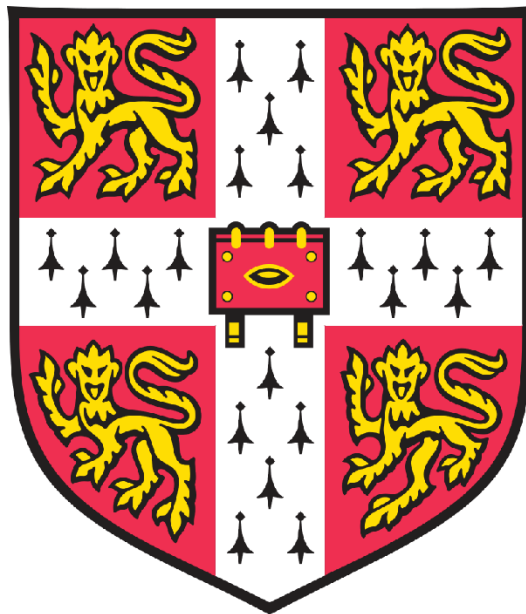# Molecular interactions of plant cell wall polymers

Oliver Morgan Terrett

Clare College

Department of Biochemistry

University of Cambridge

Dissertation submitted for the degree of
Doctor of Philosophy

September 2019

Declaration

The work described in this thesis was carried out in the Department of Biochemistry, University of Cambridge, between October 2015 and September 2019, under the supervision of Professor Paul Dupree. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. Collaborators include Dr Carlos Driemeier, Dr Clelton Dos Santos, Dr Katherine Stott and Professor Ray Dupree. It is not substantially the same as any work 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 where specifically indicated in the text. This dissertation does not exceed 60,000 words (excluding bibliography and references) the prescribed word limit for the relevant Degree Committee.

Molecular interactions of plant cell wall polymers by Oliver Morgan Terrett - Summary

Specialised plant cells produce thickened cell walls, called secondary cell walls comprised of lignocellulose. The main polymers in lignocellulose are cellulose, xylan, galactoglucomannan and lignin. Lignocellulose forms the majority of biomass on the planet and its utilization for construction, energy production, materials and pharmaceuticals may be important for a more sustainable future. For each of these applications, the interactions between the polymers in secondary cell walls are important. Previously, it was proposed that glucuronic acid side chains on xylan form ester bonds with lignin, and that this cross-linking might be important for cell wall properties. These bonds have been hypothesized to form if the glucuronic acid substitutions of xylan participate in lignin polymerization reactions and thereby cross-link xylan and lignin. Supporting this potentially important role in cell wall cross-linking, previous investigations in the model plant Arabidopsis showed that the glucuronic acid branches of xylan are crucial to the recalcitrance of lignocellulose to enzymatic digestion. In this thesis the molecular basis of this change in recalcitrance was investigated. The *gux1 gux2* mutant, which lacks glucuronic acid in secondary cell walls, was found to be more accessible to hydrolytic enzymes, likely due to an increase in the porosity of the cell wall. Investigations with solid-state NMR revealed a reduction in the interactions between lignin and xylan in the mutant plants. Specific lignin synthesis mutants, which have altered lignification chemistry preventing the xylan-lignin cross-linking, were shown to share enzyme accessibility and lignin-xylan interaction phenotypes with the *gux1 gux2* mutant. The presence of ester bonds between lignin and xylan was investigated and introduction of novel xylan-lignin cross-links was attempted. Solid-state NMR was used to extend our understanding of the interactions between the cell wall polymers, in industrially relevant conifer cell walls, which have a significantly higher content of galactoglucomannan than Arabidopsis. It was found that both xylan and galactoglucomannan bind to the cellulose surface and that these polysaccharides interact with lignin. This work demonstrates that some similarities in interactions between lignin and hemicelluloses occur in Angiosperms and Gymnosperms and that these interactions may play roles in the maintenance of plant material properties, such as the recalcitrance to enzymatic digestion.

Contents

3

5

Acknowledgements

Abbreviations

2-PB: 2-picoline borane

4CL: 4-coumarate:CoA ligase

AA: auxiliary activity

ABC: ATP-binding cassette

AFM: atomic force microscopy

AG: apiogalacturonan

AGP: arabinogalactan protein

ANOVA: Analysis of variances

ANTS: 8-aminonaphthalene-1,3,6-trisulphonic acid

APAP1: arabinoxylan pectin arabinogalactan protein1

ARAD: arabinan deficient

Ara*f*: L-arabinofuranose

Ara*p*: L-arabinopyranose

AXY: acetylated xyloglucan

BAHD: BEAT, AHCT, HCBT, and DAT

C3H: coumarate 3-hydroxylase

C4H: Cinnamate 4-Hydroxylase

CAD: Cinnamyl Alcohol Dehydrogenase

CASP: casparian strip domain protein

CAZy: carbohydrate active enzymes

CCoAOMT: caffeoyl CoA methyltransferase

CCR: Cinnamoyl-CoA Reductase

CE: carbohydrate esterase

CESA: cellulose synthase

CGR: cotton Golgi-related

CHI: chalcone isomerase

CHS: chalcone synthase

COBL: cobra-like

COMT: caffeoyl O-methyl transferase

CSC: cellulose synthase complex

CSI: cellulose synthase interacting

CSLN: cellulose synthase-like N

DASH: DNA-sequencer assisted saccharide analysis

DBDOX: dibenzodioxocin

DMSO: dimethyl sulphoxide

DPBB: double-psi β barrel

DSC: differential scanning calorimtery

DUF: domain of unknown function

ESK: eskimo

F3'H: flavonoid 3'-hydroxylase),

F5'H: flavonoid 5'-hydroxylase

F5H: ferulate 5-hydroxylase

FESEM: field emission scanning electron microscopy

FMT: feruloyl-CoA monolignol transferase

FNS: flavone synthase

FUT: fucosyltransferase

GATL: GAUT-like

GAUT: galacturonosyl transferase

GGM: galactoglucomannan

GH: glycosyl hydrolase

Glc: D-glucopyranose

GT: glycosyl transferase

GUX: glucuronic acid substitution of xylan

GXM: glucuronoxylan methyltransferase

HCALDH: hydroxycinnamaldehyde dehydrogenase

HCT: p-hydroxy cinnamoyl-CoA:quinate/shikimate p-hydroxycinnamoyl transferase

HG: homogalacturonan

HPAEC: high-performance anion exchange chromatography

HRGP: hydroxyproline rich glycoprotein

HRP: horse radish peroxidase

HSQC: heteronucelar single quantum cohenrence

HSQC-TOCSY: heteronuclear sungle quantum cohenrence total correlation spectroscopy

