microfibrils. However it does show the same trend as the CI_{XRD} values, the endocarp cellulose microfibrils being significantly more crystalline than the ones isolated from the leaves. Furthermore, the obtained order of magnitude is very comparable, and good agreement was found between the two techniques.

The *CI* for both the XRD fitting and the NMR fitting for the endocarp and the leaves 13BT are summarised in Table 5.3.

Table 5.3 Crystallinity index *CI* from XRD and NMR spectral deconvolution for the endocarp and leaf.

	Endocarp	**Leaf**
	CI	*CI*
	%	%
XRD peak deconvolution	48.75	40.79
NMR C-4 peak deconvolution	47.98	35.77

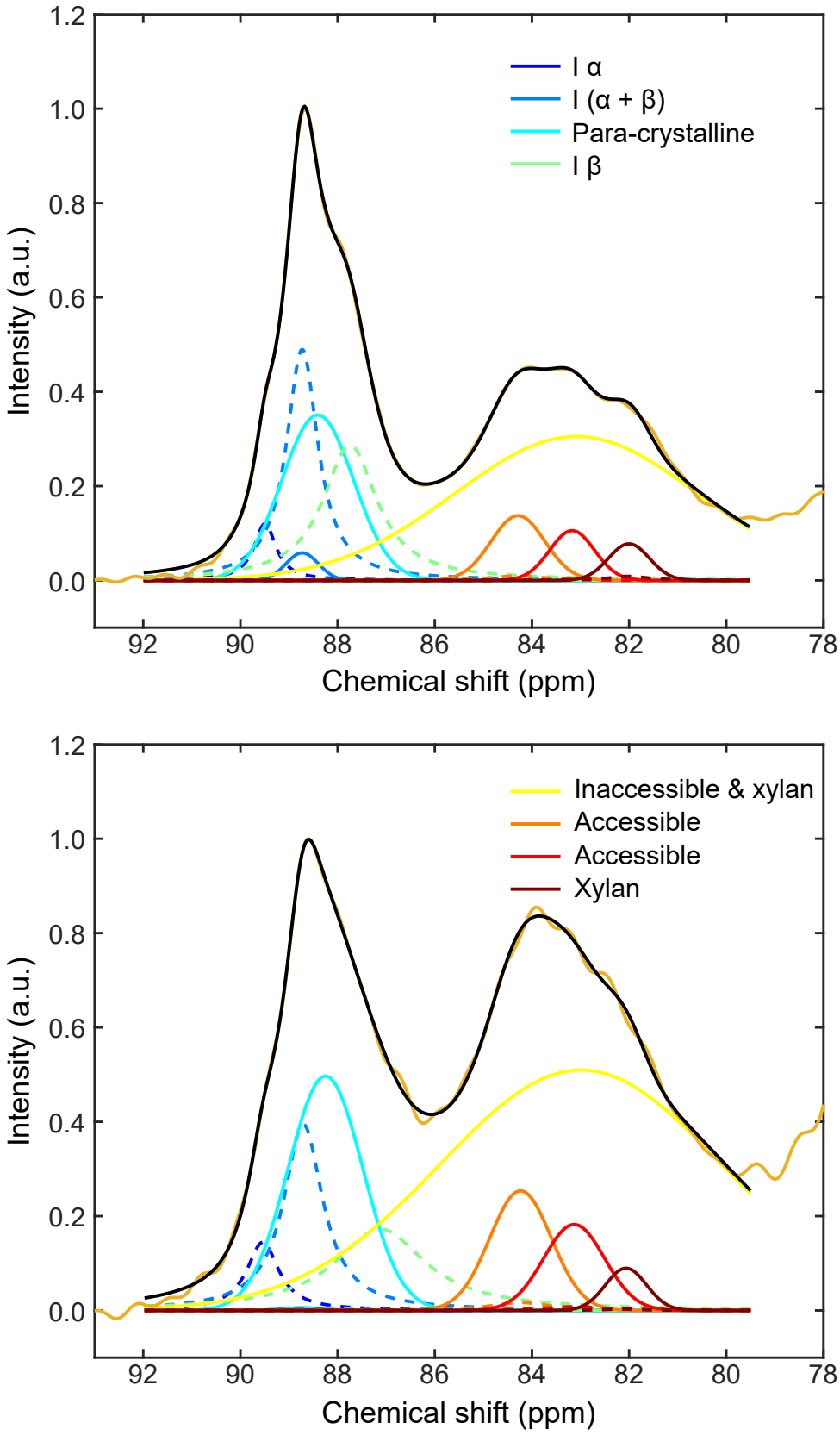


Fig. 5.15 Fitting of 13BT NMR spectra with Lorentzian (crystalline contributions, dashed lines) and Gaussian (amorphous, continuous lines) peaks, endocarp on top, leaves on bottom. The legend is spread over both plots, golden line: original data, black line: sum of all fits.

5.3.3 Isolation of cellulose microfibrils - discussion

It was surprising to find that the cellulose microfibrils from the *M. nobilis* endocarp are extremely short in length, and that their crystallinity is also above average, when comparing to e.g. cellulose microfibrils from eucalyptus wood pulp [39].

I performed a thorough literature search to decide on the mildest purification steps with the least side effects possible, to rule out any accidental degradation of the cellulose microfibrils, such as e.g. acid hydrolysis etc. during the purification process. Furthermore, I repeated the exact same procedure on a second batch of endocarp, obtaining the same results, and on the leaves of *M. nobilis*, (minus the last DMSO extraction), and found that their microfibrils length is in the range of micrometers, as expected, and their crystallinity lower than the endocarp one. In the following section, the reasoning behind the steps selected for the isolation procedure are discussed.

Firstly, the organic and aqueous extraction only remove compounds soluble in the respective medium, like oils and waxes from the cuticle, or protein and ash, and should have no effect on the cellulose microfibrils. Then, the acidic bleaching step was supposed to break down the lignin network by oxidising the phenol ring [44]. Acetic acid was added to produce the active compound, chlorine dioxide $ClO2$ from sodium chlorite $NaClO_2$ [94], while it is not a strong enough acid to induce acid hydrolysis of polysaccharides. Furthermore, only a few drops of acetic acid were added, to bring the reaction mixture to acidic pH, and therefore it is not expected to cause hydrolysis.

Next, heteropolysaccharides were dissolved during the base treatment. Higher concentrations of base can lead to the transformation of cellulose I into cellulose II via a reaction called mercerisation, but the lowest reported base concentration necessary reported in the literature was still much higher than the 2 or 4 wt% used here (8 wt% for secondary cell wall prickly pear seed pericarp cellulose [51]).

After repeating the acidic bleaching and base treatment alternatingly thrice, I expected to have removed most of the residual compounds, since this seems to be an elaborate treatment sufficient enough for other types of biomass [51]. However in the NMR spectra, the C-4 and C-1 peaks still did not look very crystalline. It looked like overlap from other polysaccharides, there was still a small bump at around 82 ppm typical for xylan interacting in a two-fold helical screw with cellulose [122], and the shape of the C-2, C-3 and C-5 region still resembled that of the starting material. Furthermore, in TEM images of the material, I observed large bundles of fibres, which seemed to be held together by some amorphous material, which according to NMR would most likely be a hemicellulose. Since xylose is the most abundant monosaccharide other than glucose based on the monosaccharide analysis, this amorphous material was most likely xylan. On the other hand, no aromatic compounds

seemed to be present anymore according to NMR, so I decided to continue with purification steps targeted at removing hemicellulose and at breaking up aggregates.

The DMSO extraction was supposed to swell the aggregates and dissolve xylan, and the added LiCl was supposed to interrupt the hydrogen bonding network to make surfaces more accessible [36]. The solid plant material was freeze-dried for this treatment, since it only works in completely anhydrous conditions. Freeze-drying is the mildest option available, where hornification of cellulose microfibrils is kept to a minimum (there was only less aggregation observed during supercritical CO_2 drying, which was not available [17].) Generally, the entire material was never dried except for the DMSO extractions to avoid hornification [93].

Finally, tip sonication was performed to help break up aggregates as well, and was carried out just before the final base treatment, to have the cellulose-xylan surfaces as accessible as possible. An energy input of 40 kJ/g cellulose was chosen based on preliminary calculations and testing different tip sonication times and checking for kinks in the cellulose microfibrils with TEM and STEM-SEM to stay below the point where fibres would get damaged [133].

Concerning the crystallinity indices determined for the isolated fibres, harsh mechanical treatment of the starting material was completely avoided. It is known that ball-milling can reduce the crystallinity, and extensive milling can even lead to obtaining entirely amorphous material [6]. For this reason, the endocarp tissue was only ground by hand in a mortar and pestle - it is curved and very brittle and thus very easy to break - until it could be passed through a sieve with approximately 1 mm pore size. The fresh leaves were simply cut into small pieces with scissors and used directly, without drying. For the NREL biomass composition analysis, they were dried and ground in a planetary ball mill at minimum time and speed necessary and with cooling periods inbetween.

All in all, any treatment potentially altering the native cellulose microfibrils was avoided.

5.3.4 Challenges in determining cellulose microfibrils dimensions

Observing so small cellulose microfibrils is very challenging, and there are limitations for every technique. A main challenge is sample aggregation: this can happen both in water, simply for the isolated cellulose microfibrils, and it could become worse when preparing the sample for microscopy. Aggregation is a particular problem for cellulose microfibrils, since they are not charged (it is easier with cellulose nanocrystals (CNC), whose surface is typically covered in negative charges derived from acid hydrolysis). To limit aggregation

of cellulose microfibrils in water as much as possible, low concentration of sample and tip-sonication treatments were used.

Subsequently, the CMF can aggregate during microscopy sample preparation. For TEM samples, very dilute samples were used (approximately 0.01 wt% concentration), and the carbon grids were always glow-discharged before depositing the sample, which makes them hydrophilic so that the entire aqueous drop can spread over the whole grid, and gives negative charges, so the CMF will adsorb on the surface, reducing their mobility upon drying. Nevertheless, aggregation on the grid cannot be fully ruled out. Cryo-TEM, where the sample is embedded in a thin electron-transparent film of vitreous ice during the observation, could circumvent these issues, but it was not possible to access a cryo-TEM, and operation is more challenging [103].

Generally, the width and possible aggregation of CMF along their length can be complicated to be evaluated from TEM observations. Since CMF only show low electron scattering contrast because they consist of light atoms, they have to be negatively stained, which could give wrong impressions of the thickness if not done properly. Furthermore, by convention, CMF images are slightly underfocused, since they are almost invisible when exactly focused. If the underfocus is too much, image interpretation might also become more difficult [103]. All the TEM images in this thesis were recorded by our collaborator Dr. Yu Ogawa, who is very experienced with electron microscopy of cellulosic samples, so the above-mentioned complications should be minimised. However, it can still be difficult to spot length-to-length aggregation. And finally, no height information is available from TEM observations.

Atomic force microscopy (AFM) could be a useful supplementation to the TEM images, but is very time-consuming beyond the scope of this thesis, and comes with its own limitations. If the AFM is well-calibrated and -operated, the height information (topography) is very precise and reliable. Moreover, it should be possible to determine whether particles are aggregated when extracting height profiles. However, AFM sample preparation on a mica surface can have similar issues as TEM sample preparation on carbon grids, and it is not possible to determine the width of CMF by AFM, since it is in the same order of magnitude as the curve radius of the tip (e.g. the tip radius of the OTESPA-R3 tip from Bruker is 7 nm [15]). This is also true for the length of the CMF, but the error should be negligible compared to the longer length [72]. To sum up, AFM measurements can be useful to accurately determine the height and possible aggregation of CMF, but is unfortunately beyond the scope of this thesis.

5.3.5 Cellulose microfibrils versus fibre bundles - discussion

The cellulose microfibrils composing the helicoidal architecture in the *M. nobilis* endocarp were isolated. One interesting question is to understand how they assemble in this helicoidal

arrangement, and how and at what point the left-handed twist is introduced. It could be at the molecular scale - D-glucose is a chiral molecule. However, a right-handed twist of the helicoidal architecture was observed in the pericarp secondary cell wall of *P. condensata*, giving raise to RCP reflection, suggesting that the intrinsic chirality of D-glucose alone is unlikely to be the final answer [147].

The twist could also happen a hierarchical level further up, at the length scale of microfibrils. Cellulose microfibrils were imaged displaying a twist, but again, it was always in the same rotation direction, making the observation of the right-handed helicoidal architecture in *P. condensata* puzzling [162].

The left-handed twist could also happen at yet another hierarchical level further up. The microfibrils interact very tightly with xylan present in the secondary cell wall, and form fibre bundles, as introduced in subsection 1.3.1. It could be a combination of factors, like the precise morphology and crystalline properties of the cellulose, and the primary structure of xylan, leading to the formation of the helicoidal architecture. If this is indeed the case, the measured dimensions of building blocks would have to be adapted to correspond to the fibre bundles.

It is very difficult to estimate the precise morphology of the bundles, since it is very hard to isolate them. Chemical purifications would not be of much use, since there is no way to stop the disintegration at the bundle level but not smaller. Possibly only carrying out bleaching to break down the lignin network, but leaving the hemicellulose matrix intact could be an option worth investigating. Another option might be mechanical treatment, like homogenisers, but it would probably take many test runs to optimise conditions, and only a small amount of endocarp material is available.

An approximate method to estimate the fibre bundles size without the need for large amounts of material was measuring the width from SEM images. The endocarp was subjected to organic extraction, to remove the cuticle covering the structure, and then areas on the sample were monitored where the top view of the helicoidal architecture was visible. The observed values will be overestimated since the sample was sputter-coated with 5 nm of Pt. Atomic force microscopy (AFM) imaging would avoid this error, but is much more time-consuming. Figure 5.16 shows SEM images that were used to estimate the fibre bundles width. Lines were drawn over bundles seen in brighter shade, that have come off the surface slightly, and the greyscale function in ImageJ was used to determine their width. In total, ten such bundles were measured in the three zoom images in B-D each, and their average width was obtained to be 59.6 ± 15.3 nm.

If the twist is indeed introduced at the fibre bundles length scale, a different morphology would have to be used to build a cell wall model and to be used e.g. in the SAXS data processing.

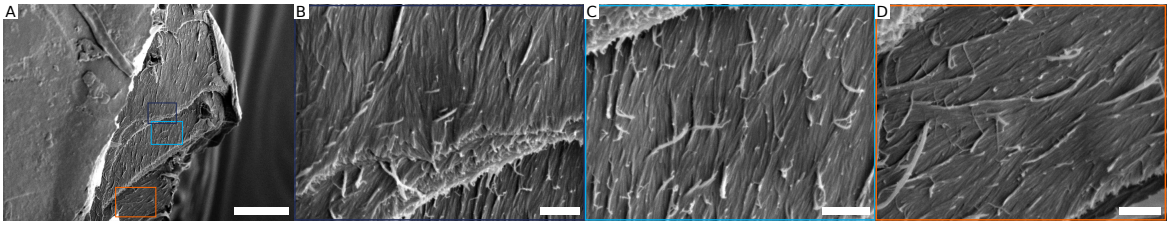


Fig. 5.16 SEM image (A) and zooms (B-D) of the defatted *M. nobilis* endocarp in top view, to estimate the width of the fibre bundles. Scale bars are 5 μm in (A), 500 nm in (B-D).

5.4 Xylan

A high amount of xylan was found in the *M. nobilis* endocarp, around 23 wt%, but only 6 wt% in the leaves of the tree, based on the NREL biomass quantifications in section 5.2. Furthermore, the xylan from the endocarp proved very difficult to be removed from the cellulose microfibrils, as described in section 5.3. These two findings for the endocarp, having high amounts of xylan in the tissue, and the xylan being hard to remove, indicate that xylan could play a major role in its helicoidal architecture of cellulose microfibrils in the secondary cell wall [112, 31].

Therefore, to ultimately build a model of the role of xylan in the cell wall helicoidal architecture and its interactions with the cellulose microfibrils, I investigated the structure of xylan. In more detail, I investigated what decorations it has, including the pattern of decorations with other monosaccharides, acetylation, the chain length and the position of (methylated) glucuronic acid groups. It is important to point out that the interpretation of the obtained results is not trivial, as the structure of xylan is not identical for every xylan molecule in a certain tissue. There will always be a distribution of xylan molecules with different backbone lengths and different decorations.

5.4.1 Backbone

The backbone of xylan is made up entirely of xylose residues, with a β-1,4-glycosidic linkage. This type of linkage in a pentapyranose configuration results in the xylan backbone being a

linear chain with equatorial configuration, equivalent to the β-1,4-glycosidic linkage in the hexapyranose glucose residues of cellulose [118].

It is very difficult to estimate the length of xylan molecules. Even though there are a few techniques available, like viscosity-based measurements, light scattering or size exclusion chromatography (SEC) [116, 82], there are various limitations to consider: xylan is usually found with some degree of acetylation, and various side chains, which can be charged, like for example glucuronic acid. All these aspects influence the interaction of xylan with its environment, like cellulose microfibrils and lignin, and also with itself, mainly its solubility. Thus it follows that the viscosity is strongly influenced by these factors, especially by potential charges. Concerning light scattering and size exclusion chromatography, solubility and unwanted precipitation are of concern, as well as a lack of appropriate xylan standards for SEC. For all these aspects, it is difficult to disentangle how big their contributions are and how they quantitatively affect the length measurements of the xylan backbone. For these reasons, I decided to estimate the xylan length from the sequencing experiments instead, subsection 5.4.4.

5.4.2 Side chains

Firstly, I examined what monosaccharides are attached to the xylose backbone. I extracted xylan with NaOH, removing any potential acetylation upon extraction. I first performed GH11 and GH30 digestions: GH11 cuts glucuronoxylan into XUXX, XX, X oligo- and monosaccharides, and GH30 cuts the GlcA-substituted xylan backbone at the -2 position of [Me]GlcA. Then, I used GH62, which cuts arabinose (Ara) decorations off the chain, and GH115, which cuts GlcA decorations off the chain. Both enzymes do not cut the xylose backbone. An image of the obtained gel is depicted in Figure 5.17. As mentioned in section 4.3, high voltage is applied to separate different oligosaccharides in gel electrophoresis, depending on their length and charges. For example, a fragment of XXXX would travel faster on the gel because of its smaller size than a fragment of XXXXX. Similarly, charged fragments migrate further than non-charged ones: XUXX will be further down than XXXX, because the charge contributes to the migration.