INADEQUATE: incredible natural abundance double quantum transfer experiment

IPTG: Isopropyl β-D-1-thiogalactopyranoside

IRX: irregular xylem

KAK: kaktus

LAC: laccase

LB: Luria Broth

LCC: lignin carbohydrate complex

MAGT: mannan alpha galactosyl transferase

MAS: Magic Angle Spinning

MLG: mixed linkage glucan

MS: mass spectrometry

MSR: mannan synthesis related

MUCI: mucilage related

MUR: murus

NMR: nuclear magnetic resonance

NOESY: nuclear Overhauser effect spectroscopy

PACE: polysaccharide analysis by carbohydrate electrophoresis

PAL/TAL: phenylalanine/tyrosine ammonia lyase

PDSD: Proton driven spin diffusion

PHT: hydroxyl cinnamoyl CoA:putrescine hydroxyl cinnamoyl transferase

PME: Pectin Methylesterase

PMT: p-coumaroyl-CoA monolignol transferase

RBOH: respiratory burst oxidase homologue

RG-I: rhamnogalacturonan-I

RG-II: rhamnogalacturonan-II

RGXT: RG-II xylosyltransferase

ROESY: rotating-frame nuclear Overhauser effect correlation spectroscopy

RWA: reduced wall acetylation

SDS-PAGE: sodium dodecyl sulfate–polyacrylamide gel electrophoresis

STS: stilbene synthase

TBL: trichome birefringence like

THT: hydroxyl cinnamoyl-CoA:tyramine N-hydroxyl cinnamoyl transferase

TMS: trimethylsilane

UAFT: UDP-arabinofuranose transporter

UAPT: UDP-arabinopyranose transporter

XAT: xylan arabinosyl transferase

XG: xylogalactruronan

X-Gal: 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

XGD: xylogalacturonan deficient

XLT: xyloglucan galactosyltransferase

XTH: xyloglucan endotransglucosylase/hydrolase

XXT: xyloglucan xylosyltransferase

Xyl: D-xylopyranose

Chapter 1:   Introduction

1.1   The Plant Cell Wall – Definition and function

The extracellular matrix of most higher plant cells is a composite of polysaccharides, polyphenolics and proteins, called the cell wall (Keegstra, 2010). These cell walls are key to many aspects of plant physiology (Burton et al., 2010). Higher plants have two types of cell walls, primary and secondary. The primary cell wall synthesised by most plant cells defines the development of cell shape and affects tissue morphogenesis (Cosgrove, 2018; Yang et al., 2016). Cell walls act as a barrier to the external world, preventing pathogen infection (Bellincampi et al., 2014) and dessication (Bennici, 2008). The specialised thickened secondary cell walls in vessels and fibres support upright growth and transport water throughout the plant, an important adaption to terrestrial ecosystems (Lucas et al., 2013).

The unique functions of each type of cell wall requires a distinct set of properties (Burton et al., 2010). The properties of cell walls, such as their indigestibility, rigidity or flexibility, are directly affected by their composition and the interactions that arise between different cell wall components (Scheller and Ulvskov, 2010; Ralph et al., 2019; Terrett and Dupree, 2019). The interactions between different cell wall components will be referred to as cell wall molecular architecture. Broadly speaking, investigating some aspects of cell wall molecular architecture and attempting to relate these interactions to some cell wall properties is the subject of this thesis.

1.2   Composition and Biosynthesis of Cell Wall Components

The majority of plant cell wall components are polysaccharides, and some proteins that are glycosylated (Keegstra, 2010). Polysaccharides and the polysaccharide component of glycoproteins are mostly synthesised in the Golgi apparatus (Scheller and Ulvskov, 2010), though some are synthesised at the plasma membrane (Somerville, 2006). Cell wall polysaccharides are synthesised by glycosyltransferase (GT) enzymes, which mostly utilise nucleotide phosphate sugars for synthesis (Lairson et al., 2008). These NDP-sugars are synthesised in the cytosol and Golgi and are transported to the site of synthesis by nucleotide sugar transporters (Temple et al., 2016). GTs are split into families based on sequence similarity (Coutinho et al., 2003). Each GT family has an inverting or retaining

catalytic mechanism, where the anomeric configuration of the sugar donor is inverted or retained. The catalytic mechanism is consistent across the GT family (Coutinho et al., 2003), meaning that the geometry of the glycosidic bond produced by all members of a particular GT family must be consistent, see Figure 1.1. The identity of the donor and acceptor utilised by GTs within a GT family varies greatly (Lairson et al., 2008). Each sugar in a specific glycosidic linkage will require a different glycosyltransferase activity and thus synthesis of the large number of different structures can be extremely complex. Section 1.2.1, 1.2.2 and 1.2.4 will describe the known structures of plant cell wall polysaccharides and glycoproteins, along with the known GTs and other enzymes involved in their synthesis.
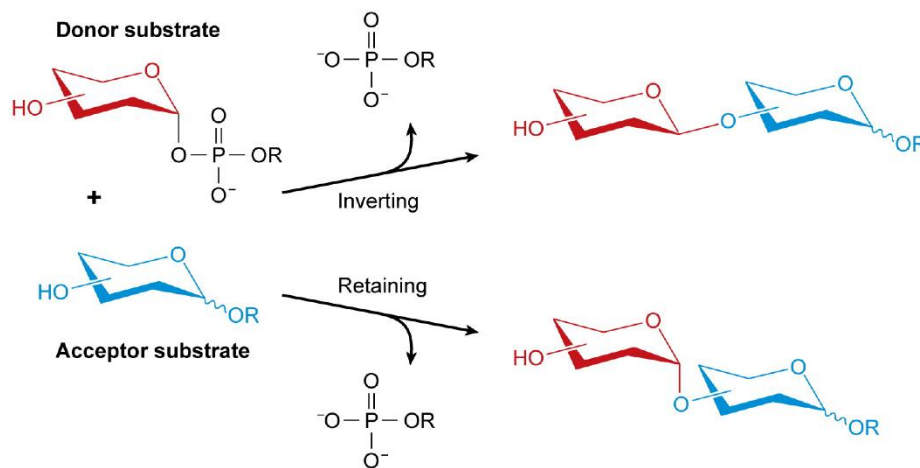
Figure 1.1: Mechanisms of glycosyltransferases. Adapted from Lairson et al., (2008) Glycosyltransferases utilise activated sugars such as NDP- sugars and, add them to acceptors, normally polysaccharides in the case of plant cell wall biology. Enzymes which utilise an enzyme-glycosyl intermediate are retaining, keeping the anomeric conifguration of the donor, while enzymes that utilise a direct displacement mechanism invert the anomeric configuration of the sugar and are inverting. Within a GT family, the mechanism is consistent, though the acceptor and donor vary.

The other major class of plant cell wall components is lignin. Lignin synthesis varies considerably from that of polysaccharides and glycoproteins (Ralph et al., 2004). The synthesis of lignin occurs in the cytoplasm, while the polymerisation is spatially separated from synthesis and occurs in the cell wall (Vanholme et al., 2019). Lignin synthesis and polymerisation incorporates a variety of enzyme families, and will be the subject of Section 1.2.4.

1.2.1 Cellulose and Hemicelluloses

By weight, cell walls are dominated by β-1,4 linked glycans, which are split into the unbranched crystalline polysaccharide cellulose (Somerville, 2006) and the hemicelluloses (Scheller and Ulvskov, 2010). Hemicelluloses also have a linear β-1,4 backbone, but unlike cellulose the backbone can be further substituted. The hemicelluloses are composed of galactoglucomannan (GGM), xyloglucan (XG), mixed linkage glucan (MLG) and evolutionarily related polysaccharides, and xylans (Scheller and Ulvskov, 2010). These polysaccharides vary in proportion and structure phylogenetically and by cell type within a species.

In Arabidopsis, mutants in cell wall synthesis have demonstrated the crucial importance of these polysaccharides. Mutants in cellulose (Taylor et al., 2003) or xylan synthesis (Brown et al., 2007), have collapsed xylem vessels that cannot support normal transpiration. Some such mutants also often have an altered hormonal balance which contributes to their phenotype (Ramírez et al., 2018). The genetic removal of xyloglucan in Arabidopsis results in developmental changes and slight dwarfing (Cavalier et al., 2008; Park and Cosgrove, 2012). GGMs may be important for the gravitropism response of Arabidopsis (Somssich et al., 2019) and embryo survival (Goubet et al., 2003). In plants where GGM forms a larger proportion of the cell wall, they may play more important functions (Scheller and Ulvskov, 2010). Thus, the β-1,4 glycans have a crucial role in plant physiology and development. This Section will describe what is known about the structure and biosynthesis of each plant cell wall β-1,4 glycan.

1.2.1.1 Cellulose

Throughout the plant kingdom, cellulose is the largest component of cell walls by weight, up to 50% (Somerville, 2006). Cellulose is composed of linear chains of β-1,4-linked glucopyranosyl residues. Each glucan chain may be many thousands of glucosyl residues long. The glucan chains crystallise together so that each glucosyl residue in each glucan chain is rotated 180° relative to the previous one, in a two-fold screw configuration (Nishiyama et al., 2003), see Figure 1.2. Hydrogen bonding occurs between hydroxyls on carbons 2, 3, 5 and 6 within and between two-fold screw glucan chains, thus forming sheets of glucan chains (Jarvis, 2018), Figure 1.2. The glucan sheets stack on top of each other through Van der Waals interactions to form the cellulose microfibril. The

conformation of the carbon 6 hydroxymethyl in each glucosyl residue is affected by the interchain hydrogen bonding, see Figure 1.2 (Jarvis, 2018). These structural differences in carbon 6 hydroxymethyl conformation are detectable by solid-state NMR, as they greatly influence the chemical shifts of carbon 4 and 6 (Phyo et al., 2018). The crystal structures differ in the relative position of glucan chains in successive glucan sheets, and the length and orientation of hydrogen bonds (Nishiyama et al., 2003; Jarvis, 2018). Despite many studies on plants, the exact crystal structure of plant cellulose is unknown.
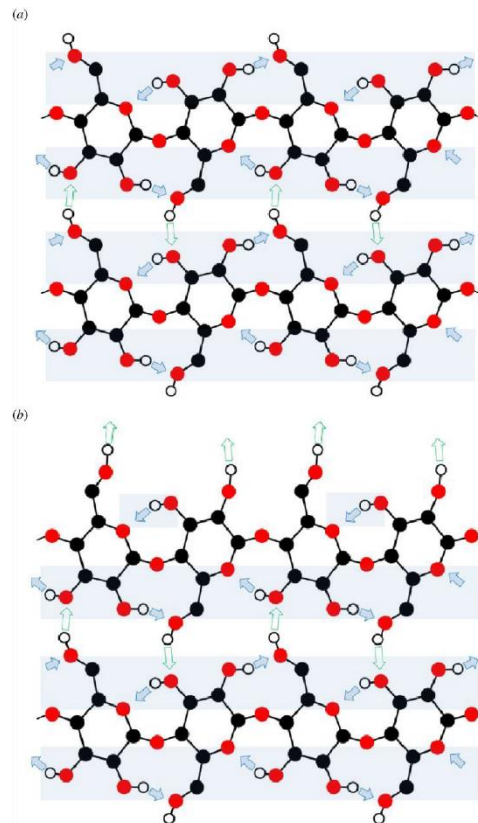
Figure 1.2: Cellulose chain and hydroxymethyl conformation. Adapted from Jarvis, (2018). The glucosyl residues of each chain are shown in a two-fold screw. a) The hydrogen bonding network of glucan chains with a *tg* carbon 6 conformation. b) The hydrogen bonding network of glucan chains with a *gt* carbon 6 conformation.

The microfibril shape and the number of glucan chains in a microfibril varies in a species-dependent manner (Ding et al., 2013; Thomas et al., 2014). These cellulose microfibril parameters may be related to cellulose synthesis (Jarvis, 2018). The glucan chains are synthesised by cellulose synthase (CESA) proteins of the glycosyltransferase family 2 (GT2) (Somerville, 2006). The multitransmembrane CESA proteins synthesise cellulose at the plasma membrane (Kimura et al., 1999). CESA proteins utilise UDP-glucose to synthesise the glucan chains in a processive manner (Morgan et al., 2016). The CESA

proteins form a six-lobed complex called the cellulose-synthase complex (CSC) (Mueller, 1980). Each lobe is thought to contain three CESA proteins (Vandavasi et al., 2016). In Arabidopsis, three CESA classes are found in each CSC, CESA4, 7 and 8 in secondary cell walls and CESA1, 3 and 6 in primary cell walls (Somerville, 2006). In Arabidopsis, the loss of one CESA results in the failure to form the CSC (Taylor et al., 2003). Genetic substitution of one CESA with another does not result in full complementation of cellulose synthesis suggesting functional differentiation in CESA classes within the CSC; CESA proteins may vary in their location in the CSC or in their catalytic activity (Kumar et al., 2018). In primary and secondary cell walls of Arabidopsis each CESA class is found in equal stoichiometry (Zhang et al., 2018). In G layer synthesis in aspen it has been found that the ratios of different CESAs vary, with one CESA dominating, suggesting certain CESA alleles can substitute for others in the CSC (Zhang et al., 2018), and that different CSC compositions can affect the behaviour of the celluloses that are synthesised.