The following conclusions could be drawn from this experiment:

- No (significant) amount of arabinose decoration is found on the xylan. GH62 would cut Ara off the chain, but (almost) no difference was observed between the respective pairs, in lanes 2 and 3, and 6 and 7. If there was any Ara decoration, a fragment of e.g. XAXX would show between X_4 and X_5 on the gel. After GH62 digestion, the XAXX

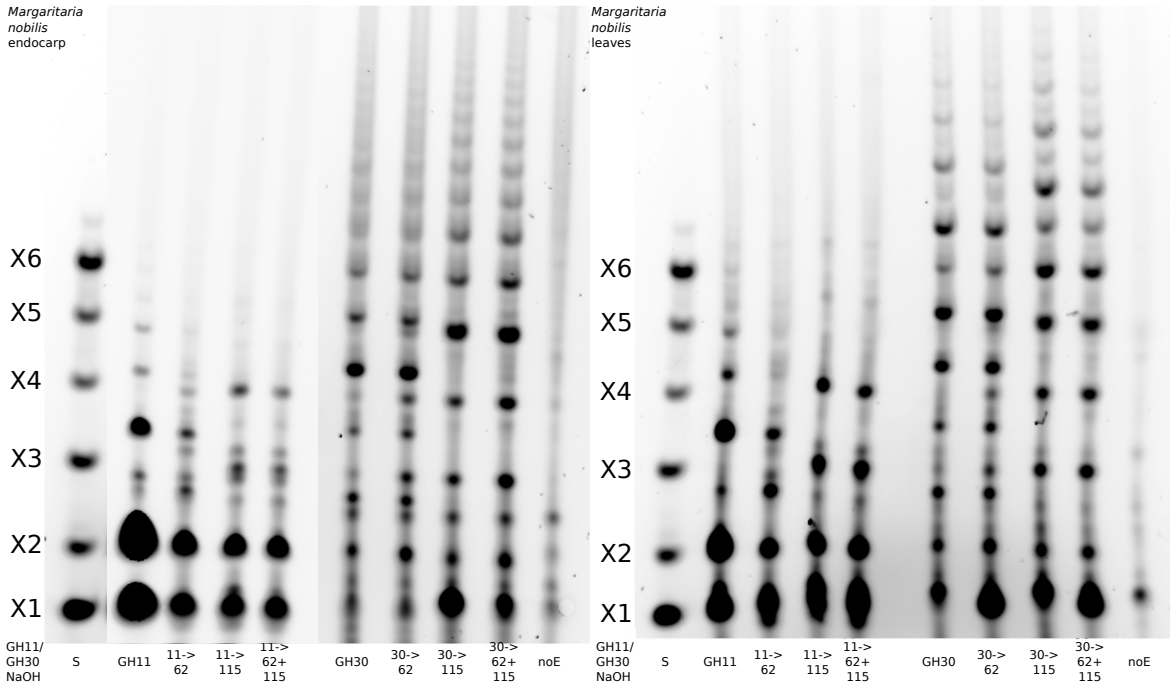


Fig. 5.17 Digestion of *M. nobilis* endocarp (left) and leaf (right) xylan with GH11 and GH30, and subsequently GH67 and GH115, analysed by PACE to investigate side chain decorations.

band would disappear and the X_4 band would appear or increase in intensity [1]. This result is in agreement with the biomass composition analysis, where only negligible amounts of Ara were found (not detected in the endocarp, and approx. 1.7 wt% in the leaves, section 5.2).

- Considerable amounts of glucuronic acid were found. GH115 cuts GlcA off the chain, and during the GH115 digestion of GH11 digestion products, XUXX was converted to XXXX. For the initial GH30 digestion, the same effects were observed upon addition of GH115. Additionally, XXXUX was reacted to XXXXX, and XXXXUX and XXXXXUX to X_6 and X_7 were observed as well.

- The amount of GlcA decoration was determined from the band sizes with the Fiji plug-in for imageJ (on an image with lower exposure, see Figure A.8): around 4 % of xylose units carry a glucuronic acid residue [120, 119]. This value is applicable for xylan that can be digested with GH11, but there might be more xylan that is not accessible for this enzyme (for example due to steric hindrance or it is in too tight interaction with cellulose). So this value could be lower when considering the entire xylan.

GUX1 and GUX2 are both glucuronyltransferases that decorate the xylan backbone with GlcA and 4-O-methyl-GlcA (summarised as [Me]GlcA) in specific patterns during biosynthesis [14]. In the endocarp, there are X_3UX, X_4UX, and X_5UX fragments found, which are characteristic fragments observed after GH30 digestion if there was GUX2 glucuronyltransferase activity during the xylan biosynthesis. However, it seems there was no GUX1 glucuronyltransferase activity during xylan biosynthesis, which produces xylans that would yield even fragments upon GH30 digestion, predominantly X_4UX, X_6UX and X_8UX. These fragments, which give a fingerprint pattern for GUX1 activity, are not observed in the endocarp sample. Hence, it is likely that during the endocarp biosynthesis, GUX2 is active and decorating the xylan backbone, but GUX1 is not.

The same experiment was also carried out for the leaf tissue, see Figure 5.17 on the right. Most bands look comparable, thus the same conclusions were drawn. Additionally, even oligosaccharides with GlcA were observed, X_4UX to X_6, X_6UX to X_8 and X_8UX to X_{10} (lane 6 to lane 8), which means there seems to be GUX1 and GUX2 glucuronyltransferase activity during biosynthesis.

The monosaccharide composition and GlcA content for the endocarp and leaves was additionally determined via HPAEC-PAD, the results are summarised below, Table 5.4. The obtained value for GlcA is probably underestimated by this technique for two reasons: firstly, the α-1,2-glycosidic bond between xylose and GlcA is considered stable, e.g. in comparison to the β-1,4-glycosidic bond along the xylan chain, so probably not all of the GlcA is released during TFA hydrolysis, and will thus not be detected via HPAEC-PAD. Secondly, there are no calibration standards available for methylated GlcA, and it is hence not known where it elutes and cannot be quantified either. The peaks suspected to be meGlcA are shown in the appendix Figure A.9. Since the value of approximately 4 % determined from the PACE gel is probably a slight overestimation, the two results can be considered to be in acceptable agreement.

5.4.3 Acetylation

The extent and pattern of acetylation was studied next. In general, xylan is in a three-fold helical screw conformation when recrystallised (based on X-ray diffraction and chain-packing models) or in solution (based on molecular dynamics modelling) [95, 18]. Recently, it was found that xylan can adopt a two-fold helical screw configuration when interacting with cellulose [18, 122].

Undecorated xylan could interact very tightly with cellulose microfibrils, but irregular decorations will introduce steric hindrance, hence making tight interaction impossible. If

Table 5.4 Concentration of monosaccharides from hemicellulose

	Endocarp		**Leaves**	
	mol%	mol%	mol%	mol%
	RMS	STDEV	RMS	STDEV
Rhamnose	1.03	0.06	24.56	0.61
Arabinose	1.00	0.04	9.22	0.10
Galactose	2.18	0.08	15.67	0.18
Glucose	16.74	1.47	6.34	0.26
Xylose	74.14	1.78	29.99	0.97
Mannose	2.07	0.29	5.59	0.32
Galacturonic acid	2.64	0.12	7.44	1.43
Glucuronic acid	0.31	0.01	0.42	0.08

acetylation is found on every other xylose residue, it means that in the two-fold screw conformation, all acetyl groups would point out on the same side, while the other side of the xylan backbone would be completely undecorated. For this undecorated side, tight interaction with cellulose microfibrils is possible and expected on the hydrophilic (010) and (020) faces [18]. This hypothesis of compatible and incompatible domains was explained in section 1.3.

As a first investigation of acetylation on the xylan backbone, I ran GH10 digestions and PACE gels of three differently extracted xylans: firstly, I isolated xylan with acetylation intact (Ac), secondly, I isolated xylan with acetylation intact and then treated it with base to remove acetylation (Ac→OH), and thirdly, I isolated xylan with base, so that the acetylation was removed upon extraction (OH). All three samples were compared, as presented in Figure 5.18, with the leaves (L) on the left, and the endocarp (E) on the right.

In Figure 5.18, lane 7 (the Ac endocarp lane) displays a long smear with only a few defined bands, but after base treatment (lane 8, Ac→OH), there were more defined bands and less smear visible. Smearing can originate for two reasons: acetylation or high salt concentration. If any xylose units carry an acetyl group, it can be on C-2 or C-3 in the pyranose ring, and it can transfer between the two. This dynamic phenomenon yields a distribution of different configurations, which would give a more continuous signal on the gel rather than defined bands. After removing the acetylation by the base treatment (lane 8, Ac→OH), the oligosaccharides are more homogeneous and thus migrate in more defined bands. The second reason, high salt concentration, is probably not the case here, since no excessive sample preparation procedures were necessary.

Since we can see a clear difference in smearing or defined bands for the Ac versus the Ac→OH sample respectively, we could conclude that there is a significant amount of

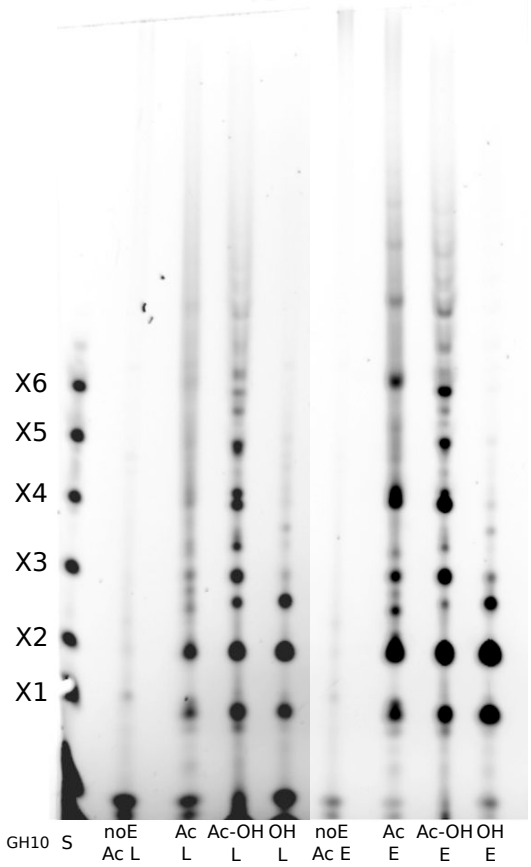


Fig. 5.18 Digestion of *M. nobilis* leaf (L) and endocarp (E) xylan with GH10 analysed by PACE to investigate acetylation. Ac: xylan extracted following the isolation of acetylated heteroxylan protocol, OH: xylan extracted by 4 M NaOH, Ac→OH: xylan first extracted as Ac followed by 4 M NaOH treatment.

acetylation in the endocarp xylan. For the leaves tissue, similar observations were made again.

In order to obtain more comprehensive information, the acetylated samples were digested with GH10 and subjected to MALDI-ToF MS (matrix-assisted laser desorption/ionisation - time of flight mass spectrometry) analysis (with support from Theodora Tryfona).

The mass spectrum in Figure 5.19 beautifully shows that a certain type of fragment has a higher peak intensity: fragments containing $2 \cdot n$ xylose units and $1 \cdot n$ acetyl groups clearly dominate the result. This phenomenon is particularly strong for the xylan from the endocarp. While for the leaf tissue (Figure 5.20), these peaks were also observed, they are by far not as dominant, and many other peaks were also present in the spectrum, showing that the

endocarp xylan is much more regular and organised, compared to the leaf xylan.

No peaks for fragments containing GlcA were found in the endocarp xylan mass spectrum, which was slightly unexpected but could have several reasons. Firstly, there is only a rather small amount of GlcA present in the xylan, secondly, GlcA-containing fragments could be difficult to evaporate for analysis, and thirdly, it could be because of the extraction method and enzyme digestion. In order to investigate acetylation, no base extraction, which would be very efficient, can be carried out. Hence, AIR was depectinated and delignified instead, to make xylan accessible. Possibly, not all of the contained xylan was rendered accessible this way, which should be verified (but could not be done in the scope of this thesis). Another option would be to carry out DMSO extractions, which are less efficient than base extractions, but do not remove the acetyl groups. Furthermore, it could be possible that there is some sort of steric hindrance for the GH10 enzyme due to the acetylation or way of preparing the material.

Finally, to further investigate whether the position of the acetyl groups is indeed on every other xylose unit, the acetylated xylan from the endocarp and from the leaf were compared to a well-characterised *gux1 gux2* double mutant from *Arabidopsis thaliana* [18]. To that end, all three samples were digested with GH10, and subjected to gel electrophoresis, the gel is shown in Figure 5.21.

The profile of the endocarp xylan looks very similar to the *A. thaliana gux1 gux2* double mutant, a very clear pattern of even degree of polymerisation (DP) is observed in both cases. This mutant does not have any GlcA decoration, but it has acetyl groups on every other xylose in the chain [88, 14]. As above, AIR was depectinated and delignified, and then subjected to GH10 digestion. Since there was no GlcA found via MALDI-ToF for this method, it can be assumed that it is not present in significant amounts here either. Thus, the gel pattern would not be compromised by GlcA contributions. For comparison, the PACE gel showing the GH10 and GH30 digestion of the *A. thaliana gux1 gux2* double mutant, published by Busse-Wicher et al. [18], is shown in Figure 5.21 B. In this case, after depectination and delignification of AIR, a DMSO extraction was carried out, yielding a comparable pattern. This result suggests that the *M. nobilis* endocarp xylan is indeed acetylated on every other xylose unit along the chain, while this regular pattern is not observed for the leaf xylan.

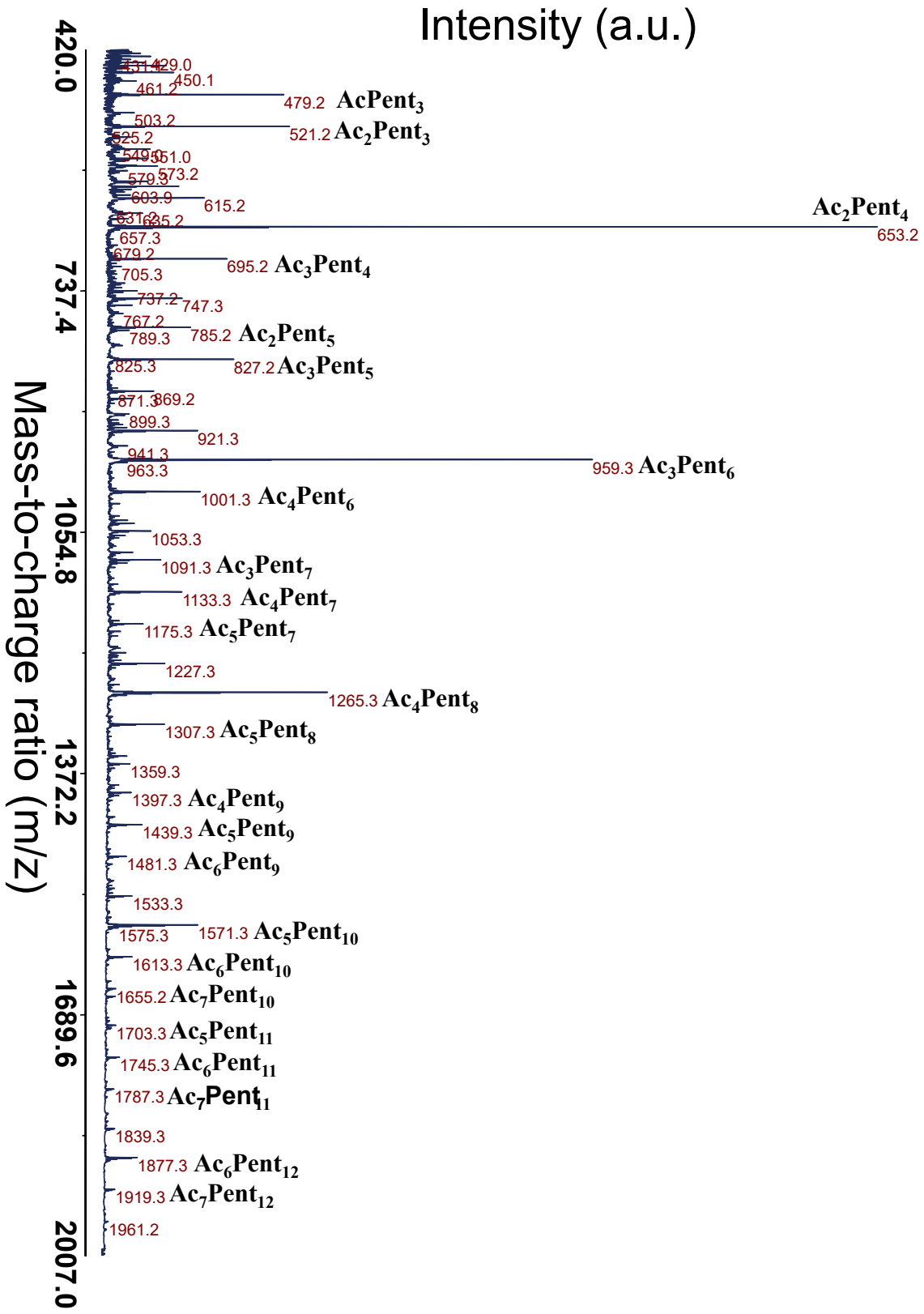


Fig. 5.19 MALDI-ToF MS spectrum of *M. nobilis* endocarp acetylated xylan digested by GH10.