There are numerous accessory activities required for cellulose synthesis. The plasma membrane localised endoglucanase KORRIGAN (Vain et al., 2014) and non-catalytic COBRA-like (COBL) proteins affect the amount of cellulose in the plant (Roudier et al., 2005). In addition KORRIGAN forms a part of the CSC (Vain et al., 2014). Unknown acyltransferases acylate cysteines within CESA proteins, and this is important for cellulose content of the plant and movement of CSCs in the plasma membrane (Kumar et al., 2016). The Golgi glycosyltransferase STELLO affects the movement rate and localisation of CSCs, and the plant cellulose content through an unknown mechanism (Zhang et al., 2016). CESA proteins track along microtubules during cellulose synthesis, and their density and coordinated movement regulates the density and relative orientation of cellulose microfibrils in the cell wall (Lewis, 2016). The associations of the CSC with microtubules are partly mediated by cellulose synthase interactive proteins (CSI1 and 3) (Gu and Somerville, 2010; Lei et al., 2013), and may also be affected by the endoglucanase KORRIGAN (Lei et al., 2014). Though all these activities are important for proper cellulose synthesis, their functions are poorly understood.

1.2.1.2 Galactoglucomannans

Galactoglucomannans (GGMs) are the dominant secondary cell wall polysaccharides in the economically and ecologically important conifer species (Del Lungo et al., 2006; Bradshaw et al., 2009; Scheller and Ulvskov, 2010). GGM is also found in lower amounts

in the primary and secondary cell walls of Angiosperms (Scheller and Ulvskov, 2010), and as storage polymers in seeds and roots (Daud and Jarvis, 1992). GGMs have a β-1,4-linked backbone of glucopyranosyl and mannopyranosyl residues. The ratio of mannose to glucose can vary dramatically from 1:1 in Arabidopsis mucilage (Yu et al., 2018), to 8:1 in spruce wood (Hannuksela and Hervé du Penhoat, 2004), the conifer type GGM is shown in Figure 1.3. In addition to the linear backbone, there are two types of substitutions found on GGM, α-1,6 linked galactopyranosyl and O-2 and 3 linked acetylation (Hannuksela and Hervé du Penhoat, 2004). It has been reported that the α-1,6 linked galactosyl can be β-1,2 substituted with another galactopyranosyl residue in tobacco cell cultures (Sims et al., 1997).
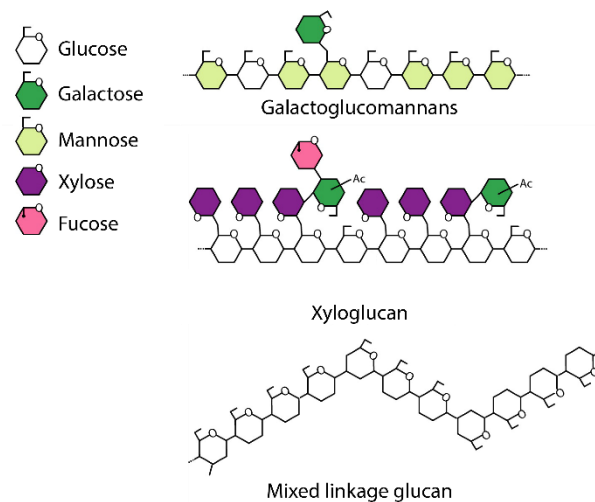
Figure 1.3: Structures of polysaccharides synthesised by CSL enzymes. Adapted from Scheller and Ulskov, (2010). The structures of some CSL-synthesised hemicelluloses are shown.

GGM synthesis differs significantly from the synthesis of cellulose, despite also being synthesised by GT2 enzymes, the <u>c</u>ellulose <u>s</u>ynthase-<u>l</u>ike family <u>A</u> (CSLA) (Davis et al., 2010). *csla2/3/9* mutant plants lack stem GGM (Goubet et al., 2009). GGM is synthesised by CSLAs in the Golgi (Davis et al., 2010; De Caroli et al., 2014), rather than at the plasma membrane. In addition, the CSLA proteins utilise GDP-mannose and GDP-glucose (Liepman et al., 2005) rather than UDP-sugars. It has been reported that the CSLA active site is luminal (Davis et al., 2010; De Caroli et al., 2014), suggesting it would utilise Golgi-localised GDP-sugars, rather than the cytosolic pool (like CESAs). This proposal is controversial however, as related proteins such as CESAs and CSLCs have a cytosolic active site (Davis et al., 2010). In addition, Golgi GDP-mannose transporter

mutants have normal amounts of GGM, suggesting Golgi GDP-mannose is not required for GGM synthesis (Mortimer et al., 2013; Jing et al., 2018). In addition to CSLA proteins, the GT75-related mannan synthesis related (MSR) proteins are important for *in vivo* mannan content (Wang et al., 2013). MSR proteins regulate the total activity of CSLA proteins expressed in *Pichia pastoris*, and they may affect the composition of the GGM produced (Voiniciuc et al., 2019).

The GGM backbone is decorated by glycosyltransferases and acetyltransferases. Mannan α-galactosyltransferases (MAGT) (Yu et al., 2018) from GT34 add galactosyl substitutions to mannosyl residues (Voiniciuc et al., 2015). The MAGT enzymes exhibit strong polysaccharide acceptor specificity. AtMAGT1 will only galactosylate mannose in the backbone context GMGMGM - where G is a glucosyl and M is a mannosyl residue - (Yu et al., 2018), while fenugreek MAGT can galactosylate pure mannan oligosaccharides (Edwards et al., 2002). *In vivo,* the MAGT1 decorated GGM in Arabidopsis mucilage has 60% of the mannosyl residues galactosylated (Yu et al., 2018), in fenugreek, nearly every mannosyl residue is galactosylated, reflecting *in vitro* specificity differences (Edwards et al., 2002). The trichome birefringence like proteins (TBL) are acetyltransferases from the DUF231 family. TBL23/24/25/26 utilise acetyl CoA to acetylate mannosyl residues at carbons 2 and 3 (Zhong et al., 2018). The proteins exhibited significantly different K_m and V_{max} parameters on mannohexaose; this could suggest that the enzymes acetylate mannosyl residues in different backbone contexts.

1.2.1.3 Xyloglucans

Xyloglucans are the most common primary cell wall hemicellulose in eudicots and gymnosperms, and also form a small proportion of monocot primary cell walls (Scheller and Ulvskov, 2010). Xyloglucans have a β-1,4 linked glucosyl backbone with α-1,6 xylosyl substitutions. Interestingly, xyloglucan has a repeating tetrasaccharide in the backbone, either XXGG or XXXG, where X is a xylosylated glucosyl residue and G is an unsubstituted glucosyl residue. The xylosyl substitutions can be further substituted at carbon 2 with a diverse array of sugars, including galactose, arabinose, xylose or at carbon 4 with galactose or galacturonic acid (Zabotina, 2012). Xyloglucan side chains can be even more complex, the β-1,2-galactosylated xylose can be substituted with fucose or galactose, while the arabinosylated xylose can be substituted with an arabinose. In addition to glycosylation, the galactosyl substitutions on the backbone xylosyl

19

substitutions can be acetylated (Schultink et al., 2015). The xyloglucan structure varies phylogenetically (Pauly and Keegstra, 2016) and the dominant Arabidopsis xyloglucan structure is shown in Figure 1.3.