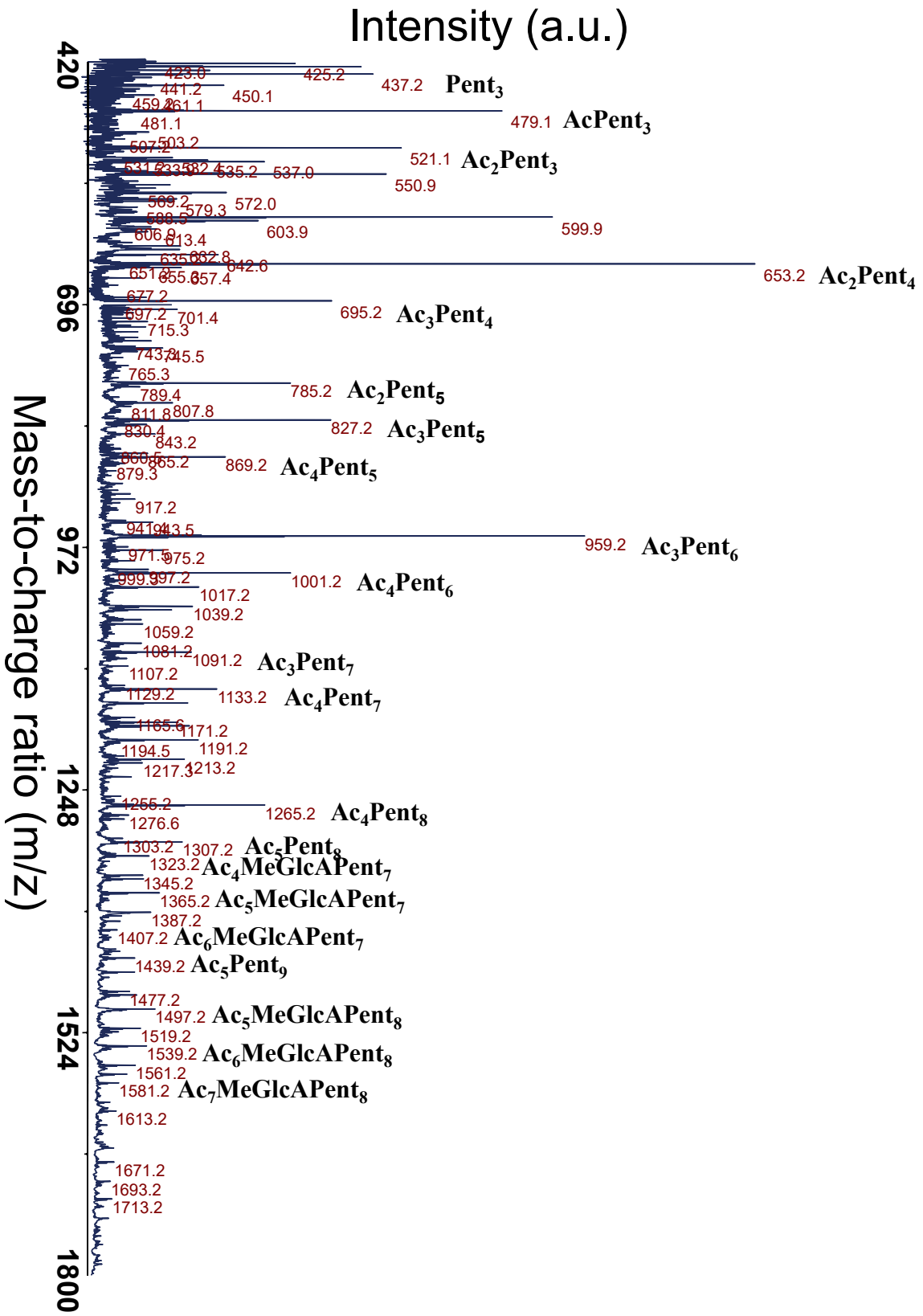


Fig. 5.20 MALDI-ToF MS spectrum of *M. nobilis* leaf acetylated xylan digested by GH10.

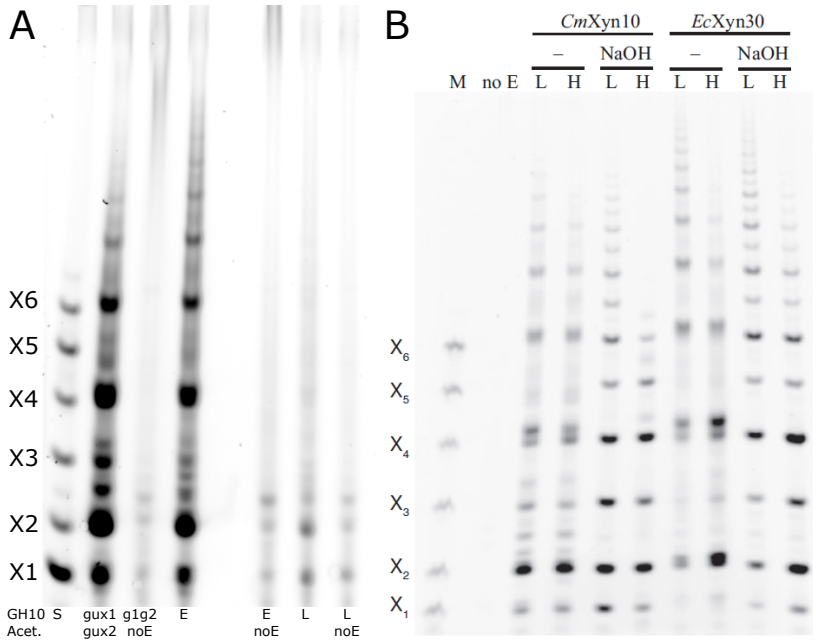


Fig. 5.21 (A) *Arabidopsis thaliana gux1 gux2* double mutant and *M. nobilis* endocarp (E) and leaf (L) xylan digested with GH10 and analysed by PACE to compare position of acetyl groups. (B) *A. thaliana gux1 gux2* double mutant digested with GH10 and GH30 and analysed by PACE, published by Busse-Wicher et al. [18].

5.4.4 Chain length and position of glucuronic acid decoration

The position of glucuronic acid decoration could play an important role in the helicoidal assembly of cellulose microfibrils: glucuronic acid can provide a negative charge, causing electrostatic repulsion, and might even influence twisting behaviour between the cellulose microfibrils or bundles [112, 31].

In this sequencing experiment, xylan was pre-labelled before any enzyme digestions and excess labelling agent was washed out. Then, sequential digestions by GH3 xylosidase and GH67 were carried out alternatingly, taking aliquots after each step. GH3 xylosidase cuts off xylose units from the unlabelled chain end, the non-reducing end, but only as long as there are no decorations on the chain. GH67 cuts off GlcA from the xylose backbone from the non-reducing end, only from the last xylose residue. By using those two enzymes alternatingly, and taking aliquots between each step, we might be able to estimate the distance between the glucuronic acid decorations.

This experiment is a very new approach and has not been intensively studied and tested with other systems yet. There are still many other possibilities of how to interpret the results, and all the ideas and conclusions are very preliminary. It is additionally complicated by the

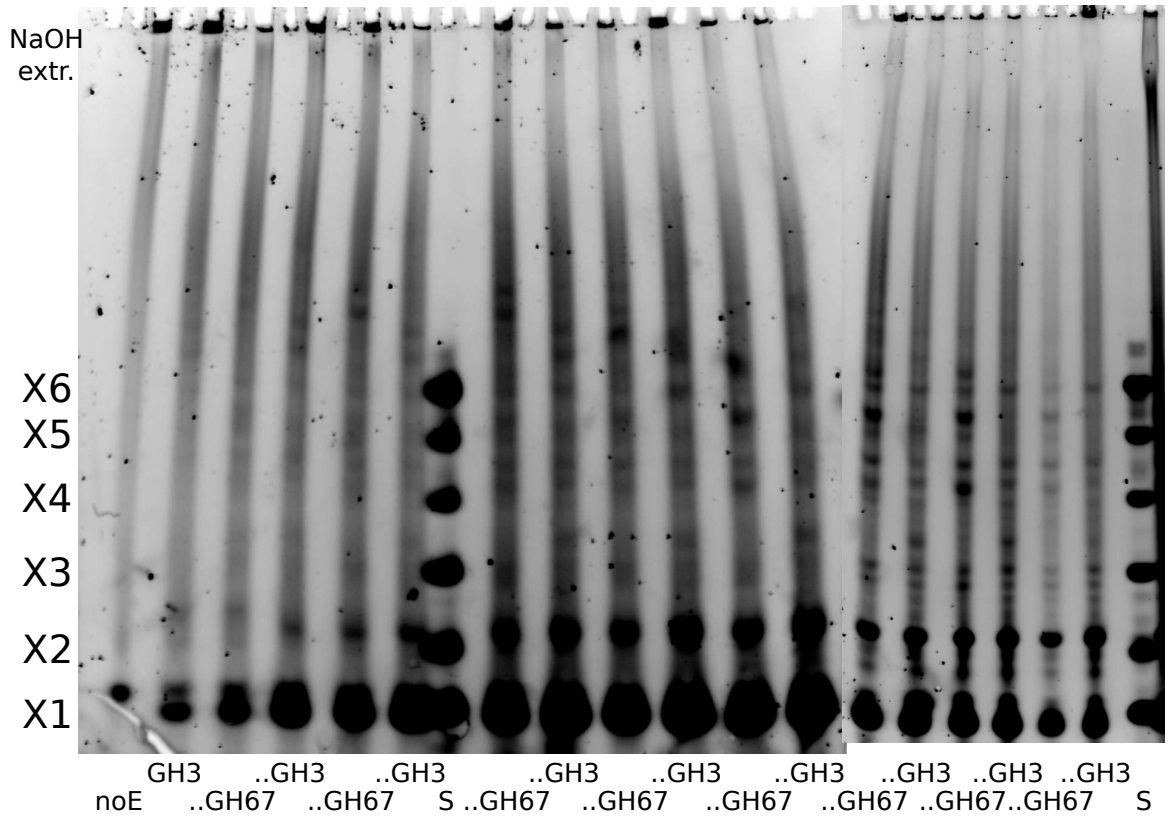


Fig. 5.22 Sequential digestion of *M. nobilis* endocarp xylan by GH3 xylosidase and GH67 analysed by PACE to estimate chain length and position of GlcA decoration along the chain.

distribution of different populations of xylan, which will probably lead to broadening of all the bands.

Figure 5.22 and Figure 5.23 show the obtained PACE gels for the endocarp and the leaves, respectively. For the endocarp xylan, possibly the chains became shorter with every cycle of enzyme digestions, since it seems that dark bands move down a bit further after every cycle. Furthermore, the dark bands seem to move down fairly evenly and there does not seem to be any abrupt steps between cycles, which could indicate that the GlcA is possibly distributed along the chain fairly evenly. Moreover, after around 8 cycles, there were already bands in the range of X_6 visible. This could indicate that the xylan is maybe fairly short and there are possibly not more than 8 or 9 GlcA decorations along the chain. For the leaf xylan, the bands did not seem to move down much at all. Possible reasons could be that it is initially much longer, it could have additional decorations that hinder the enzyme efficiency, etc.. In both cases, we also observed that some proportion of xylan stayed in the wells, instead of migrating during electrophoresis, one possible reason being precipitation.

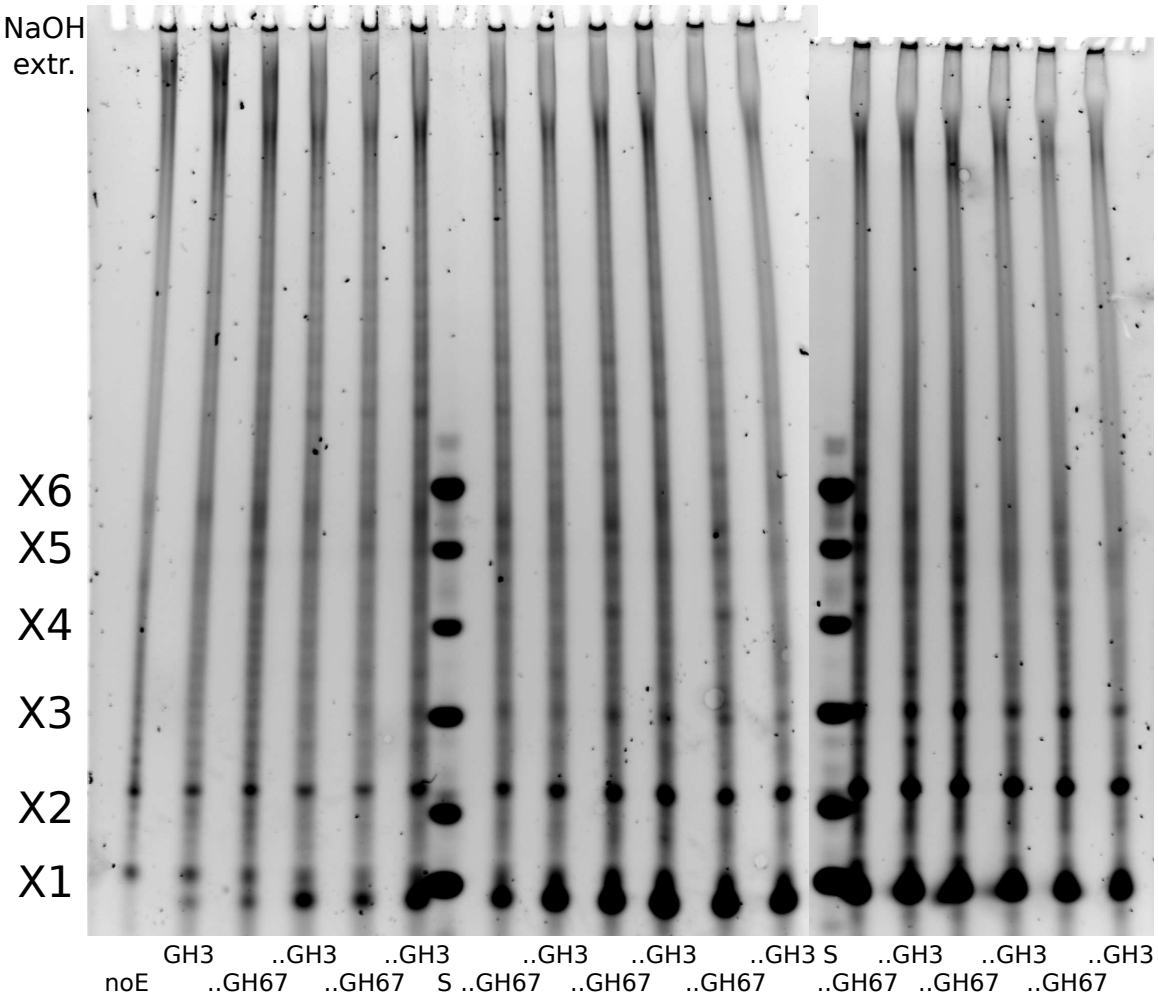


Fig. 5.23 Sequential digestion of *M. nobilis* leaf xylan by GH3 xylosidase and GH67 analysed by PACE to estimate chain length and position of GlcA decoration along the chain.

However, all these ideas are highly speculative. There could be many other influences, like for example there could simply be thermal degradation of the xylan chains, since the samples are heated to 100 °C after every digestion to inactivate the enzymes. Also, there could be a high proportion of broken chains, and we do not know how long the initial, undigested chains are. Moreover, issues with the pre-labelling of the sample and washing out the excess labelling agent cannot be ruled out either.

5.4.5 Summary of experimental results

- Firstly, xylan from the structurally coloured endocarp tissue is more regular and organised in terms of decorations than the xylan from the leaves.

- For the endocarp xylan, acetylation is found on every other xylose unit along the chain, as speculated based on MALDI-ToF MS experiments and co-migration with the *A. thaliana gux1 gux2* double mutant.
- There is no arabinose decoration, as concluded by GH62 digestions and monosaccharide quantification after acid hydrolysis (section 5.2).
- There is about 4 % of glucuronic acid decoration, overall every 1 in 25 xylose units should carry a GlcA. Furthermore, the substitution pattern characteristic for GUX2 glucuronyltransferase activity during xylan biosynthesis is observed for the endocarp xylan, namely X_3UX, X_4UX and X_5UX identified from GH30 and GH115 digestions.
- All these findings for the endocarp xylan combined mean that there is alternating acetylation and no arabinose decoration. One hypothesis for the GlcA substitution pattern is that there are longer stretches with only even acetylation but no glucuronation, and then stretches with clustered GlcA decoration - there should be some clustered areas since a GH30 pattern of X_3UX, X_4UX and X_5UX was observed. At the same time, there have to be stretches without GlcA decoration, since a continuous GUX2 decoration would require more than approximately 4 % of glucuronation. From these implications, a tentative xylan structure is suggested, depicted in Figure 5.24. No conclusion about the position of these potential clusters of GlcA decoration can be drawn, so a chain with clusters somewhere along the chain, at or close to the non-reducing end, or at or close to the reducing end are possible.
- (Another option is that there is a population of xylan with very low or no GlcA decoration but only even acetylation, and another population with more GlcA. This hypothesis could be tested by ion exchange chromatography and eluting xylan with buffer with increasing ionic strength, but is beyond the scope of this thesis [14]. An indication for this scenario is that no acidic sugars could be detected in the MALDI-ToF and acetylation experiments.)
- For the leaves xylan, all the results are much less distinct and pronounced, meaning that it is much less regularly organised and probably longer.

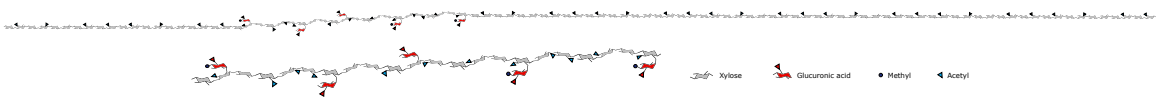


Fig. 5.24 Tentative structure of *M. nobilis* endocarp xylan: an example chain, and magnification of the proposed area of clustered GlcA decorations.

5.5 Small-angle X-ray scattering analysis

The SAXS profile obtained for the *M. nobilis* endocarp, taken tangentially to the endocarp surface, is depicted in Figure 5.25. Using this configuration, we aimed to obtain the structural information from the cross section of the cells. Data processing and interpretation is still ongoing. We aim to extract i) the form factor of the individual scattering body (which corresponds more or less to the cellulose microfibrils), and ii) the structural factor giving information about the organisation of the scattering bodies (which should correspond to the interfibrillar distance). In this case, the finite length of the cellulose microfibrils has to be considered as well, and we hope to facilitate SAXS data processing by eventually providing morphology parameters obtained from TEM and AFM measurements, to decrease the amount of unknown parameters. SAXS analysis is carried out by Dr. Yu Ogawa.

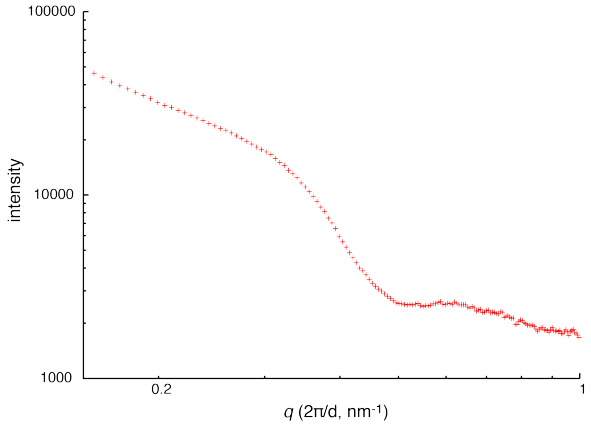


Fig. 5.25 SAXS profile of the *M. nobilis* endocarp, taken tangentially to the endocarp surface.

5.6 Coarse grain molecular dynamics modelling of cellulose-xylan assembly

The morphological and chemical informations provided in this thesis were then used as starting parameters to model the cellulose-xylan interactions in the helicoidal cell wall by Dr. Yu Ogawa. To set up the simulation, the cell wall deposition of polysaccharides was approximately imitated by having a fixed layer of cellulose microfibrils, a mobile layer of cellulose microfibrils, and xylan inbetween, as shown in Figure 5.26 A and B. The system was hydrated, and then slowly dried: After 20 ns equilibration at 300 K with gradual drying, a rotation angle of about 8° was observed, as shown in Figure 5.26 C and D. However, the angle varied between simulations with different starting configurations.