The synthesis of xyloglucan is somewhat similar to GGM. The glucan backbone is synthesised by CSLC proteins (Cocuron et al., 2007). Phylogenetic analysis suggests *CSLA* and *CSLC* are the most closely related clades in the cellulose synthase superfamily (Little et al., 2018). Unlike CSLAs, the CSLC proteins utilise UDP-glucose (Liepman et al., 2005) and have a cytosolic active site (Davis et al., 2010). The xylosyl residues of xyloglucan are added by GT34 enzymes, the xyloglucan xylosyltransferases (XXT) (Chou et al., 2012). XXTs are from the same GT family as MAGTs. Protein interactions between XXTs and CSLCs boost CSLC glucan synthase activity (Cocuron et al., 2007). In Arabidopsis, the structural differences between XXT proteins enable regiospecific activity. XXT1 and 2 add xylose to GGGG and XGGG, while XXT5 can accommodate XXGG and produce the XXXG tetrasaccharide (Culbertson et al., 2018; Zabotina et al., 2012). In Arabidopsis, the galactosyl substitutions of the xylosyl substitutions are added by the GT47 enzymes, xyloglucan galactosyltransferase 2 (XLT2) and murus 3 (MUR3) (Jensen et al., 2012; Madson et al., 2003). XLT2 converts XXXG to XLXG (where L is a galactosylated xylosyl substitution), while MUR3 converts XLXG to XLLG. GT47s from other species add different sugars to the xylosyl substitution (Schultink et al., 2013; Zhu et al., 2017). Xyloglucan fucosyltransferase (FUT1) from GT37 adds fucosyl resdiues to the second L in XLLG to make XLFG (where F is a fucosylated galactosylated xylosyl substitution) (Perrin et al., 1999). In addition, instead of the backbone residues being acetylated as in GGM, the galactosyl substitutions can be O-6 acetylated, by altered xyloglucan 4 (AXY4) from the TBL family (Gille et al., 2011).

1.2.1.4 Other hemicelluloses synthesised by CSL proteins

The mixed linkage glucans (MLG) form a significant component of monocot cell walls (Scheller and Ulvskov, 2010), especially in growing cells and senesced tissues (Vega-Sánchez et al., 2013). The β-1,4 glucan backbone is interrupted by single β-1,3-linked glucopyranosyl residues every three or four β-1,4-linked residues, although longer stretches of β-1,4-linked residues exist in grass MLG (Sørensen et al., 2008). The frequency of the β-1,3 linked glucosyl residues varies phylogenetically (Sørensen et al., 2008). Other structurally similar polysaccharides exist, for instance an arabinoglucan,

containing β-1,4-linked glucopyranosyl resdiues and α-1,3 linked arabinofuranosyl residues (Roberts et al., 2018). In addition, a mixed composition glucoxylan with β-1,4-linked glucopyranosyl and xylopyranosyl residues has been observed (Little et al., 2019). The structure of mixed linkage glucan is shown in Figure 1.3.

The structures described above are largely synthesised by enzymes from the CESA superfamily. The CESA superfamily has 11 clades within it, (Little et al., 2018), including the previously discussed CESAs, CSLAs, and CSLCs. The CSLD family is thought to synthesise a mannan (Verhertbruggen et al., 2011) or cellulose polysaccharide (Park et al., 2011; Douchkov et al., 2016). Two CESA superfamily clades have been shown to synthesise mixed linkage glucans in grasses. The CSLF (Burton et al., 2006) and CSLH (Doblin et al., 2009) clades have all been reported to synthesise mixed linkage glucans, when expressed in Arabidopsis. In addition to mixed linkage glucan, some CSLF genes from Barley have been shown to synthesise glucoxylan when expressed in tobacco leaves (Little et al., 2019), suggesting the diversity of this single CSL sub-clade allows the synthesis of different glycosidic linkages and the utilisation of different UDP-sugars. A moss GT2 enzyme, more closely related to fungal enzymes than CESA, synthesises the mixed linkage arabinoglucan (Roberts et al., 2018).

1.2.1.5 Xylans

Xylan is a common hemicellulose throughout gymnosperm and angiosperm cell walls (Scheller and Ulvskov, 2010). Its structure and synthesis differs significantly from the CSL synthesised polysaccharides. In gymnosperms and eudicots, xylan forms between 10-25% of secondary cell walls by weight. In the monocots, xylan forms a significant component of both primary and secondary cell walls, up to 50% by weight. Xylans have a β-1,4-linked xylosyl backbone. In eudicots, gymnosperms and some monocots this is capped by a reducing end oligosaccharide, with a xylose β-1,3 linked to a rhamnosyl residue, α-1,3 linked to a galacturonosyl residue, α-1,4 linked to a xylosyl residue, see Figure 1.4. The backbone xylosyl residues can be substituted at carbons 2 or 3. In monocot, eudicots and gnetophytes, both carbons can be acetylated (Busse-Wicher et al., 2014; Tryfona et al., 2019), while in other gymnosperms there is no xylan acetylation (Busse-Wicher et al., 2016). Across angiosperms and gymnosperms xylans have α-1,2 glucuronosyl substitutions (Mortimer et al., 2010; Busse-Wicher et al., 2016; Tryfona et al., 2019), though some tissues lack this substitution (Mares and Stone, 1973). In some