So far, this preliminary result suggests that the formation of a helicoidal architecture composed of cellulose microfibrils and xylan can be reproduced by molecular modelling. Further study is needed to address the effect of decoration on xylan molecules (like the charge provided by glucuronic acid decoration, and the steric effect from acetylation), the effect of the hydration-drying procedure on the modelling, and the size effect of the *M. nobilis* endocarp cellulose microfibrils and xylan, as well as the system size itself.

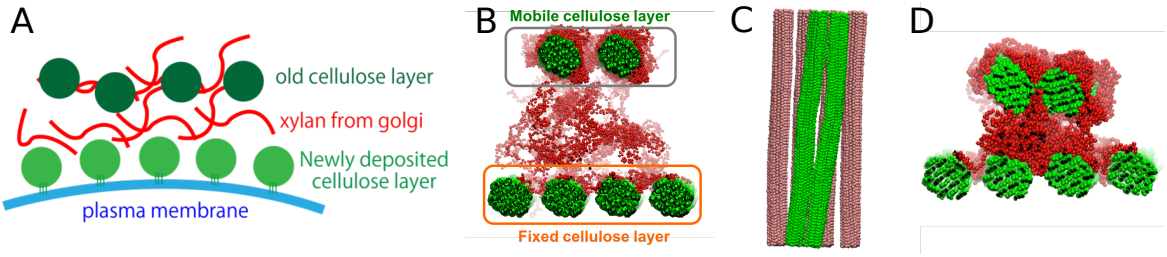


Fig. 5.26 Coarse grain molecular dynamics modelling of the cellulose-xylan assembly. (A, B) Starting configuration. (C, D) Rotation angle of about 8° observed after 20 ns equilibration at 300 K with gradual drying.

5.7 Outlook and discussion

5.7.1 Outlook

A few possible experiments, some of them already ongoing, to elucidate the ultrastructure of the helicoidal architecture in more detail, are discussed in this section.

A very important parameter to determine is the precise morphology of the cellulose microfibrils composing the helicoidal cell wall architecture, and then the distance between the microfibrils in the tissue. This information will allow us to better infer the role of xylan in the structure. Future plans include further investigating which of the two (or both) identified populations of cellulose microfibrils constitute the helicoidal structure. Current results indicate that the second population of shorter microfibrils is the more suitable candidate. These microfibrils are found in batch 1 and batch 2 endocarp, but most of them are sparse in TEM imaging as they are probably mainly in aggregates. This second population was still monitored when I determined the (200) crystallite size of 4.02 nm in X-ray diffraction measurements [98].

Unfortunately more base treatments, DMSO extractions or tip sonication did not seem to help concerning possible aggregation, and the second population of microfibrils are probably much more likely to aggregate because of their short length [51]. An option to investigate

whether most of the second population of microfibrils is aggregated is to add sulphuric acid to the sample mixture. It will introduce unwanted hydrolysis, but it might also be able to break up the aggregates slightly and add charges to the surface of the microfibrils, which could then stabilise them in water and make their detection by microscopy easier.

Once we have a better understanding of the distribution of the two microfibril populations in the cell wall, or at least their relative ratios, thorough investigations of the morphology should be carried out - length distributions via TEM and the microfibril height via atomic force microscopy (AFM).

These results will help with processing and interpretation of SAXS data, which is very complex. Generally, when SAXS is carried out on plant material, infinite fibril length in one dimension is assumed, allowing simplification of data processing. However, for both cellulose microfibril populations found in *M. nobilis* endocarp, this assumption is not valid. Furthermore, the cell shape and the alignment of the endocarp with respect to the beam need to be taken into account as well, making the experiment very challenging. Extracting the interfibrillar distance from the measurement data will allow us to propose a cell wall model for the endocarp. Moreover, it will also possibly allow us to answer several of the open questions regarding the interaction of cellulose with xylan: whether it will be possible for the xylan to span the distance between the microfibrils, whether the microfibrils align directly next to each other or in a xylan matrix, how many microfibrils make up a fibre bundle, what the layer distance between layers in a helicoidal pitch is and finally what their rotation angle is.

A way to investigate this question would be to isolate the microfibrils and the xylan and then see if the helicoidal architecture can be replicated *in vitro*. However, even if the structure is indeed obtained *in vitro*, it does not necessarily mean that the same interactions regulate the growth of the SCW. Only if it does not happen, it can be excluded as a mechanism - or some aspects of the experimental set-up did not work out and have to be adapted.

Moreover, coarsed-grain modelling of the system might help to shed light on the interaction between cellulose and xylan. Again, only if there is no helicoidal arrangement observed, we can conclude that this is not a mechanism in the cell wall (or the parameters have to be optimised). If it does work, it merely means that helicoidal arrangement could be possible in this way. Currently, our collaborator Dr. Yu Ogawa developed a working model for the simulation, based on a fixed layer of cellulose microfibrils that are anchored at the plasma membrane, and a mobile layer of cellulose microfibrils on top of it. However, it does seem difficult that the entire part of the cell wall, that is already laid down, can rotate in the cell, when not only finite cellulose microfibrils are investigated but the entire plant cell wall. It seems much more likely that the cellulose microfibrils and xylan are already arranged in a

helicoidal way to begin with.

5.7.2 Discussion

The biosynthesis of the secondary plant cell wall is a very complex process, of the precisely orchestrated deposition of cellulose microfibrils by the CSC, the deposition of hemicelluloses transported from the Golgi, and subsequent lignification. Even for model species like *A. thaliana*, many questions are unanswered and many pathways amd mechanisms are probably not even discovered yet. Understanding the helicoidal arrangement of cellulose microfibrils in the secondary plant cell wall is a humongous research question, and even though this thesis presents a detailed study of the system, we can only speculate about some of the mechanisms.

There are many open questions that will be discussed in this section:

- **What are the prerequisites for the helicoidal arrangement?** Which are the key parameters playing a role in the helicoidal arrangement ?
- Is there any specific type of morphology and crystalline properties of cellulose microfibrils required to build a helicoidal cell wall structure ?
- How does the primary structure of hemicellulose, xylan in this case, affect the arrangement? Do the decorations play a role, what substitutions there are and in what pattern?
- What is the role of lignification? Is it essential to shift the reflection into the visible range? Is the lignin content crucial to obtain blue colouration? Is the position of enzymes initiating polymerisation crucial?
- Is the exact composition of the endocarp essential? The ratios of cellulose to hemicellulose to lignin to moisture content?
- What other parameters play a role?
- **What induces the twist?**
- At what level of the hierarchy is the twist induced? Is it the cellulose microfibril itself, is it the xylan, is it the fibre bundles?

5.7.3 Cellulose

The cellulose microfibrils making up the helicoidal secondary cell wall were found to be rather exceptional because of their short length. There are two populations of fibrils, and it was not yet possible to understand whether they have distinct roles at different locations or whether they are mixed together and interact to form the helicoidal structure. One hypothesis could be that the two populations constitute either the primary or the secondary cell wall. Unfortunately during the isolation procedure, both are purified together and there is no way to separate the two beforehand.

Both populations are exceptional because they are surprisingly short: less than about 1 or 2 micrometres for the first population, and only a few hundred nanometres for the second population. Cellulose microfibrils from many other sources, like cotton, wood or other fruit, are found to be at least a few micrometers long, and it is usually challenging to determine their length, since it is hard to see their beginnings and ends [96]. Moreover, the second population of microfibrils found in the endocarp seems to have a significantly bigger width than the first population.

Considering the morphologies of the two populations, it might be more likely that the second population of microfibrils composes the helicoidal architecture. Generally, the width of SCW microfibrils is bigger than PCW microfibrils. More CSC populate the plasma membrane during SCW biosynthesis than during PCW biosynthesis, and they are more clustered and move in the same direction along the microtubules. This way, it is more feasible to produce thicker microfibrils, since directly after biosynthesis, they can aggregate more easily [78]. The second population of microfibrils found in the endocarp seems to have a significantly bigger width than the first population, thus hinting at their prevalence in the SCW.

Either way, the next obvious question to ask is, what makes the cellulose microfibrils so short? Are they biosynthesised in such a short way? Or are there enzymes in the SCW that cleave them short after being extruded? Cellulose hydrolysis is a difficult task, because of the high crystallinity and hierarchical structure, and usually requires a combination of various cellulases. Most often, enzymatic cleavage does not result in a clear cut through the cellulose microfibril, but rather in thinned areas or ends [56]. Furthermore, accessibility of the microfibrils might be difficult in the SCW, since they are found in much bigger bundles than is common for PCW microfibrils, and the SCW is generally much thicker. For these reasons, it seems unlikely that cellulose microfibrils are cut short after biosynthesis, suggesting that they are produced in such short dimensions.

If it is not enzymes cleaving the cellulose microfibrils after biosynthesis, more questions arise. Is this short length correlated to their relatively high crystallinity? Are they so short

because they are highly crystalline, or are they highly crystalline because they are so short? Or are these two parameters not interdependent at all? The second hypothesis would be favourable when considering that it is statistically less likely to have defects in a shorter than in a longer stretch. However, the first hypothesis might be more plausible when thinking that the highly cyrstalline fibrils are so stiff that they cannot follow the microtubule tracks along too much curvature and thus have to be terminated.

I think it is more likely that the cellulose microfibrils are produced in such a short way. Cellulase activity is a very complicated process, and the accessibility of the cellulose microfibrils also seems hindered in the SCW. Furthermore, the length distribution seems fairly narrow, which appears very difficult to obtain by enzymatic cleavage after biosynthesis.

5.7.4 Hemicellulose

It is likely that hemicellulose does play a key role in this helicoidal structure for two reasons.

Firstly, it seems very tightly associated with the cellulose microfibrils, since it took many purification steps to dissolve and remove it from the endocarp tissue. It was much easier to remove the hemicellulose from the leaves tissue than from the endocarp tissue.

Secondly, the endocarp xylan does seem to be exceptionally regular when compared to the leaves xylan. It has acetylation on every other xylose unit, the glucuronic acid is probably distributed fairly evenly along the chain, and there does not seem to be any other decoration present, like e.g. L-arabinose. This could be an indication that the tight control over its biosynthesis is necessary to build up this helicoidal structure. If not, it would not be necessary to regulate the xylan biosynthesis so tightly and have so much control over it.

5.7.5 Lignin

Lignin is possibly less likely to influence the development of the helicoidal architecture. Lignification only occurs once the cellulose-hemicellulose composite has been laid down [10, 140]. It is unlikely that this polysaccharide structure is rearranged after its initial biosynthesis, it rather has to be made this way in the first place. It seems very difficult to rearrange all the components along the entire thickness of the SCW in the entire cell after deposition. For this reason, it seems that lignin only shows up after the helicoidal arrangement has started.

However lignin could have an influence on the development of colour. A helicoidal architecture in general is very widespread in many plant tissues, it might be caused by the way the cells grow and expand, and it might improve mechanical stability, among other

reasons. In most plants, the helicoidal pitch p, the length of the helicoidal axis after a 180° rotation of the cellulose microfibrils, is in the range of micrometres, by far exceeding the visible range.

This scenario is clearly not the case for the *M. nobilis* endocarp, but the opposite situation could be possible. It might be possible that during the first stages of SCW biosynthesis, the helicoidal pitch p is too short for reflection in the visible and that cellulose and hemicellulose are arranged so tightly that reflection would only occur in the UV range or at even lower wavelengths. It might be that only when lignin is deposited in the structure, the distance between the layers of cellulose microfibrils is increased and the architecture starts to reflect in the visible range.

This hypothesis could be tested by investigating immature fruits: if lignification does not contribute in that way, LCP reflection would be observed from the beginning, and it would only increase in intensity as the SCW grows and more layers are added. However if it did play a role, no LCP reflection would be observed at first, and then it would appear and probably shift in wavelength from the UV to slightly longer wavelengths towards the blue and greenish region of the visible spectrum as more lignin is added [68]. In practice, it would be hard to observe unambigously, since many other factors influence the reflection as the SCW develops: the cell shape could still change slightly, or levels of hydration may fluctuate and influence the reflection, to give just two examples.

5.7.6 Hierarchy

Another open question is at what length scale the twist originates? Is it at the molecular level of glucose chains, at the level of cellulose microfibrils, or is it imparted at the fibre bundles [59]?

Glucose and the glucose chains are chiral molecules, so it might be possible that the twist simply stems from this shortest of the hierarchy levels. However I do not think it is likely, while it might be a prerequisite for a chiral structure, it cannot be the only determining factor, since then reflection in the visible light range from helicoidally arranged SCW would be a much more abundant phenomenon in the plant kingdom. Furthermore, plant tissues showing opposite chirality have been observed [147].

I think a similar conclusion applies in the case of the twist originating from the cellulose microfibrils, that if it were the case, reflection in the visible light range would be more abundant. However this hypothesis might be more likely since more parameters could be optimised for twisting: the precise shape and dimensions of the cross-section of the microfibrils, their crystalline properties, the presence of defects along the chain, their lengths and aspect ratio to name but a few. However even if the cellulose microfibrils were the

only reason to obtain such a helicoidal architecture, a plethora of parameters and pathways would have to be tightly controlled to produce cellulose microfibrils with exactly the required parameters and a multitude of pathways and cell biology processes would need to be involved.

If the twist originates at the level of fibre bundles, all the previous factors discussed for microfibrils might still be relevant, as well as the properties and interactions of cellulose with xylan.

5.7.7 Twisting

Finally, what causes the structure to twist? There are many possibilities, and it could be a combination of all these points and more factors I have not even thought about. Firstly, what role do cortical microtubules play? A simple explanation for the helicoidal arrangement of cellulose microfibrils could be that the CMT move in this pattern. The fibrils might simply be deposited more or less parallel, and then for the following layer, the CMT would move at a slight angle to the first track they did. CMT movement is still being investigated - it is still under discussion whether they have some sort of active propulsion or whether they are only pushed along the plasma membrane by the nascent cellulose microfibrils [104]. Furthermore, even if they moved in a parallel fashion in the first layer, there is little evidence of control mechanisms used to set about the orientation in the second layer and how to move in a slight angle towards it? And finally, even if the twist was caused by the CMT movement, additional processes would be necessary to keep the microfibrils short, either by stopping further extension or cleavage by enzymes. If they were helicoidally arranged but very long, the structure would most likely be too disordered to give this strong and defined optical response.

Finally, at what point is the twist introduced? Is it introduced due to biological control, or is it purely physical? What role does the xylan play? Might the glucuronic acid groups provide a charge causing electrostatic repulsion that leads to the initiation of the twisting? Is the precise decoration and pattern of xylan essential? Does it maybe, at the point of a newly nascent cellulose microfibril, determine its direction of deposition by rotating away from glucuronic acid charges from the previous layer? This idea was explored for seed mucilage in quince seeds by Reis and Vian, comparing it to the self-assembly of chiral nematic suspensions of cellulose nanocrystals [112].

Most likely, in my opinion, it is a combination of many factors, like the precise morphology and crystalline properties of the cellulose microfibrils, the precise structure of xylan and the exact timescale of deposition of the polysaccharides, moisture, positioning of enzymes for lignification, ratio of the different polymers to each other.

It will be important to investigate systems with similar optical properties in the same depth as this one, to establish whether there are any factors they have in common in different locations of the phylogenetic tree. For example, preliminary results on the monocot *P. condensata* suggest a similar distribution of different populations of cellulose microfibrils. At the same time, not much extractable xylan was found in its pericarp in preliminary experiments, so it might either not play a role, or its role might be fulfilled by another type of hemicellulose with a similar structure (predominant hemicelluloses in the SCW of monocots are glucuronoarabinoxylans, but they might need different conditions to be extracted and digested for PACE analysis [118]).

All in all, the components of the SCW of the iridescent endocarp of *M. nobilis* have been thoroughly investigated in this chapter. Some hypotheses could be made with regards to requirements to build up such a helicoidal architecture. At the same time, like any good fundamental research, it has opened up many more questions, and it is very exciting to hope for them to be looked into in the future.

Chapter 6

Conclusion

In this thesis, structural colour in plants was investigated. I focused on a specific architecture, the helicoidal arrangement of cellulose microfibrils in the secondary cell wall. This structure was characterised thoroughly from an optical point of view in the epidermal cells of the fronds of the fern *Microsorum thailandicum* (chapter 2, chapter 3), and from a chemical point of view in the fruit endocarp of the eudicot *Margaritaria nobilis* (chapter 4, chapter 5).

I found that the fronds of *M. thailandicum* exhibit striking optical properties: both the adaxial and the abaxial epidermis reflect strongly in the visible light range. Every cell reflects a different wavelength, and the maxima of the adaxial surface lie in the blue range of the spectrum, 400-550 nm, while the maxima of the abaxial surface are much more varied and cover almost the entire visible spectrum, 400-650 nm. Furthermore, the spectra for the abaxial epidermal cells are much broader with lower intensity, and show more different features compared to the adaxial epidermal cells. The same extent of variation as observed between different cells was also confirmed within individual cells, by mapping the spectral response with high resolution of about 20 spectra per single cell.

The optical behaviour was furthermore related to the ultrastructure of the fronds, and the helicoidal pitch p was determined from electron microscopy images and related to the observed reflections. Finally, disorder in the helicoidal structure was assessed by correlating optical reflection and ultrastructure, and by modelling the system with the Berreman 4x4 approximation.

I can conclude that there is disorder in the helicoidal structure which is not caused by defined, localised defects but is more complex. Moreover, the fact that the adaxial epidermal secondary cell wall exhibits a much narrower reflection and has thus much tigher control on its biosynthesis than the abaxial epidermal secondary cell wall makes *M. thailandicum* an

incredibly interesting plant to study cell wall biosynthesis and composition.

I characterised the composition and molecular building blocks of the secondary cell wall of the endocarp of *M. nobilis* in great detail, to correlate it to the formation of the helicoidal architecture.

I determined the composition of the endocarp as 9 wt% lipids, 36 wt% cellulose, 23 wt% hemicellulose, and 30 wt% lignin. As the most abundant compound, I found that there are two distinct populations of cellulose microfibrils in the endocarp tissue: the first population is less than 1 or 2 μm long and with a thin diameter, and the second population is only a few hundred nanometres long, and with a slightly thicker width. Furthermore they exhibit high crystallinity as determined by nuclear magnetic resonance spectroscopy and X-ray diffraction.

Moreover, I found that the xylan in the secondary cell wall is interacting very tightly with the cellulose microfibrils and unusually regular. It exhibits an even pattern of acetylation, no arabinose decoration, and approximately 4 % of glucuronic acid decoration in regular intervals along the backbone.

Furthermore, coarsed-grain modelling and small-angle X-ray scattering studies are ongoing, which will help shed light on the interactions and assembly mechanism of this helicoidal architecture. A myriad of factors and hypotheses are discussed to conclude this study on *M. nobilis*.

References

[1] Anders, N., Wilkinson, M. D., Lovegrove, A., Freeman, J., Tryfona, T., Pellny, T. K., Weimar, T., Mortimer, J. C., Stott, K., Baker, J. M., Defoin-Platel, M., Shewry, P. R., Dupree, P., and Mitchell, R. A. C. (2012). Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proceedings of the National Academy of Sciences*, 109(3):989–993.

[2] Anderson, C. T. (2018). Finding order in a bustling construction zone: quantitative imaging and analysis of cell wall assembly in plants. *Current Opinion in Plant Biology*, 46:62–67.

[3] Andersson, S., Serimaa, R., Paakkari, T., Saranpää, P., and Pesonen, E. (2003). Crystallinity of wood and the size of cellulose crystallites in Norway spruce (*Picea abies*). *Journal of Wood Science*, 49(6):531–537.

[4] Arens, N. C. (1997). Responses of leaf anatomy to light environment in the tree fern *Cyathea caracasana* (Cyatheaceae) and its application to some ancient seed ferns. *Palaios*, pages 84–94.

[5] Atmodjo, M. A., Hao, Z., and Mohnen, D. (2013). Evolving views of pectin biosynthesis. *Annual review of plant biology*, 64.

[6] Avolio, R., Bonadies, I., Capitani, D., Errico, M., Gentile, G., and Avella, M. (2012). A multitechnique approach to assess the effect of ball milling on cellulose. *Carbohydrate Polymers*, 87(1):265–273.

[7] Baley, C. (2002). Analysis of the flax fibres tensile behaviour and analysis of the tensile stiffness increase. *Composites Part A: Applied Science and Manufacturing*, 33(7):939–948.

[8] Barnett, J. R. and Bonham, V. A. (2004). Cellulose microfibril angle in the cell wall of wood fibres. *Biological Reviews*, 79(2):461–472.

[9] Berreman, D. W. (1972). Optics in stratified and anisotropic media: 4×4-matrix formulation. *Journal of the Optical Society of America*, 62(4):502–510.

[10] Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annual Review of Plant Biology*, 54(1):519–546.

[11] Bonawitz, N. D. and Chapple, C. (2010). The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype. *Annual Review of Genetics*, 44(1):337–363.

[12] Boonkerd, T. and Nooteboom, H. P. (2001). A new species of Microsorum (Polypodiaceae) from Thailand. *Blumea*, 46:581–583.

[13] Bouligand, Y. (1972). Twisted fibrous arrangements in biological materials and cholesteric mesophases. *Tissue and Cell*, 4(2):189–217.

[14] Bromley, J. R., Busse-Wicher, M., Tryfona, T., Mortimer, J. C., Zhang, Z., Brown, D. M., and Dupree, P. (2013). GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. *The Plant Journal*, 74(3):423–434.

[15] Bruker Nano Inc. (2019). Bruker AFM probes. https://www.brukerafmprobes.com/p-3864-otespa-r3.aspx.

[16] Buchanan, B. B., Gruissem, W., and Jones, R. L. (2009). *Biochemistry & molecular biology of plants*. American Soc. of Plant Physiologists, Rockville, Md, 7. reprint. edition.

[17] Buchtová, N. and Budtova, T. (2016). Cellulose aero-, cryo- and xerogels: towards understanding of morphology control. *Cellulose*, 23(4):2585–2595.

[18] Busse-Wicher, M., Gomes, T. C. F., Tryfona, T., Nikolovski, N., Stott, K., Grantham, N. J., Bolam, D. N., Skaf, M. S., and Dupree, P. (2014). The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of *Arabidopsis thaliana*. *The Plant Journal*, 79(3):492–506.

[19] Busse-Wicher, M., Grantham, N. J., Lyczakowski, J. J., Nikolovski, N., and Dupree, P. (2016a). Xylan decoration patterns and the plant secondary cell wall molecular architecture. *Biochemical Society Transactions*, 44(1):74–78.

[20] Busse-Wicher, M., Li, A., Silveira, R. L., Pereira, C. S., Tryfona, T., Gomes, T. C. F., Skaf, M. S., and Dupree, P. (2016b). Evolution of xylan substitution patterns in gymnosperms and angiosperms: implications for xylan interaction with cellulose. *Plant Physiology*, page pp.00539.2016.

[21] Carter, I. E., Weir, K., McCall, M. W., and Parker, A. R. (2016). Variation in the circularly polarized light reflection of *Lomaptera* (Scarabaeidae) beetles. *Journal of The Royal Society Interface*, 13(120):20160015.

[22] Castany, O. (2016). Berreman4x4. https://github.com/Berreman4x4/Berreman4x4. Accessed 2016-08-21.

[23] Cazetta, E., Zumstein, L. S., Melo-Júnior, T. A., and Galetti, M. (2008). Frugivory on *Margaritaria nobilis* l.f. (Euphorbiaceae): poor investment and mimetism. *Brazilian Journal of Botany*, 31(2):303–308.

[24] Cordón, G. B. and Lagorio, M. G. (2007). Optical properties of the adaxial and abaxial faces of leaves. Chlorophyll fluorescence, absorption and scattering coefficients. *Photochemical & Photobiological Sciences*, 6(8):873.

[25] Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology*, 6(11):850–861.

[26] Cosgrove, D. J. (2018a). Diffuse Growth of Plant Cell Walls. *Plant Physiology*, 176(1):16–27.

[27] Cosgrove, D. J. (2018b). Nanoscale structure, mechanics and growth of epidermal cell walls. *Current Opinion in Plant Biology*, 46:77–86.

[28] Cosgrove, D. J. and Jarvis, M. C. (2012). Comparative structure and biomechanics of plant primary and secondary cell walls. *Frontiers in Plant Science*, 3.

[29] Denef, J.-F., Cordier, A. C., Mesquita, M., and Haumont, S. (1979). The influence of fixation procedure, embedding medium and section thickness on morphometric data in thyroid gland. *Histochemistry and Cell Biology*, 63(2):163–171.

[30] Dinand, E., Vignon, M., Chanzy, H., and Heux, L. (2002). Mercerization of primary wall cellulose and its implication for the conversion of cellulose I to cellulose II. *Cellulose*, 9(1):7–18.

[31] Donaldson, L. A. and Knox, J. P. (2012). Localization of Cell Wall Polysaccharides in Normal and Compression Wood of Radiata Pine: Relationships with Lignification and Microfibril Orientation. *Plant Physiology*, 158(2):642–653.

[32] Donza, M. R., Monteiro, P. H., Peixoto, R. d. N., Ferreira, N. S., Leal, D. L., Santos, L. S., Arruda, M. S., Silva, M. N., Arruda, A. C., and Guilhon, G. (2008). Chemical study of *Margaritaria nobilis* (Phyllanthaceae) leaves. *Sociedade Brasileira de Quimica*, 34.

[33] Doucet, S. M. and Meadows, M. G. (2009). Iridescence: a functional perspective. *Journal of The Royal Society Interface*, 6(Suppl_2):S115–S132.

[34] Dumanli, A. G., van der Kooij, H. M., Kamita, G., Reisner, E., Baumberg, J. J., Steiner, U., and Vignolini, S. (2014). Digital Color in Cellulose Nanocrystal Films. *ACS Applied Materials & Interfaces*, 6(15):12302–12306.

[35] Elazzouzi-Hafraoui, S., Nishiyama, Y., Putaux, J.-L., Heux, L., Dubreuil, F., and Rochas, C. (2007). The shape and size distribution of crystalline nanoparticles prepared by acid hydrolysis of native cellulose. *Biomacromolecules*, 9(1):57–65.

[36] Falcoz-Vigne, L., Ogawa, Y., Molina-Boisseau, S., Nishiyama, Y., Meyer, V., Petit-Conil, M., Mazeau, K., and Heux, L. (2017). Quantification of a tightly adsorbed monolayer of xylan on cellulose surface. *Cellulose*, 24(9):3725–3739.

[37] Fauré, R., Courtin, C. M., Delcour, J. A., Dumon, C., Faulds, C. B., Fincher, G. B., Fort, S., Fry, S. C., Halila, S., and Kabel, M. A. (2009). A brief and informationally rich naming system for oligosaccharide motifs of heteroxylans found in plant cell walls. *Australian Journal of Chemistry*, 62(6):533–537.

[38] Fernandes, A. N., Thomas, L. H., Altaner, C. M., Callow, P., Forsyth, V. T., Apperley, D. C., Kennedy, C. J., and Jarvis, M. C. (2011). Nanostructure of cellulose microfibrils in spruce wood. *Proceedings of the National Academy of Sciences*.

[39] Flauzino Neto, W. P., Putaux, J.-L., Mariano, M., Ogawa, Y., Otaguro, H., Pasquini, D., and Dufresne, A. (2016). Comprehensive morphological and structural investigation of cellulose I and II nanocrystals prepared by sulphuric acid hydrolysis. *RSC Advances*, 6(79):76017–76027.

[40] French, A. D. (2014). Idealized powder diffraction patterns for cellulose polymorphs. *Cellulose*, 21(2):885–896.

[41] French, A. D. (2017). Glucose, not cellobiose, is the repeating unit of cellulose and why that is important. *Cellulose*, 24(11):4605–4609.

[42] Frka-Petesic, B., Kamita, G., Guidetti, G., and Vignolini, S. (2018). The angular optical response of cellulose nanocrystal films explained by the structural distortions of the arrested suspension upon drying. *In preparation.*

[43] Gardens, K. (2019). World Check List of Selected Plant Families.

[44] Gierer, J. (1986). Chemistry of delignification. *Wood Science and Technology*, 20(1):1–33.

[45] Gitzendanner, M. A., Soltis, P. S., Wong, G. K.-S., Ruhfel, B. R., and Soltis, D. E. (2018). Plastid phylogenomic analysis of green plants: a billion years of evolutionary history. *American Journal of Botany*, 105(3):291–301.

[46] Gonçalves, V. M., Evtuguin, D. V., and Domingues, M. R. M. (2008). Structural characterization of the acetylated heteroxylan from the natural hybrid Paulownia elongata/Paulownia fortunei. *Carbohydrate Research*, 343(2):256–266.

[47] Goren, A., Ashlock, D., and Tetlow, I. J. (2018). Starch formation inside plastids of higher plants. *Protoplasma*, 255(6):1855–1876.

[48] Goubet, F., Jackson, P., Deery, M. J., and Dupree, P. (2002). Polysaccharide Analysis Using Carbohydrate Gel Electrophoresis: A Method to Study Plant Cell Wall Polysaccharides and Polysaccharide Hydrolases. *Analytical Biochemistry*, 300(1):53–68.

[49] Gould, K. S. and Lee, D. W. (1996). Physical and ultrastructural basis of blue leaf iridescence in four Malaysian understory plants. *American Journal of Botany*, 83(1):45–50.

[50] Graham, R. M., Lee, D. W., and Norstog, K. (1993). Physical and Ultrastructural Basis of Blue Leaf Iridescence in Two Neotropical Ferns. *American Journal of Botany*, 80(2):198.

[51] Habibi, Y., Heux, L., Mahrouz, M., and Vignon, M. R. (2008). Morphological and structural study of seed pericarp of Opuntia ficus-indica prickly pear fruits. *Carbohydrate Polymers*, 72(1):102–112.

[52] Hames, B., Ruiz, R., Scarlata, C., Sluiter, A., Sluiter, J., and Templeton, D. (2008a). Preparation of samples for compositional analysis. *Laboratory Analytical Procedure (LAP)*, 1617.

[53] Hames, B., Scarlata, C., and Sluiter, A. (2008b). Determination of protein content in biomass. *National Renewable Energy Laboratory*, pages 1–5.

[54] Herth, W. (1983). Arrays of plasma-membrane "rosettes" involved in cellulose microfibril formation of Spirogyra. *Planta*, 159(4):347–356.

[55] Hovenkamp, P., Ravensberg, W. J., and Hennipman, E. (1986). *A monograph of the fern genus pyrrosia (Polypodiaceae)*. Brill : Leiden University Press, Leiden. OCLC: 14211052.

[56] Imai, T., Boisset, C., Samejima, M., Igarashi, K., and Sugiyama, J. (1998). Unidirectional processive action of cellobiohydrolase Cel7a on Valonia cellulose microcrystals. *FEBS letters*, 432(3):113–116.

[57] infoncbi.nlm.nih.gov (2019). NCBI Taxonomy Browser.

[58] Jacobs, M., Lopez-Garcia, M., Phrathep, O.-P., Lawson, T., Oulton, R., and Whitney, H. M. (2016). Photonic multilayer structure of Begonia chloroplasts enhances photosynthetic efficiency. *Nature Plants*, 2(11).

[59] Jarvis, M. C. (2018). Structure of native cellulose microfibrils, the starting point for nanocellulose manufacture. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2112):20170045.

[60] Johansen, V. E., Onelli, O. D., Steiner, L. M., and Vignolini, S. (2017). Photonics in Nature: From Order to Disorder. In *Functional Surfaces in Biology III*, volume 10, pages 53–89. Springer International Publishing, Cham.

[61] Jordan, D., Bowman, M., Braker, J., Dien, B., Hector, R., Lee, C., Mertens, J., and Wagschal, K. (2012). Plant cell walls to ethanol. *Biochemical Journal*, 442(2):241–252.

[62] Kang, X., Kirui, A., Widanage, M. C. D., Mentink-Vigier, F., Cosgrove, D. J., and Wang, T. (2019). Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nature communications*, 10(1):347.

[63] Karlen, S. D., Smith, R. A., Kim, H., Padmakshan, D., Bartuce, A., Mobley, J. K., Free, H. C., Smith, B. G., Harris, P. J., and Ralph, J. (2017). Highly Decorated Lignins in Leaf Tissues of the Canary Island Date Palm *Phoenix canariensis*. *Plant Physiology*, 175(3):1058–1067.

[64] Kaul, R. B. (1977). The role of the multiple epidermis in foliar succulence of Peperomia (Piperaceae). *Botanical Gazette*, 138(2):213–218.

[65] Kück, U. and Wolff, G. (2014). Das Blatt. In *Botanisches Grundpraktikum*, pages 91–116. Springer.

[66] Keegstra, K. (2010). Plant Cell Walls. *Plant Physiology*, 154(2):483–486.

[67] Kinoshita, S. (2008). *Structural colors in the realm of nature*. World Scientific.

[68] Kinoshita, S. (2016). *Bionanophotonics: an introductory textbook*. CRC Press.

[69] Klug, H. P. and Alexander, L. E. (1974). *X-ray diffraction procedures: for polycrystalline and amorphous materials*. Wiley-VCH, 2 edition.

[70] Kreier, H.-P., Zhang, X.-C., Muth, H., and Schneider, H. (2008). The microsoroid ferns: Inferring the relationships of a highly diverse lineage of Paleotropical epiphytic ferns (Polypodiaceae, Polypodiopsida). *Molecular Phylogenetics and Evolution*, 48(3):1155–1167.

[71] Krogh, K. B. R. M., Harris, P. V., Olsen, C. L., Johansen, K. S., Hojer-Pedersen, J., Borjesson, J., and Olsson, L. (2010). Characterization and kinetic analysis of a thermostable GH3 β-glucosidase from Penicillium brasilianum. *Applied Microbiology and Biotechnology*, 86(1):143–154.

[72] Lahiji, R. R., Xu, X., Reifenberger, R., Raman, A., Rudie, A., and Moon, R. J. (2010). Atomic force microscopy characterization of cellulose nanocrystals. *Langmuir*, 26(6):4480–4488.

[73] Lampugnani, E. R., Khan, G. A., Somssich, M., and Persson, S. (2018). Building a plant cell wall at a glance. *Journal of Cell Science*, 131(2):jcs207373.

[74] Larsson, P. T., Hult, E.-L., Wickholm, K., Pettersson, E., and Iversen, T. (1999). CP/MAS 13c-NMR spectroscopy applied to structure and interaction studies on cellulose I. *Solid state nuclear magnetic resonance*, 15(1):31–40.

[75] Larsson, P. T., Wickholm, K., and Iversen, T. (1997). A cp/mas ^{13}c nmr investigation of molecular ordering in celluloses. *Carbohydrate Research*, 302(1-2):19–25.

[76] Lee, D. W. (1991). Ultrastructural basis and function of iridescent blue colour of fruits in Elaeocarpus. *Letters to Nature*, 349:260–261.

[77] Lee, D. W. (2007). *Nature's palette the science of plant color*. University of Chicago Press, Chicago. OCLC: 1045535906.

[78] Li, S., Bashline, L., Zheng, Y., Xin, X., Huang, S., Kong, Z., Kim, S. H., Cosgrove, D. J., and 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*, 113(40):11348–11353.

[79] Lloyd, J. R. and Kossmann, J. (2015). Transitory and storage starch metabolism: two sides of the same coin? *Current opinion in biotechnology*, 32:143–148.

[80] López, C. A., Rzepiela, A. J., de Vries, A. H., Dijkhuizen, L., Hünenberger, P. H., and Marrink, S. J. (2009). Martini Coarse-Grained Force Field: Extension to Carbohydrates. *Journal of Chemical Theory and Computation*, 5(12):3195–3210.

[81] Lyczakowski, J. J., Wicher, K. B., Terrett, O. M., Faria-Blanc, N., Yu, X., Brown, D., Krogh, K. B. R. M., Dupree, P., and Busse-Wicher, M. (2017). Removal of glucuronic acid from xylan is a strategy to improve the conversion of plant biomass to sugars for bioenergy. *Biotechnology for Biofuels*, 10(1).

[82] Martínez-Abad, A., Giummarella, N., Lawoko, M., and Vilaplana, F. (2018). Differences in extractability under subcritical water reveal interconnected hemicellulose and lignin recalcitrance in birch hardwoods. *Green Chemistry*, 20(11):2534–2546.

[83] Maslen, S. L., Goubet, F., Adam, A., Dupree, P., and Stephens, E. (2007). Structure elucidation of arabinoxylan isomers by normal phase HPLC–MALDI-TOF/TOF-MS/MS. *Carbohydrate Research*, 342(5):724–735.

[84] Meents, M. J., Watanabe, Y., and Samuels, A. L. (2018). The cell biology of secondary cell wall biosynthesis. *Annals of Botany*, 121(6):1107–1125.

[85] Middleton, R., Steiner, U., and Vignolini, S. (2017). Chapter 17 Bio-mimetic Structural Colour using Biopolymers. In *Bio-inspired Polymers*, pages 555–585. The Royal Society of Chemistry. 10.1039/9781782626664-00555.

[86] Mohnen, D. (2008). Pectin structure and biosynthesis. *Current opinion in plant biology*, 11(3):266–277.

[87] Mollenhauer, H. H. (1993). Artifacts caused by dehydration and epoxy embedding in transmission electron microscopy. *Microscopy research and technique*, 26(6):496–512.