21

tissues and plants the glucuronic acid can be substituted with α-1,2 arabinopyranosyl or β-1,2 galactopyranosyl residues (Mortimer et al., 2015; Peña et al., 2016). In addition, α-1,2 and 3 linked arabinofuranosyl backbone substitutions are present throughout higher plants (Scheller and Ulvskov, 2010; Busse-Wicher et al., 2016). The backbone arabinosyl substitutions can be esterified at carbon 5 by *p*-coumaric acid or ferulic acid in commelinid monocots (Scheller and Ulvskov, 2010), and the esterified arabinosyl substitutions can also have β-1,2 xylopyranosyl substitutions. Some plants, such as rice and Arabidopsis may have α-1,2 linked xylosyl substitutions on their xylan (Zhong et al., 2018), though the evidence for this is more indirect, such as glycosidic linkage analysis (Voiniciuc et al., 2015). There may also be some galactopyranosyl substitutions on highly substituted maize xylans, attached with an unknown linkage (Chanliaud et al., 1995; Rogowski et al., 2015). The structure of commelinid monocot, eudicot and conifer xylan is shown in Figure 1.4.
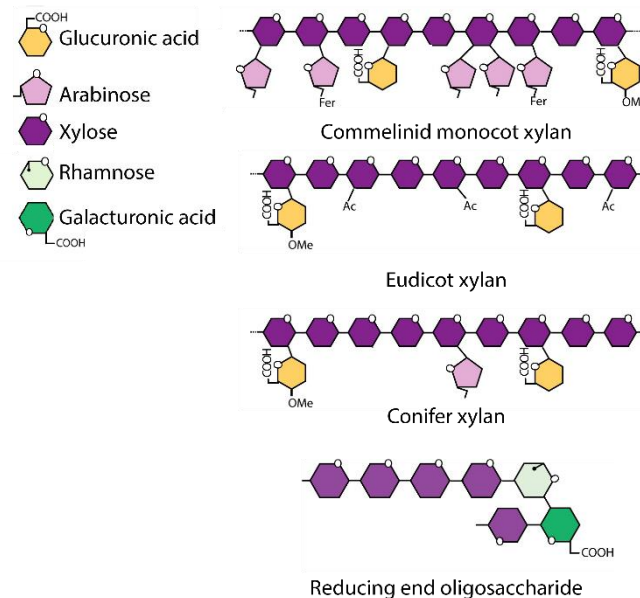
Figure 1.4: Representative structures of canonical higher plant xylans and the reducing end oligosacchride of xylans. Adapted from Scheller and Ulskov, (2010).

The structure of xylan is noticeably different to the other β-1,4 glycans in plant cell walls, due to the presence of the reducing end oligosaccharide (Scheller and Ulvskov, 2010). The reducing end oligosaccharide is thought to be synthesised by the IRX7, IRX8 and PARVUS proteins (Brown et al., 2007), as the corresponding mutants lack the reducing end oligosaccharide. In addition, the mutants have a reduced xylan content, despite only having slightly reduced microsomal xylosyltransferase activity. Thus the reducing end

oligosaccharide may act as a primer for xylan extension. The reducing end oligosaccharide has been suggested to be added to an arabinogalactan protein acceptor, APAP1 (Tan et al., 2013), though the proposed structure linking the xylan to rhamnogalacturonan-I (RG-I) or AGP is not the same as the xylan reducing end oligosaccharide. *IRX8* and *PARVUS* are both from the GT8 family, IRX8 is a member of the GAUT clade, while PARVUS is a member of the GATL clade (Peña et al., 2007). The GT8 family are retaining glycosyltransferases (Sterling et al., 2006; Mortimer et al., 2010), and the GAUT proteins have been reported to have *in vitro* galacturonoysltransferase activity, suggesting IRX8 is a strong candidate to transfer the galacturonic acid onto the reducing end oligosaccharide. The GATL clade has no reported UDP-sugar specificity yet, but it could add the α-linked rhamnose to the reducing end oligosaccharide, or the reducing end xylose to an unknown acceptor (Yin et al., 2010; Yin et al., 2010). IRX7 is a member of the GT47 family, inverting glycosyltransferases, and thus generates β-linked sugars (Busse-Wicher et al., 2014). IRX7 could add the β-1,3 xylosyl residue. There is currently no biochemical evidence for the activity of these three proteins, and the acceptor for xylan synthesis is unknown.

The xylan backbone elongation is dependent on at least three glycosyltransferases, IRX9, IRX10 and IRX14 (Scheller and Ulvskov, 2010). IRX10 is a GT47 enzyme (Brown et al., 2009), while IRX9 and 14 belong to family GT43. *irx9*, *10* or *14* Arabidopsis mutant microsomes have reduced xylosyltransferase activity (Brown et al., 2007; 2009), suggesting all proteins are required for normal xylan backbone elongation. When heterologously expressed alone, Arabidopsis (Urbanowicz et al., 2014), psyllium and *Physcomitrella* IRX10 have xylosyltransferase activity (Jensen et al., 2014). These results suggest IRX10 extends the xylan backbone *in vivo*. When heterologously expressed altogether, the three Asparagus proteins have greatly enhanced xylosyltransferase activity over IRX10 alone (Zeng et al., 2016). *In vivo*, wheat IRX9, 10 and 14 form a complex (Zeng et al., 2010). Thus interactions between IRX10 and the other two proteins may be important for boosting IRX10 activity. While IRX9 does not contain a DxD motif (important for binding NDP-sugars), IRX10 and IRX14 do have putative DxD motifs that, when mutated, abolish xylosyltransferase activity *in vitro* (Ren et al., 2014; Zeng et al., 2016). Due to a lack of NDP-sugar binding, IRX9 probably plays a structural role in the complex, rather than a direct glycosyltransferase activity. Xylan backbone elongation would require the addition of a β-1,4 xylosyl residue to the reducing end oligosaccharide β-1,3 xylosyl residue, followed by extension of the β-1,4 linked xylosyl residue identical

23

residues. Based on *in vitro* activity, IRX10 is responsible for xylan elongation (Urbanowicz et al., 2014), while the initial extension of the reducing end oligosaccharide could be performed by IRX14. However this does not explain why a complex with a DxD mutated IRX14 has reduced xylosyltransferase activity onto a β-1,4 xylo-oligosaccharide (Ren et al., 2014).

Different xylan synthesis mutants have different effects on the length of the xylan produced. The reducing end oligosaccharide mutants have more heterodisperse xylan length (Peña et al., 2007) than wild type, while the backbone elongation mutants have shorter xylans (Brown et al., 2009). Interestingly, other mutants affect the xylan length. The IRX15/15L proteins are in the DUF579 family, which have been characterised to be glucuronic acid methyltransferases (Urbanowicz et al., 2012; Temple et al., 2019). The *irx15/15l* double mutant has shorter xylan than wild type, while the overexpressor synthesises longer xylans than the wild type (Jensen et al., 2011). Thus these putative methyltransferases may regulate the activity of either the elongation or reducing end oligosaccharide synthetic enzymes to affect xylan length, or they may methylate an unknown structure on the xylan backbone.

Decoration of the xylan backbone involves numerous enzyme families. The GT8 glucuronic acid substitution of xylan (GUX) enzymes add glucuronosyl substitutions to the xylan backbone (Mortimer et al., 2010). The Arabidopsis *gux1 gux2* mutant lacks glucuronic acid substitutions on its secondary cell wall xylan, and heterologously expressed GUX enzymes transfer glucuronic acid to acetylated xylan from the *gux1 gux2* mutant (Lyczakowski et al., 2017). Most glucuronic acid substitutions are methylated at carbon 4 by glucuronoxylan methylation (GXM) enzymes (Urbanowicz et al., 2012). Interestingly, different GUX enzymes add glucuronosyl substitutions in different patterns on the xylan backbone (Bromley et al., 2013). The GUX1 enzyme adds glucuronic acid substitutions to the backbone with an even number of xylosyl residues between consecutive substitutions, while the GUX2 enzyme produces more clustered substitutions with no preference for odd or even spacings. The spacing of acetyl substitutions also has a preference for even spacings, with acetylations occurring every other xylosyl residue (Busse-Wicher et al., 2014). The even spacing of xylan substitutions is also found in gymnosperms (Busse-Wicher et al., 2016).