[88] Mortimer, J. C., Miles, G. P., Brown, D. M., Zhang, Z., Segura, M. P., Weimar, T., Yu, X., Seffen, K. A., Stephens, E., Turner, S. R., and Dupree, P. (2010). Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. *Proceedings of the National Academy of Sciences*, 107(40):17409–17414.

[89] Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J., and Mansfield, S. D. (2016). Designer lignins: harnessing the plasticity of lignification. *Current Opinion in Biotechnology*, 37:190–200.

[90] Moyroud, E., Wenzel, T., Middleton, R., Rudall, P. J., Banks, H., Reed, A., Mellers, G., Killoran, P., Westwood, M. M., Steiner, U., Vignolini, S., and Glover, B. J. (2017). Disorder in convergent floral nanostructures enhances signalling to bees. *Nature*, 550(7677):469–474.

[91] Neville, A. C. (1985). Molecular and mechanical aspects of helicoid development in plant cell walls. *BioEssays*, 3(1):4–8.

[92] Neville, A. C. and Levy, S. (1985). The helicoidal concept in plant cell wall ultrastructure and morphogenesis. In *Biochemistry of plant cell walls*, number 28 in Society for Experimental Biology. Brett C.T., Hillmann J.R.

[93] Newman, R. H. (2004). Carbon-13 NMR evidence for cocrystallization of cellulose as a mechanism for hornification of bleached kraft pulp. *Cellulose*, 11(1):45–52.

[94] Ni, Y., Kubes, G. J., and Van Heiningen, A. R. P. (1993). Mechanism of chlorate formation during bleaching of kraft pulp with chlorine dioxide. *Journal of pulp and paper science*, 19(1).

[95] Nieduszynski, I. A. and Marchessault, R. H. (1972). Structure of β-D-(1$\rightarrow$4')-xylan hydrate. *Biopolymers: Original Research on Biomolecules*, 11(7):1335–1344.

[96] Niimura, H., Yokoyama, T., Kimura, S., Matsumoto, Y., and Kuga, S. (2010). AFM observation of ultrathin microfibrils in fruit tissues. *Cellulose*, 17(1):13–18.

[97] Nishiyama, Y., Kim, U.-J., Kim, D.-Y., Katsumata, K. S., May, R. P., and Langan, P. (2003). Periodic Disorder along Ramie Cellulose Microfibrils. *Biomacromolecules*, 4(4):1013–1017.

[98] Nishiyama, Y., Langan, P., and Chanzy, H. (2002). Crystal Structure and Hydrogen-Bonding System in Cellulose Iβ from Synchrotron X-ray and Neutron Fiber Diffraction. *Journal of the American Chemical Society*, 124(31):9074–9082.

[99] Nishiyama, Y., Langan, P., O'Neill, H., Pingali, S. V., and Harton, S. (2014). Structural coarsening of aspen wood by hydrothermal pretreatment monitored by small- and wide-angle scattering of X-rays and neutrons on oriented specimens. *Cellulose*, 21(2):1015–1024.

[100] 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., Roberts, A. W., Yingling, Y. G., and Haigler, C. H. (2016). Comparative Structural and Computational Analysis Supports Eighteen Cellulose Synthases in the Plant Cellulose Synthesis Complex. *Scientific Reports*, 6(1).

[101] NREL (2019). NREL Biomass Compositional Analysis Laboratory Procedures. https://www.nrel.gov/bioenergy/biomass-compositional-analysis.html.

[102] Oehme, D. P., Downton, M. T., Doblin, M. S., Wagner, J., Gidley, M. J., and Bacic, A. (2015). Unique Aspects of the Structure and Dynamics of Elementary I β Cellulose Microfibrils Revealed by Computational Simulations. *Plant Physiology*, 168(1):3–17.

[103] Ogawa, Y. and Putaux, J.-L. (2019). Transmission electron microscopy of cellulose. Part 2: technical and practical aspects. *Cellulose*, 26(1):17–34.

[104] Paredez, A. R., Somerville, C. R., and Ehrhardt, D. W. (2006). Visualization of cellulose synthase demonstrates functional association with microtubules. *Science*, 312(5779):1491–1495.

[105] Park, S., Baker, J. O., Himmel, M. E., Parilla, P. A., and Johnson, D. K. (2010). Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnology for biofuels*, 3(1):10.

[106] Parker, A. R. (2000). 515 million years of structural colour. *Journal of Optics A: Pure and Applied Optics*, 2(6):R15–R28.

[107] Pell, G., Taylor, E. J., Gloster, T. M., Turkenburg, J. P., Fontes, C. M. G. A., Ferreira, L. M. A., Nagy, T., Clark, S. J., Davies, G. J., and Gilbert, H. J. (2004). The Mechanisms by Which Family 10 Glycoside Hydrolases Bind Decorated Substrates. *Journal of Biological Chemistry*, 279(10):9597–9605.

[108] Petchsri, S., Boonkerd, T., and Baum, B. R. (2009). A first record of *Microsorum musifolium* Copel.(Polypodiaceae) from Thailand. *Tropical Natural History*, 9(1):99–104.

[109] Petchsri, S., Boonkerd, T., Baum, B. R., Karladee, D., Suriyong, S., and Lungkaphin, A. (2012). Phenetic study of the *Microsorum punctatum* complex (Polypodiaceae). *ScienceAsia*, 38(1):1–12.

[110] Pfister, B. and Zeeman, S. C. (2016). Formation of starch in plant cells. *Cellular and Molecular Life Sciences*, 73(14):2781–2807.

[111] Rasband, W. (1997). ImageJ. http://imagej.net/Welcome.

[112] Reis, D. and Vian, B. (2004). Helicoidal pattern in secondary cell walls and possible role of xylans in their construction. *Comptes Rendus Biologies, Academie des sciences, Elsevier SAS*, 327(9-10):785–790.

[113] Reis, D., Vian, B., Chanzy, H., and Roland, J.-C. (1991). Liquid crystal-type assembly of native cellulose-glucuronoxylans extracted from plant cell wall. *Biology of the Cell*, 73(2-3):173–178.

[114] Rogowski, A., Baslé, A., Farinas, C. S., Solovyova, A., Mortimer, J. C., Dupree, P., Gilbert, H. J., and Bolam, D. N. (2014). Evidence That GH115 α-Glucuronidase Activity, Which Is Required to Degrade Plant Biomass, Is Dependent on Conformational Flexibility. *Journal of Biological Chemistry*, 289(1):53–64.

[115] Roland, J. C., Reis, D., Vian, B., Satiat-Jeunemaitre, B., and Mosiniak, M. (1987). Morphogenesis of plant cell walls at the supramolecular level: internal geometry and versatility of helicoidal expression. *Protoplasma*, 140(2-3):75–91.

[116] Saake, B., Kruse, T., and Puls, J. (2001). Investigation on molar mass, solubility and enzymatic fragmentation of xylans by multi-detected SEC chromatography. *Bioresource technology*, 80(3):195–204.

[117] Saffer, A. M. (2018). Expanding roles for pectins in plant development. *Journal of integrative plant biology*, 60(10):910–923.

[118] Scheller, H. V. and Ulvskov, P. (2010). Hemicelluloses. *Annual Review of Plant Biology*, 61(1):263–289.

[119] Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682.

[120] Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods*, 9(7):671.

[121] Scott, K. and Willis, K. (2016). *Botanicum*. Big Picture Press, UK, limited edition hardback edition.

[122] Simmons, T. J., Mortimer, J. C., Bernardinelli, O. D., Pöppler, A.-C., Brown, S. P., deAzevedo, E. R., Dupree, R., and Dupree, P. (2016). Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nature Communications*, 7:13902.

[123] Sluiter, A., Hames, B., Hyman, D., Payne, C., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., and Wolfe, J. (2008a). Determination of total solids in biomass and total dissolved solids in liquid process samples. *National Renewable Energy Laboratory, Golden, CO, NREL Technical Report No. NREL/TP-510-42621*, pages 1–6.

[124] Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., and Templeton, D. (2005). Determination of Ash in Biomass. *National Renewable Energy Laboratory*.

[125] Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., and Crocker, D. (2008b). Determination of structural carbohydrates and lignin in biomass. *Laboratory analytical procedure*, 1617:1–16.

[126] Sluiter, A., Ruiz, R., Scarlata, C., Sluiter, J., and Templeton, D. (2008c). Determination of extractives in biomass. *National Renewable Energy Laboratory*.

[127] Sluiter, A. and Sluiter, J. (2008). Determination of starch in solid biomass samples by HPLC. *Colorado: NREL*.

[128] Sluiter, J. and Sluiter, A. (2011). Summative mass closure. *National Renewable Energy Laboratory, Golden*.

[129] Smith, A. R. and Hoshizaki, B. J. (2000). *Microsorum whiteheadii* (Polypodiaceae), an Attractive New Species from Sumatra. *Novon*, 10(4):411.

[130] Steiner, L. M., Ogawa, Y., Johansen, V. E., Lundquist, C. R., Whitney, H., and Vignolini, S. (2018). Structural colours in the frond of Microsorum thailandicum. *Journal of the Royal Society Interface Focus*, 9(1):20180055.

[131] Strout, G., Russell, S. D., Pulsifer, D. P., Erten, S., Lakhtakia, A., and Lee, D. W. (2013). Silica nanoparticles aid in structural leaf coloration in the Malaysian tropical rainforest understorey herb Mapania caudata. *Annals of Botany*, 112(6):1141–1148.

[132] Sultanova, N., Kasarova, S., and Nikolov, I. (2009). Dispersion properties of optical polymers. *Acta Physica Polonica-Series A General Physics*, 116(4):585.

[133] Tanaka, R., Kuribayashi, T., Ogawa, Y., Saito, T., Isogai, A., and Nishiyama, Y. (2017). Ensemble evaluation of polydisperse nanocellulose dimensions: rheology, electron microscopy, X-ray scattering and turbidimetry. *Cellulose*, 24(8):3231–3242.

[134] Terrett, O. M. and Dupree, P. (2019). Covalent interactions between lignin and hemicelluloses in plant secondary cell walls. *Current opinion in biotechnology*, 56:97–104.

[135] Thomas, K. R., Kolle, M., Whitney, H. M., Glover, B. J., and Steiner, U. (2010). Function of blue iridescence in tropical understorey plants. *Journal of The Royal Society Interface*, 7(53):1699–1707.

[136] Thomas, L. H., Forsyth, V. T., Martel, A., Grillo, I., Altaner, C. M., and Jarvis, M. C. (2014). Structure and spacing of cellulose microfibrils in woody cell walls of dicots. *Cellulose*, 21(6):3887–3895.

[137] Thomas, L. H., Forsyth, V. T., Martel, A., Grillo, I., Altaner, C. M., and 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(1).

[138] Tomlinson, P. B., Horn, J. W., and Fisher, J. B. (2011). *The anatomy of palms: Arecaceae-Palmae*. Oxford University Press, Oxford, UK, 1 edition.

[139] Urbániková, u., Vršanská, M., Mørkeberg Krogh, K. B. R., Hoff, T., and Biely, P. (2011). Structural basis for substrate recognition by Erwinia chrysanthemi GH30 glucuronoxylanase: X-ray structure of xylanase A-ligand complex. *FEBS Journal*, 278(12):2105–2116.

[140] Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin Biosynthesis and Structure. *Plant Physiology*, 153(3):895–905.

[141] Vardakou, M., Dumon, C., Murray, J. W., Christakopoulos, P., Weiner, D. P., Juge, N., Lewis, R. J., Gilbert, H. J., and Flint, J. E. (2008). Understanding the Structural Basis for Substrate and Inhibitor Recognition in Eukaryotic GH11 Xylanases. *Journal of Molecular Biology*, 375(5):1293–1305.

[142] Vaughan, T. (2017a). Icytree. https://icytree.org/. Accessed 2019-04.

[143] Vaughan, T. G. (2017b). IcyTree: rapid browser-based visualization for phylogenetic trees and networks. *Bioinformatics*, 33(15):2392–2394.

[144] Vian, B., Reis, D., Darzens, D., and Roland, J. C. (1994). Cholesteric-like crystal analogs in glucuronoxylan-rich cell wall composites: experimental approach of acellular re-assembly from native cellulosic suspension. *Protoplasma*, 180(1-2):70–81.

[145] Vignolini, S., Gregory, T., Kolle, M., Lethbridge, A., Moyroud, E., Steiner, U., Glover, B. J., Vukusic, P., and Rudall, P. J. (2016). Structural colour from helicoidal cell-wall architecture in fruits of *Margaritaria nobilis*. *Journal of The Royal Society Interface*, 13(124):20160645.

[146] Vignolini, S., Moyroud, E., Glover, B. J., and Steiner, U. (2013). Analysing photonic structures in plants. *Journal of The Royal Society Interface*, 10(87):20130394–20130394.

[147] Vignolini, S., Rudall, P. J., Rowland, A. V., Reed, A., Moyroud, E., Faden, R. B., Baumberg, J. J., Glover, B. J., and Steiner, U. (2012). Pointillist structural color in Pollia fruit. *Proceedings of the National Academy of Sciences*, 109(39):15712–15715.

[148] Voiniciuc, C., Pauly, M., and Usadel, B. (2018). Monitoring Polysaccharide Dynamics in the Plant Cell Wall. *Plant Physiology*, 176(4):2590–2600.

[149] Vukusic, P. and Sambles, J. R. (2003). Photonic structures in biology. *Nature*, 424(6950):852.

[150] Vukusic, P. and Stavenga, D. (2009). Physical methods for investigating structural colours in biological systems. *Journal of The Royal Society Interface*, 6(Suppl_2):S133–S148.

[151] Wang, T., Park, Y. B., Cosgrove, D. J., and 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(3):871–884.

[152] Wang, T., Yang, H., Kubicki, J. D., and 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(6):2210–2222.

[153] Whitney, H. M., Kolle, M., Andrew, P., Chittka, L., Steiner, U., and Glover, B. J. (2009). Floral Iridescence, Produced by Diffractive Optics, Acts As a Cue for Animal Pollinators. *Science*, 323(5910):130–133.

[154] Wickett, N. J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., Ayyampalayam, S., Barker, M. S., Burleigh, J. G., and Gitzendanner, M. A. (2014). Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences*, 111(45):E4859–E4868.

[155] Wickholm, K., Larsson, P. T., and Iversen, T. (1998). Assignment of noncrystalline forms in cellulose I by NMR. *Carbohydrate Research*, 312:123–129.

[156] Wilkens, C., Andersen, S., Dumon, C., Berrin, J.-G., and Svensson, B. (2017). GH62 arabinofuranosidases: Structure, function and applications. *Biotechnology Advances*, 35(6):792–804.

[157] Wilts, B. D., Dumanli, A. G., Middleton, R., Vukusic, P., and Vignolini, S. (2017). Invited Article: Chiral optics of helicoidal cellulose nanocrystal films. *APL Photonics*, 2(4):0408011–0408017.

[158] Wilts, B. D., Whitney, H. M., Glover, B. J., Steiner, U., and Vignolini, S. (2014). Natural Helicoidal Structures: Morphology, Self-assembly and Optical Properties. *Materials Today: Proceedings*, 1:177–185.

[159] Wohlert, J. and Berglund, L. A. (2011). A Coarse-Grained Model for Molecular Dynamics Simulations of Native Cellulose. *Journal of Chemical Theory and Computation*, 7(3):753–760.

[160] Wolf, F. T. (1958). Chlorophylls A and B in the Pteridophytes. *Bulletin of the Torrey Botanical Club*, 85(1):1.

[161] Zeeman, S. C., Smith, S. M., and Smith, A. M. (2007). The diurnal metabolism of leaf starch. *Biochemical Journal*, 401(1):13–28.

[162] Zhang, T., Zheng, Y., and Cosgrove, D. J. (2016). Spatial organization of cellulose microfibrils and matrix polysaccharides in primary plant cell walls as imaged by multichannel atomic force microscopy. *The Plant Journal*, 85(2):179–192.

Appendix A

Appendix

A.1 Structural colour from helicoidal architectures in plants

There are different ways to construct phylogenetic trees: one way is based on plastid phylogenomic analysis, as discussed in chapter 1. Figure A.1 depicts a phylogenetic tree of plants based on phylogenetic analysis via nuclear genes by Wickett et al. [154]. *Microsorum thailandicum* (discussed in detail in chapter 2 and chapter 3), *Lindsaea lucida*, *Diplazium tomentosum*, and *Danaea nodosa* are all found in the clade of monilophytes. Next, both *Pollia condensata* and *Mapania caudata* belong to the clade of monocots, while *Margaritaria nobilis* belongs to the clade of eudicots.

A.2 NREL quantifications

A.2.1 Calculations

An overview of all the reactions and measurements is given in Figure 4.1, where all the steps are explained as well. Here, all the formulae are summarised. All the calculations and values are based on the dried starting material, the so-called oven dried weight (ODW), and they are all in weight percent.

$$TS\% = \frac{sample_{dried}}{sample_{milled}} \cdot 100 \tag{A.1}$$

$$ODW\% = \frac{sample \cdot TS\%}{100} \tag{A.2}$$

First, the protein content and the ash content were determined from the starting material, as well as the total solids content. It is recommended to determine the so-called nitrogen

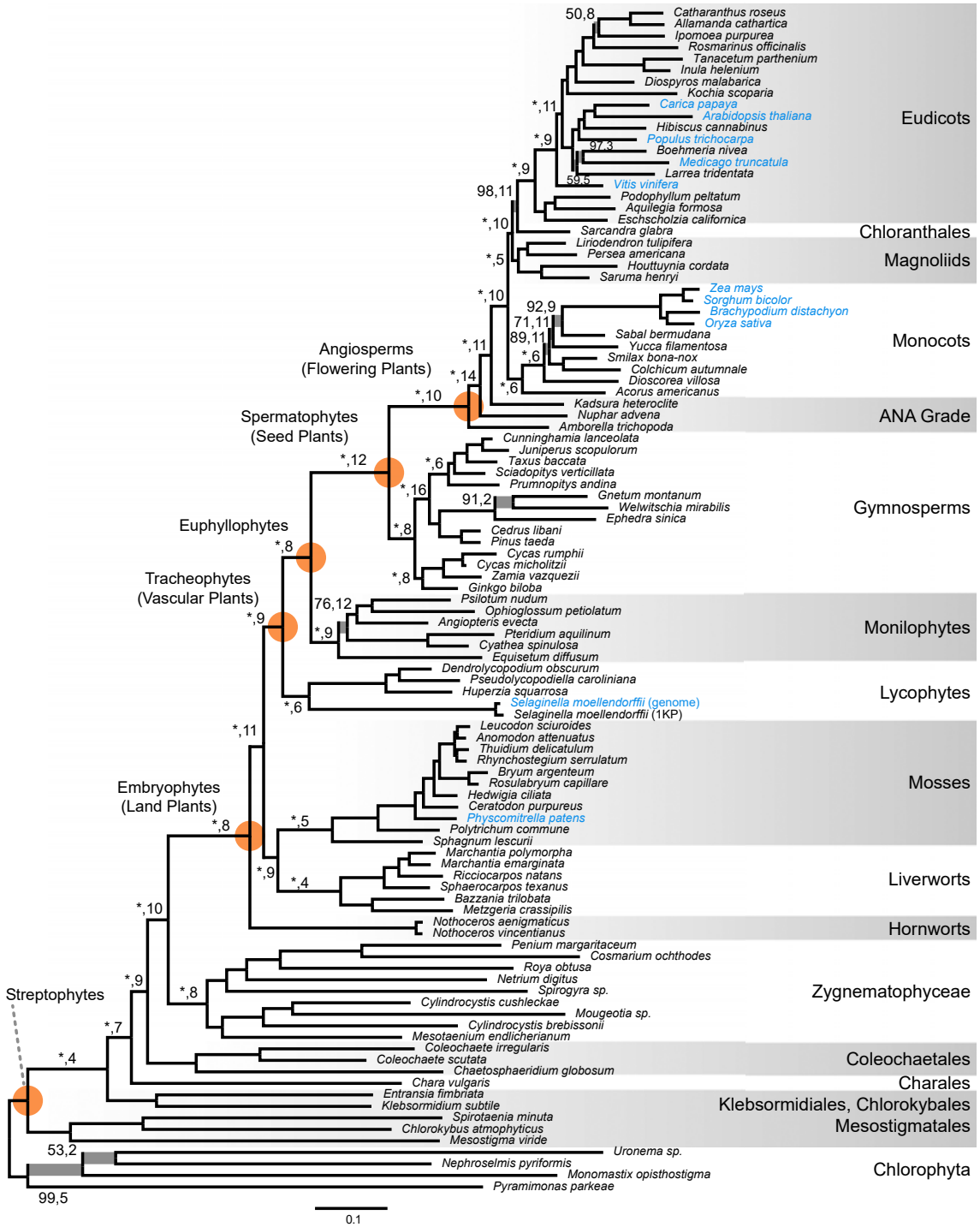


Fig. A.1 A phylogenetic tree of plants based on phylogenetic analysis via nuclear genes by Wickett et al. [154].

factor for each sample, based on the individual amino acid composition of the protein content. However, the protein content is low, only around 10 wt% in the leaves, and 1-2 % in the endocarp, so this step was omitted and the commonly used factor of 6.25 was used for simplicity. Both the nitrogen and ash content were determined by the microanalysis service technician in the department, via combustion analysis with a CHN analyser. Their contents are calculated as follows:

$$protein\% = nitrogen\%_{ODW} \cdot 6.25 \tag{A.3}$$

$$ash\% = \frac{ash}{ODW} \cdot 100 \tag{A.4}$$

The TS% obtained is used to correct any further sample weight in the following procedures. The protein% and ash% are the final results of the ODW sample, all labelled with 1 in the scheme. The protein% and ash% were also determined again later in the procedure, labelled with 2 and 3, to correct for protein and ash that was still present in the extracted solid (2), and that condensed into the filtrate, the acid insoluble lignin (AIL) (3).

Two samples were measured twice each, and the root mean square (RMS) and the root mean square standard deviation (σ) were calculated:

$$RMS = \sqrt{\frac{1}{n} \cdot \sum x_i^2} \tag{A.5}$$

$$\sigma_{RMS} = \sqrt{\frac{1}{n} \cdot \sum (x_i - RMS)^2} \tag{A.6}$$

Next, the aqueous and ethanol extractions were carried out. The sample was not previously oven-dried, but used as milled, and corrected with the TS% to continue calculations with the ODW. It is not recommended to oven-dry samples before experiments, since the high temperatures might lead to changes that could cause decomposition in the following reaction steps, but rather correct the amounts with the TS%. The respective extractives contents were calculated via:

$$extractives\% = \frac{extractives}{ODW} \cdot 100 - ((protein\%_2 - protein\%_1) + (ash\%_2 - ash\%_1)) \tag{A.7}$$

Here, the protein%(2) and ash%(2) of the extracted solid were determined, and used for correcting the extractives%: the percentage of extractives not found in the extracted solid

anymore compared to before extraction (protein%(2)-protein%(1) and ash%(2)-ash%(1)) must have been extracted into the extract, and were therefore subtracted from the extractives.

Furthermore, another TS% value of the extracted solids was determined via Equation A.1, TS%(2), to correct the amount of extracted sample used for hydrolysis in the subsequent step with Equation A.2.

Next, the extracted sample was split into two equal parts and acid hydrolysis was carried out, the RMS of the results from both runs was calculated. The lignin content was determined by summing up the acid insoluble lignin (AIL)and the acid soluble lignin (ASL). Protein%(3) and ash%(3) were used for correcting their contributions to AIL. As opposed to AIL and the extractives, ASL and the carbohydrates are not determined gravimetrically and need not be corrected for protein and ash.

$$AIL\% = \frac{AIL}{ODW} \cdot 100 - (protein\%_3 - ash\%_3) \tag{A.8}$$

$$ASL\% = \frac{UV_{abs} \cdot Vol \cdot dil}{\varepsilon \cdot (ODW \cdot 1000) \cdot L} \cdot 100 \tag{A.9}$$

Symbol		Unit
UV_{abs}	UV absorbance at 240 nm	-
Vol	volume of filtrate	ml
dil	dilution	-
ε	absorptivity at 240 nm	L / g cm
L	path length	cm

The volume of filtrate is 86.73 ml for 0.3 g of extracted sample used for hyrdolysis, or adapted accordingly. Dilution is calculated via Equation A.10, and should be chosen so that the UV absorbance at 240 nm lies between 0.7 and 1. (interference from carbohydrate degradation products should be at a minimum at 240 nm). ε was chosen as 23 L/g cm as the average value given in the NREL protocol for different types of biomass, and path length is 1 cm.

$$dil = \frac{Vol_{sample} + Vol_{milliQ}}{Vol_{sample}} \tag{A.10}$$

Finally, the content of different monosaccharides, namely glucose, xylose, galactose, L-arabinose and mannose were determined via HPLC quantifications. Furthermore, cellobiose was monitored, since residual cellobiose would mean incomplete hydrolysis, so that hydrolysis time would have to be extented, which was never observed in my samples.

Calibration curves were made for each monosaccharide (and cellobiose) and corrected with a calibration verification standard, all equations are given in Table A.1. Furthermore, for every set of two samples of acid hydrolysis, a set of monosaccharides solution was run through the second hydrolysis step (in the autoclave), to calculate the sugar recovery standard rate R_{sugar}, to correct for monosaccharide degradation during hydrolysis. $Conc_{sugar}$ is the sugar concentration in mg/ml, in the same unit as the calibration equations, as well as $\left[SRS_{after}\right]$, the concentration of SRS obtained via HPLC analysis after hydrolysis, and $\left[SRS_{before}\right]$, before hydrolysis, respectively.

$$R_{sugar} = \frac{\left[SRS_{after}\right]}{\left[SRS_{before}\right]} \cdot 100 \tag{A.11}$$

$$Conc_{sugar} = \frac{conc_{HPLC}}{R_{sugar}} \cdot 100 \tag{A.12}$$

$$Sugar\% = \frac{conc_{sugar} \cdot Vol \cdot \frac{1}{1000}}{ODW} \cdot 100 \tag{A.13}$$

Table A.1 Calibration equations and coefficients of determination of R^2 for cellobiose and all monosaccharides

Sugar	**Calibration equation**	R^2
Cellobiose	y = 126133.2106481650 · x + 0.0035923090	0.99997
Glucose	y = 121617.1584628570 · x + 0.0068769877	0.99964
Xylose	y = 123237.2807789320 · x - 0.0062065462	1.00000
Galactose	y = 111650.2328272470 · x + 0.0000628057	0.99998
L-arabinose	y = 118143.1625255570 · x - 0.0039469106	0.99996
Mannose	y = 170475.2153934030 · x - 0.0540435628	0.99759

A.2.2 Results

The results for the quantification of the components of the endocarp are summarised in Table A.2. note that monosaccharide composition was based on 3 runs rather than 4, see below...

A.2.3 HPLC elution diagrams

Figure A.2 shows an example of the HPLC elution diagrams obtained after acid hydrolysis and neutralising the sample. In (A), the original measurement data are shown in orange, and

Table A.2 Detailed results for the quantification of the components of the endocarp

	Run 1		**Run 2**		**RMS**	**STDEV**
Aqueous extractives	6.74	6.74	5.72	5.72	**6.25**	0.51
EtOH extractives	2.81	2.81	2.26	2.26	**2.55**	0.28
Protein	n.d.	n.d.	1.26	1.26	**0.89**	0.68
Ash	0.55	0.55	0.30	0.30	**0.44**	0.13
	Run 1a	**Run 1b**	**Run 2a**	**Run 2b**		
Acid insoluble lignin	19.88	24.38	26.45	28.80	**25.10**	3.29
Acid soluble lignin	6.38	2.67	5.59	3.99	**4.87**	1.45
D-glucose	29.25	35.57	36.11	35.72	**35.80**	0.22
D-xylose	15.35	23.18	22.11	22.40	**22.57**	0.45
D-galactose	1.35	1.55	0.00	0.00	**0.89**	0.82
L-arabinose	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
D-mannose	0.35	0.22	1.21	1.21	**1.00**	0.48

Table A.3 Detailed results for the quantification of the components of the leaves

	Run 1		**Run 2**		**RMS**	**STDEV**
Aqueous extractives	33.39	33.39	34.89	34.89	**34.14**	0.75
EtOH extractives	14.04	14.04	10.68	10.68	**12.47**	1.68
Protein	11.30	11.30	11.23	11.23	**11.27**	0.03
Ash	5.26	5.26	5.27	5.27	**5.26**	0.00
	Run 1a	**Run 1b**	**Run 2a**	**Run 2b**		
Acid insoluble lignin	24.97	23.13	15.82	14.76	**20.17**	4.47
Acid soluble lignin	3.44	4.92	5.51	8.05	**5.73**	1.68
D-glucose	16.32	15.47	15.30	14.84	**15.49**	0.54
D-xylose	4.95	4.54	7.60	6.95	**6.15**	1.30
D-galactose	2.35	2.72	2.99	3.08	**2.80**	0.28
L-arabinose	1.72	2.23	1.31	1.17	**1.66**	0.42
D-mannose	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

a blue line that is subtracted, which removes the baseline drift stemming from the sulphuric acid. The slope and extent were adjusted for every sample, since it varied slightly between different batches. The processed elution diagram is shown in (B), and the height of the D-glucose peak (16.05 min, orange), D-xylose (17.27 min, yellow), and D-mannose (20.81 min, green) were detected automatically in this case, D-galactose (18.52 min) was read manually, and no peak for L-arabinose was detected in this sample.

For one of the endocarp quantifications, the monosaccharide concentrations could not be determined, since the sulphuric acid baseline drift dropped in the middle of the xylose peak, and no way could be found to correct for it. I also tried to neutralise residual sample in another attempt, but it did not influence the position of the drift drop.

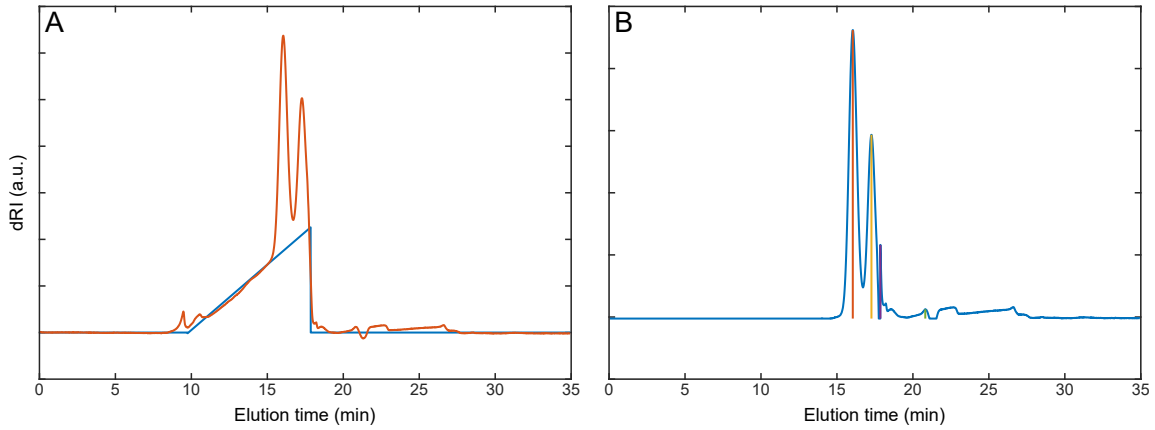


Fig. A.2 HPLC elution diagrams of NREL monosaccharide quantification. (A) Original diagram and with line to subtract baseline drift. (B) Corrected diagram with peak heights.

A.3 Cellulose isolation procedure

A.3.1 STEM SEM imaging to optimise tip-sonication energy input

STEM SEM imaging does not have as good resolution as TEM imaging, but it is a faster way to scan a larger amount of samples. In this case STEM SEM was used to monitor whether tip-sonication leads to mechanical damage of the cellulose microfibrils, i.e. by introducing kinking. Various tip-sonication times, corresponding to different amounts of energy input, were tested on small amounts of sample, and their morphology monitored.

Figure A.3 shows the endocarp sample after the second DMSO extraction and before tip-sonication, 11DE, in A, where some microfibrils were visible, but they were covered in large amounts of aggregated xylan. (The image was taken next to the TEM grid, which is

the black structure on the right.) In (B), a TEM grid was prepared by only depositing the supernatant of the 11DE, which contained more aggregated xylan. However there were also some microfibrils contained, and washing the sample after reactions is a balance between removing xylan and accidentally removing too much cellulose with the supernatant. Finally, Figure A.3 C shows the endocarp sample after tip-sonication, 12TS, with 40 kJ/g cellulose. No xylan was removed during the tip-sonication, but it seems that the sonication did help break up xylan aggregates and remove them from the microfibrils surface. Furthermore, it was confirmed that the sonication energy of 40 kJ/g cellulose was not too high since no kinks were observed on the microfibrils.

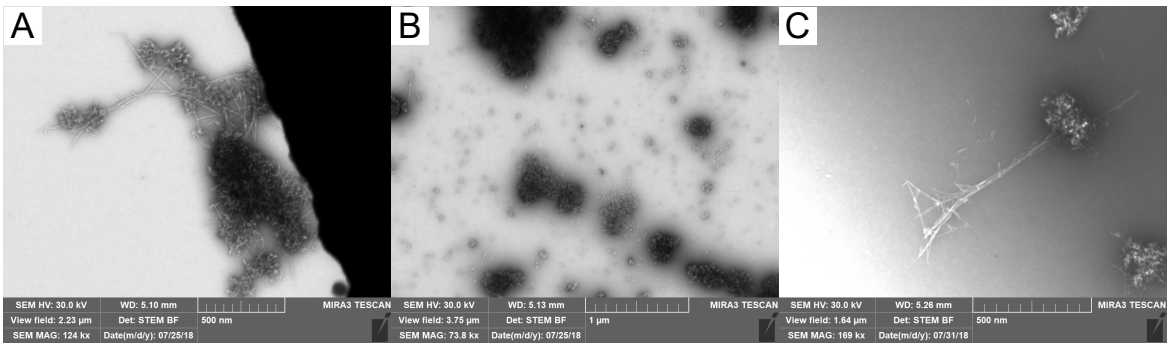


Fig. A.3 STEM SEM images monitoring tip-sonication energy input for the endocarp 12TS sample. (A) Endocarp 11DE, microfibrils covered in large amounts of aggregated xylan. (B) Endocarp 11DE supernatant, primarily containing aggregated xylan. (C) Endocarp 12TS after tip-sonication with 40 kJ/g cellulose, no kinks were introduced.

A.3.2 Xylan in NMR spectrum

As mentioned in section 5.3, regenerated xylan was found in the endocarp 13BT NMR spectrum and TEM imaging. It was already observed when optimising tip-sonication energy input, see Figure A.3. During the base treatments and DMSO extractions, it presumably got dissolved, but then it was very difficult to wash it out during centrifuge washes. If too much of the supernatant were disposed of, also cellulose microfibrils would be removed, and it would shift the length distribution to longer lengths. It did not seem that there was still a lot of xylan interacting with the cellulose, since the peak at 82 ppm, indicating xylan in two-fold helical screw conformation is very small. However peaks corresponding to regenerated xylan in three-fold helical screw conformation were observed at approximately 101, 74, and 63 ppm. An NMR spectrum of extracted birch xylan was kindly provided by Lea Falcoz-Vigne for comparison. It might be slightly different from regenerated xylan from *M. nobilis* endocarp, but it was subtracted from the endocarp 13BT spectrum in Figure 5.6

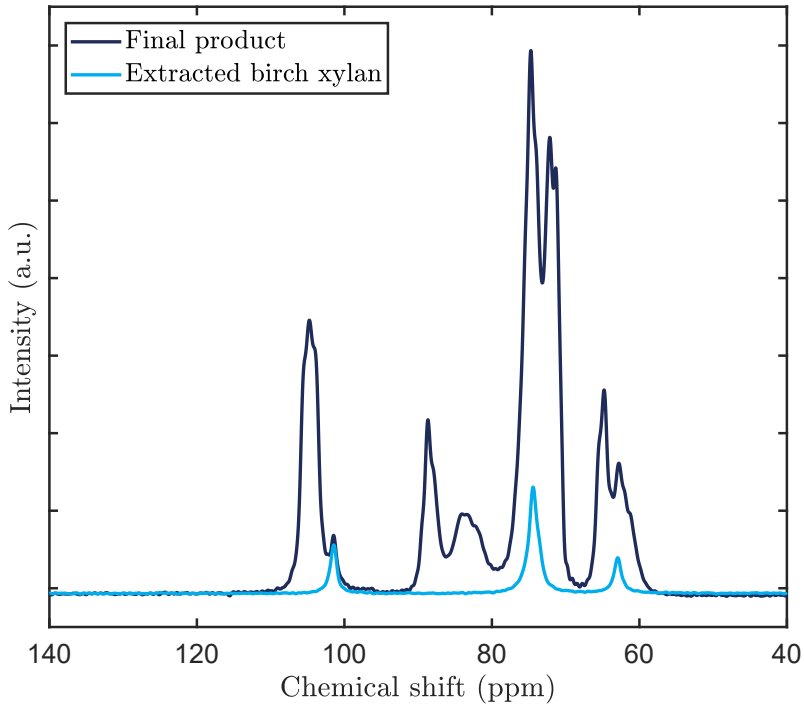


Fig. A.4 NMR spectrum of endocarp 13BT and of extracted birch xylan, provided by Lea Falcoz-Vigne [36].

top for simplicity. The original spectrum and the accordingly scaled xylan spectrum are presented in Figure A.4.