The spacing of glucuronic acid substitutions is dependent upon the patterning of acetyl substitutions in Arabidopsis. Xylosyl residues are acetylated by enzymes from the TBL

family (Lefebvre et al., 2011; Xiong et al., 2013; Yuan et al., 2016; Yuan et al., 2016). The TBLs utilise acetyl CoA (Urbanowicz et al., 2014) thought to be imported into the Golgi by reduced wall acetylation (RWA) proteins (Lee et al., 2011; Manabe et al., 2013). TBL29 adds the majority of acetyl substituents to the xylan backbone; the *tbl29/eskimo1* (*esk1*) mutant has a 60% reduction in xylan acetylation (Lefebvre et al., 2011). The majority of ESK1 acetylations are evenly spaced acetyl substituents (Grantham et al., 2017). Interestingly, the *esk1* mutant has altered glucuronic acid substitution patterns, with a loss of the even spacing of the glucuronic acid substitutions added by GUX1 (Grantham et al., 2017). *TBL28* complements the *tbl29* mutant growth phenotype, but other TBLs do not (Zhong et al., 2017). However, they do add acetate to xylan, but the full details of their activity are not clear yet (Yuan et al., 2016; Yuan et al., 2016). Interestingly, AXY9, a protein distantly related to TBLs, affects the acetylation of xyloglucan and xylan, and has been proposed to act as an intermediate in polysaccharide acetylation by TBL proteins (Schultink et al., 2015).

Xylan acetylation is also regulated by esterases from the GDSL family (Zhang et al., 2017). Two GDSL esterases in rice have been reported to have effects on xylan structure (Zhang et al., 2017; Zhang et al., 2019). One of these esterases has been reported to remove acetyl susbtitutions directly from the xylan backbone (Zhang et al., 2017), based on *in vitro* enzyme activities and *in vivo* xylan structural changes. The other reported GDSL esterase removes acetates that are found on the arabinosyl substitutents of grass xylan (Zhang et al., 2019; Ishii, 1991). The role of these esterases in xylan synthesis is not understood.

The arabinosyl substitutions on the xylan backbone are added by clade A enzymes from the GT61 family (Anders et al., 2012). The α-1,3 arabinosyltransferase function of the wheat GT61 xylan arabinosyltransferase (XAT) was demonstrated by expression in Arabidopsis. An Arabidopsis clade B GT61 and a rice clade A GT61 have been proposed to substitute xylosyl residues at the 2 position (Voiniciuc et al., 2015; Zhong et al., 2018). Based on linkage analysis (Voiniciuc et al., 2015), and NMR (Zhong et al., 2018), it was concluded that the substitutions were themselves xylosyl residues. Interestingly, another clade A GT61, XAX1, has been reported to have β-1,2 xylosyltransferase activity onto arabinosyl substitutions of xylan (Chiniquy et al., 2012). However, this activity has been disputed. In unpublished work, it was proposed that XAX1 adds ferulated arabinosyl substitutions to xylan (Feijão, 2016), and because these are the only xylosylated

arabinosyl substitutions, the reduction in ferulated substitutions results in fewer xylosylated arabinosyl substitutions.

In addition to GT61 enzymes, the synthesis of esterified arabinosyl substitutions involves the cytosolic BAHD enzymes (Terrett and Dupree, 2019). BAHD genes were originally proposed to be involved in xylan ferulation due to their co-expression with xylan synthetic genes (Mitchell et al., 2007). BAHD genes utilise phenylpropanoids acylated with coenzyme A (Moglia et al., 2016) as a substrate to acylate numerous substrates including sugars and other phenylpropanoids (Delporte et al., 2018; Karlen et al., 2018). RNAi of four BAHD genes in rice resulted in a reduction in ferulate content (Piston et al., 2010), while RNAi of specific BAHD01 genes leads to reduced ferulate content of numerous species (de Souza et al., 2018; de Souza et al., 2019). Overexpression of BAHD genes from other clades leads to an increase in coumarylation of arabinose, and a decrease in ferulation (Bartley et al., 2013).

1.2.2 Pectins

Pectins are a group of structurally related acidic polysaccharides (Atmodjo et al., 2013). The common structural feature of pectins is the presence of galacturonic acid in the backbone, though this varies in proportion. Pectins are a key component of primary cell walls, especially in non-graminaceous higher plants. The importance of pectin is shown by mutants in their synthesis. Mutants that have less pectin, or alterations in pectin structure, can have severe developmental phenotypes (O'Neill et al., 2001; Wang et al., 2013).

1.2.2.1 Homogalacturonan, xylogalacturonan and apiogalacturonan

Homogalacturonan (HG) is the most abundant pectin fraction. HG is composed of a backbone of α-1,4 linked galacturonopyranosyl residues (Atmodjo et al., 2013). The galacturonosyl resides can be methyl esterified and acetylated at carbon 2 or 3. β-1,3 xylosylation and β-1,2/3 apiosylation convert the HG to xylogalacturonan (XG) and apiogalacturonan (AG), respectively. These molecules may form one continuous polymer with different domains (Mohnen, 2008).

Multiple glycosyltransferase activities are involved in the synthesis of HG, XG and AG. The backbone of homogalacturonan is synthesised by galacturonosyl transferase (GAUT)

enzymes from family GT8 (Sterling et al., 2006). GAUT1 and GAUT7 form an immunoprecipitable complex, with GAUT1 providing biosynthetic activity and the non-catalytic GAUT7 anchoring the two proteins to the Golgi membrane (Atmodjo et al., 2011). Mutants of some GAUT genes affect the level of galacturonic acid in the cell wall (Caffall et al., 2009). A GT47 family member, xylogalacturonan deficient 1 (XGD1), adds xylosyl residues to HG to synthesise XG (Jensen et al., 2008), though other enzymes are likely to be involved, since the xgd1 mutant retains some XG. The GT47 family could also include the AG synthase, but this has not been identified yet.

HG is methylated as it is synthesised in the Golgi (Atmodjo et al., 2013). The cotton-Golgi related (CGR) proteins are strong candidates for galacturonosyl residue methylation (Kim et al., 2015). cgr2 cgr3 double mutants have reduced Golgi homogalacturonan methylation activity and methylation, while overexpression results in the opposite phenotype. In addition, the QUASIMODO2 (QUA2) family has been proposed to methylate homogalacturonan, but there is contradictory evidence, with the qua2 mutant having a reduced homogalacturonan content, but normal levels of methylation (Atmodjo et al., 2013). The degree of methylation is altered in muro by a large family of pectin methylesterase (PME) proteins (Micheli, 2001). The activity of PMEs is tuned by PME inhibitors produced by the plant (Wormit and Usadel, 2018).