A.4 Xylan analysis

A.4.1 Enzymes

Cartoons depicting the mode of action, obtained oligo- and monosaccharide fragments and the resulting PACE gels are shown below. GH10 (already depicted in the main text in Figure 4.5), GH11 (Figure A.5), GH30 (Figure A.6) and GH3 (Figure A.7) are all (glucurono-) xylanases, meaning that they hydrolyse the xylan backbone [107, 141, 139, 71]. More specifically, GH10 and GH11 will cleave next to GlcA decorations (at the +1 or +2 position, respectively, see the cartoons for details), and also along the chain. On the other hand, GH30 will (predominantly) only cleave next to a GlcA decoration, at the -2 position. And finally, GH3 only cleaves xylose residues off the chain from the non-reducing end, but only as long as there are no decorations.

Both GH115 and GH67 are glucuronidases, they cut GlcA decorations off the main chain [114]. While GH115 will act anywhere along the chain, GH67 (Figure A.7) will only remove GlcA if it is directly at the last xlyose residue at the non-reducing end of the chain.

Furthermore, GH62 is an arabinosidase, it would cleave arabinose decorations off the xylan chain [156].

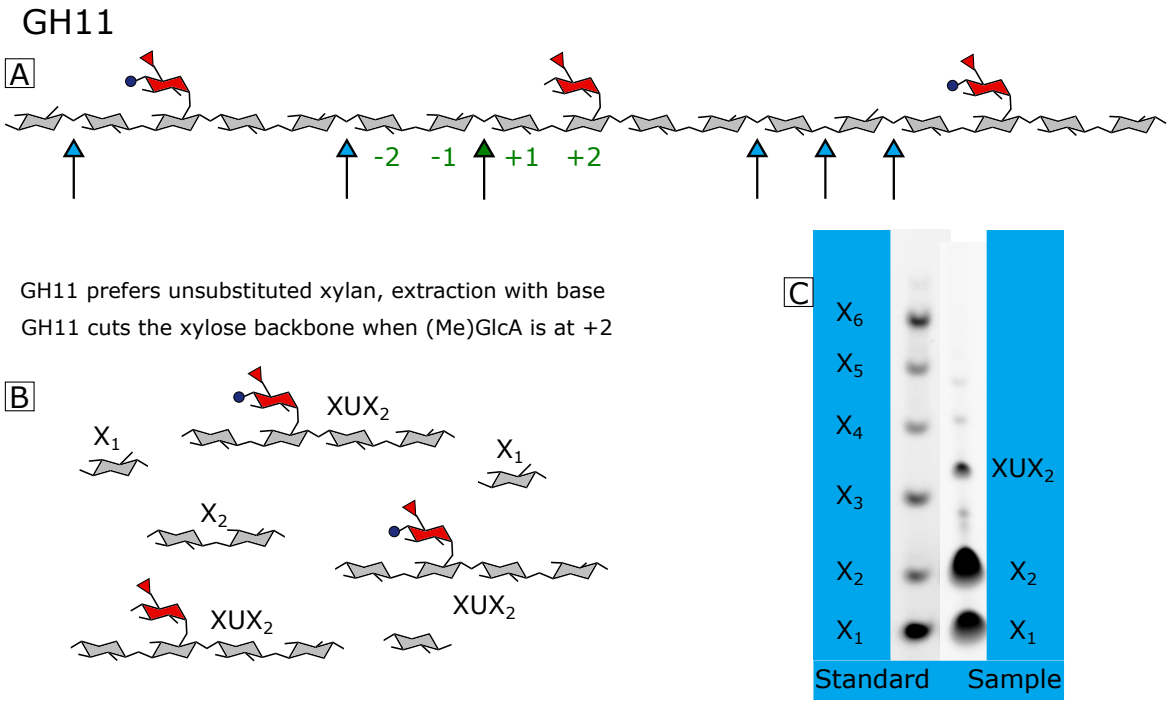


Fig. A.5 (A) Short example xylan chain. Arrows show the specific sites of cleavage for GH11. (B) Fragments obtained from digesting the short example xylan chain with GH11. (C) PACE gel showing all the characteristic fragments of GH11 digestion.

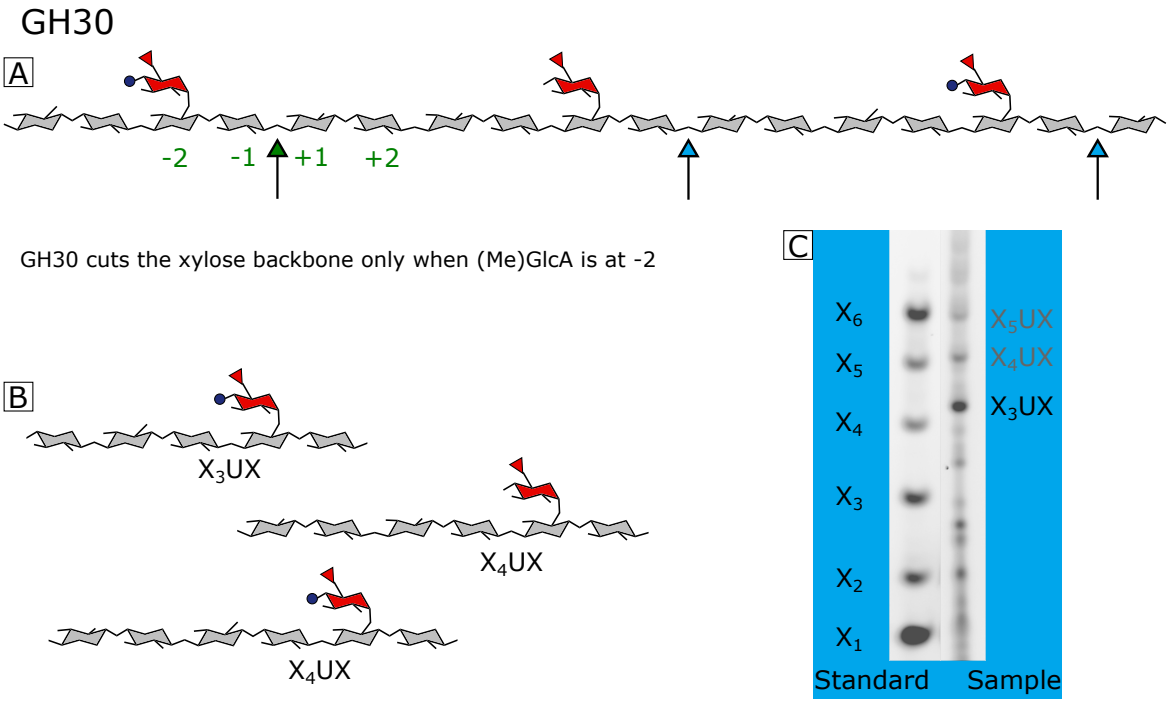


Fig. A.6 (A) Short example xylan chain. Arrows show the specific sites of cleavage for GH30. (B) Fragments obtained from digesting the short example xylan chain with GH11. (C) PACE gel showing all the characteristic fragments of GH30 digestion.

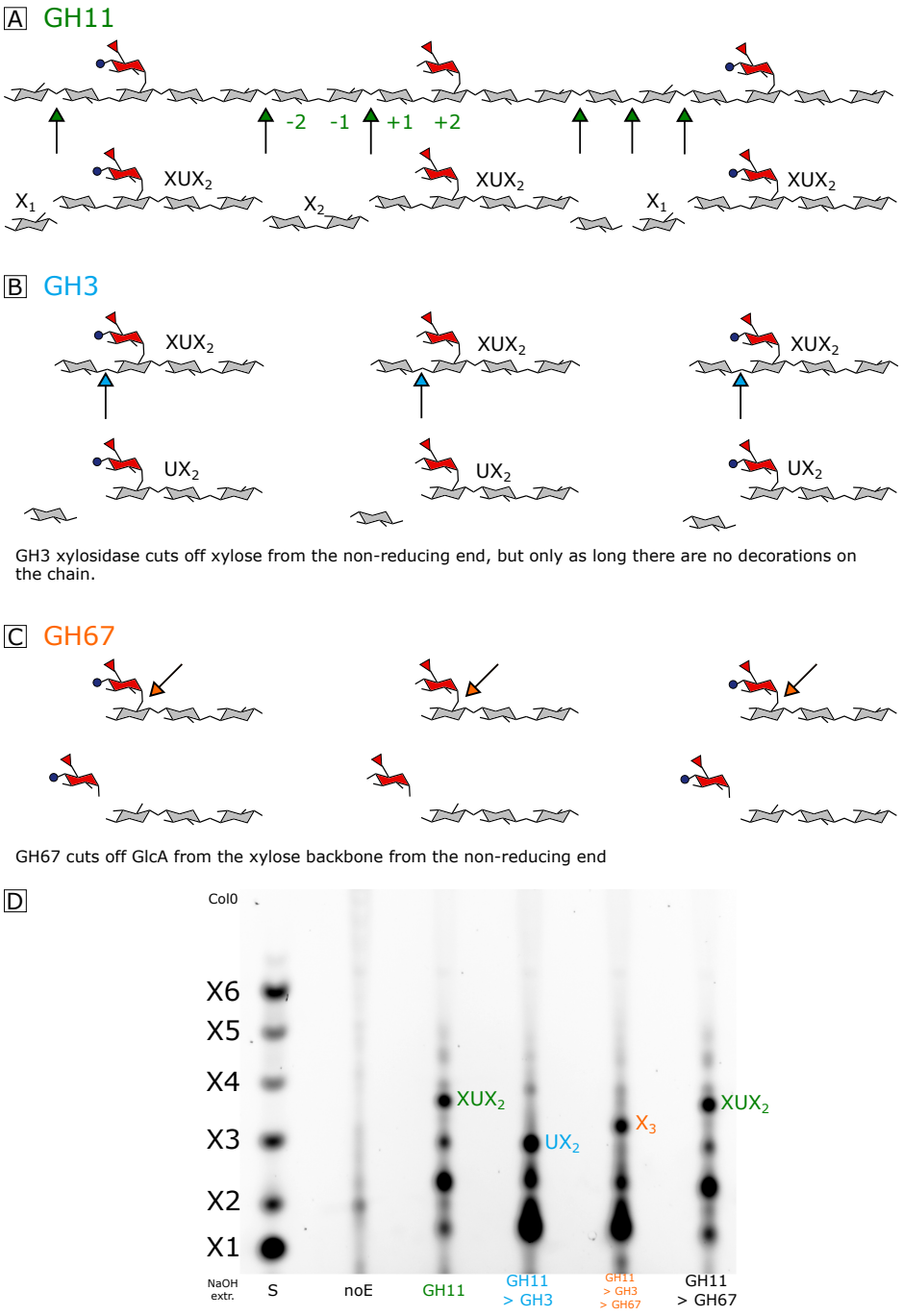


Fig. A.7 Sequential digestion with GH11, GH3 and GH67. (A) Initial digestion by GH11: Short example xylan chain. Arrows show the specific sites of cleavage for GH11. Fragments obtained from digesting the short example xylan chain with GH11. (B) Subsequent digestion by GH3: Only relevant fragments from GH11 digestion are depicted. Arrows show the specific sites of cleavage for GH3. Fragments obtained from digesting the oligosaccharides with GH3. (C) Final digestion by GH67: Oligosaccharides from GH3 digestion are depicted. Arrows show the specific sites of cleavage for GH67. Fragments obtained from digesting the oligosaccharides with GH67. (D) PACE gel showing all the characteristic fragments of the consecutive digestions. Note no GH67 digestion occurs when used straight after GH11 digestion in the last lane. S: xylan standard, noE: no enzyme control

A.4.2 Quantification of glucuronic acid content

The ratios of different fragments to each other on a PACE gel can be quantified using the Fiji plug-in in the ImageJ software [120, 119], which can be used to determine the GlcA content. The respective lane is chosen, here the GH11 digestion of deacetylated xylan (same gel as in Figure 5.17, but lower exposure time to exclude issues with oversaturation). Then, the intensity of greyscale and area are calculated, as depicted below, Figure A.8. The fragments surrounded by the yellow rectangle, obtained by GH11 digestion after NaOH extraction, are X_1, X_2 and XUX_2. The area for xylose is calculated by multiplying the respective fragment area with the respective amount of xylose atoms: $A_X = A_{X1} + 2 \cdot A_{X2} + 4 \cdot A_{X4}$, while the amount of GlcA corresponds to A_{X4}. Finally, the % GlcA of total area was obtained to be approximately 4 %.

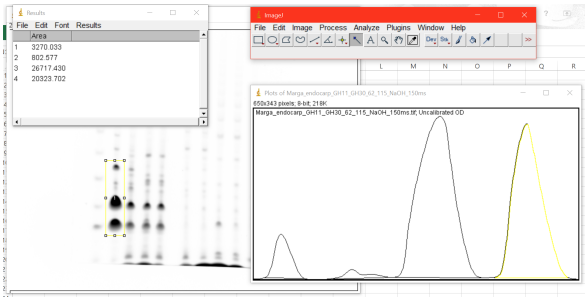


Fig. A.8 Determination of GlcA content after GH11 digestion of deacetylated xylan. The bands in the yellow rectangle are quantified and their ratio calculated.

Mild acid hydrolysis and HPAEC-PAD was carried out by Henry Temple to determine the monosaccharide sugar composition of hemicelluloses contained in the sample, the results are summarised in Table 5.4.

Some of the GlcA in the xylan could be methylated. However, we did not have calibration standards available for methylated GlcA, and it is hence not known where it elutes during HPAEC-PAD analysis and cannot be quantified either. The suspected peaks are shown in Figure A.9. If either of them is indeed MeGlcA, the calculated GlcA content of the sample is underestimated.

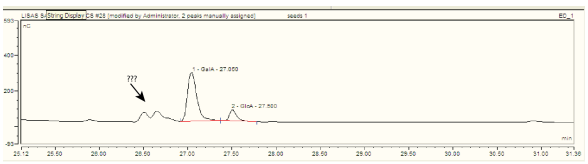


Fig. A.9 HPAEC-PAD elution diagram showing the peaks for galacturonic acid and GlcA (for the leaves). The two small peaks at 26.5 min could be meGlcA.

A.5 Coarse grain molecular dynamics modelling of cellulose-xylan assembly

The procedures to obtain coarse grain parameters for xylan for molecular dynamics modelling of the cellulose-xylan helicoidal assembly were explained in section 4.5. The optimisation of the dihedral angles most relevant for the chain conformation, B1-B2-B1'-B2' and B1-B2-B1'-B3', was depicted in Figure 4.7. The residual angles, B1-B2-B3-B1', B2-B3-B1'-B3', B2-B3-B1'-B2, and B3-B1'-B2'-B3', are depicted below, Figure A.10.

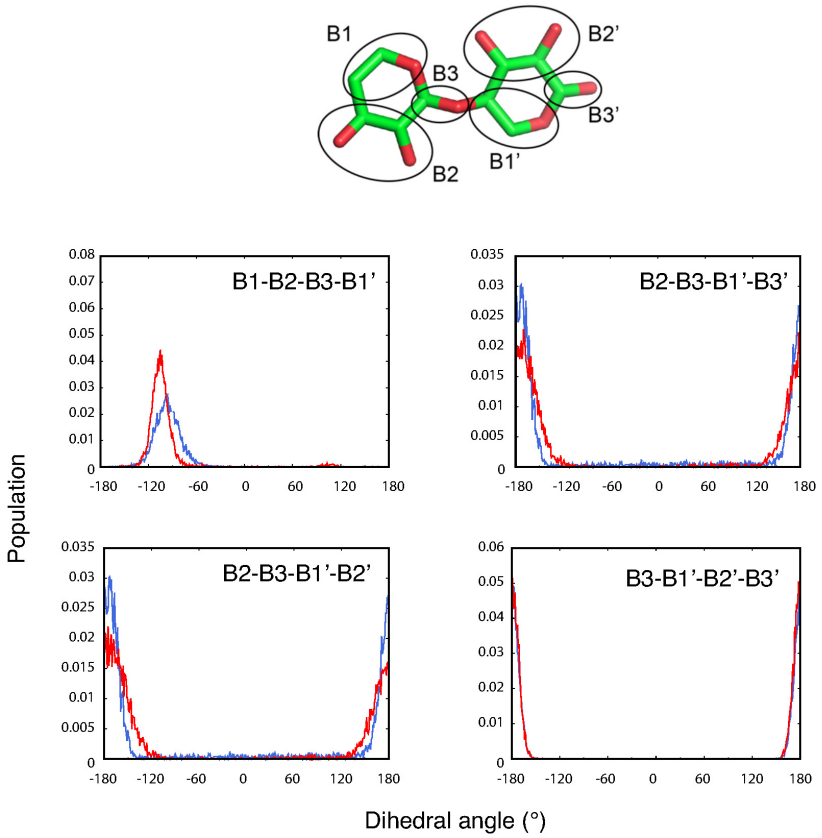


Fig. A.10 Definition of coarse grain beads and adjustment of angular and dihedral parameters. Blue line: all-atom, red line: coarse grain dihedral angle distributions.

A.6 Abbreviations

Table A.4 List of abbreviations used in this thesis

AA	all atom
AIL	acid insoluble lignin
AIR	alcohol insoluble residue
ANTS	8-aminonaphthalene-1,3,6-trisulfonic acid
APS	ammonium persulfate
Ara	arabinose
ASL	acid soluble lignin
CESA	cellulose synthase
CG	coarse grain
CI	crystallinity index
CMF	cellulose microfibrils
CMT	cortical microtubules
CP	cross polarisation
CSC	cellulose-synthesising complex
DHB	2,5-dihydroxybenzoic acid
EtOH	ethanol
GlcA, U	glucuronic acid
LCC	lignin-carbohydrate complex
MALDI-ToF MS	Matrix-assisted laser desorption/ionisation - time of flight mass spectrometry
MAS	magic angle spinning
MeOH	methanol
NMR	nuclear magnetic resonance
ODW	oven dried weight
o/n	overnight
PACE	polysaccharide analysis by carbohydrate gel electrophoresis
2-PB	2- picoline-borane
PCW	primary cell wall
PM	plasma membrane
RMS	root mean square
RT	room temperature
SANS	small-angle neutron scattering
SAXS	small-angle X-ray scattering
SCW	secondary cell wall
SEC	size exclusion chromatography
SRS	sugar recovery standard
ss	solid state
stdev	standard deviation
STEM	Scanning transmission electron microscopy bright field detector
TEMED	tetramethylethylenediamine
TMS	tetramethylsilane
TS	total solids
WAXS	wide-angle X-ray scattering
Xyl,X	xylose