1.2.2.2 Rhamnogalacturonan-I

Rhamnogalacturonan-I (RG-I) forms up to 30% of pectin (Atmodjo et al., 2013). RG-I has an alternating backbone of α-1,4 linked galacturonosyl residues and α-1,2 linked rhamnosyl residues. Both carbon 2 and 3 of the galacturonosyl residues can be acetylated. RG-I can have long and complex side chains from carbon 4 of rhamnosyl residues. These side chains can be α-1,5 arabinans, themselves modified at carbon 2 with arabinan. In addition there can be β-1,4 galactan, itself branched with arabinan at carbon 3, or β-1,3 galactan at carbon 3. Some RG-I may also be linked to AGPs (Tan et al., 2013).

The RG-I backbone is thought to be synthesised by enzymes from family GT8 and GT106. Enzymes from two separate GT8 clades, GAUT11 and GATL5, have been suggested to add galacturonic acid to the RG-I backbone (Caffall et al., 2009; Voiniciuc et al., 2018; Kong et al., 2013). GAUT11 has been demonstrated in vitro to synthesise HG, so may not

be directly involved in RG-I backbone synthesis. The MUCI70 gene, which is distantly related to GT8, may also be involved in RG-I backbone synthesis (Voiniciuc et al., 2018). The RG-I rhamnosyltransferases (RRT) from family GT106 can transfer rhamnose to RG-I oligosaccharides, and the *rrt1* mutant has a reduction in mucilage RG-I content (Takenaka et al., 2018). The backbone residues of RG-I are acetylated by the TBL10 protein (Stranne et al., 2018). The *tbl10* mutant has reduced acetic acid in pectic fractions containing RG-I.

The synthesis of RG-I side chains involves several GT families. The β-1,4 galactan side chains are synthesised by the GT92 protein galactan synthase 1 (GALS1) (Liwanag et al., 2012) The α-1,5 arabinan side chains are partly synthesised by the GT47 <u>ara</u>binan <u>d</u>eficient 1 and 2 (ARAD1/2) proteins (Harholt et al., 2006). Although arabinan epitopes were reduced in *arad1* mutant plants, the glycosyl linkage data indicated that the number of RG-I branches was the same, suggesting the chains are initiated by other enzymes, or arabinan chains are replaced by galactan chains. The genes involved in substitution of the side chains are currently unknown.

1.2.2.3 Rhamnogalacturonan-II

Rhamnogalacturonan-II (RG-II) is a minor component of pectin with important biological functions. RG-II has a HG backbone, but with four unique and structurally conserved side chains (Atmodjo et al., 2013). RG-II includes 13 different sugars combined together in 21 different glycosidic linkages. The structure is shown in Figure 1.5, but is not written out here due to its complexity. The other pectin structures are also included in Figure 1.5.

RG-II synthesis is poorly understood. The only GTs that have been confirmed to be involved are <u>RG</u>-II <u>x</u>ylosyl<u>t</u>ransferases (RGXTs), from GT77 (Egelund et al., 2006). The *rgxt4* mutant has reduced methyl xylose substitutions (Liu et al., 2011). However, IRX10, the GT47 with β-1,4 xylosyltransferase activity (Zeng et al., 2016), was previously identified as an RG-II glucuronosyltransferase, illustrating the importance of verifying *in vitro* GT activity (Iwai et al., 2002).
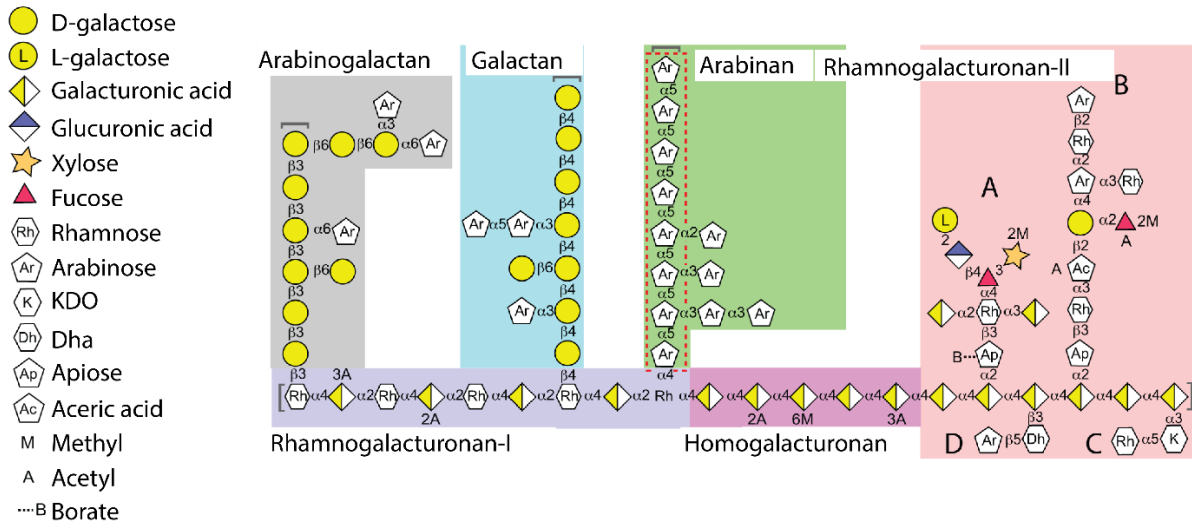
Figure 1.5: Structures of pectic polysaccharides. Adapted from Martens et al., (2011). The three main components of pectin are shown, HG, RG-I and RG-II. RG-II side chains are labelled A-D. In addition, arabinan, galactan and arabinogalactan structures are shown as RG-I side chains.

1.2.3 Lignin

Lignins form up to 30% of plant biomass by weight (Ralph et al., 2004). Lignins are composed of metabolites from phenylpropanoid and flavonoid biosynthesis synthesised in the cytoplasm (Vanholme et al., 2019). Lignins are then polymerised in the extracellular space (Ralph et al., 2019). The genetic removal or modification of lignin can result in extremely dwarfed plants with collapsed vessels (Van Acker et al., 2013), though some of these effects may not be due to the function of lignin in the cell wall but rather the accumulation of phytotoxic intermediates (Bonawitz et al., 2014).

1.2.3.1 Lignin monomers and their synthesis

35 different compounds have so far been found to be incorporated in lignins from different plants (Vanholme et al., 2019), see the red boxes in Figure 1.6. All these monomers share a core of a phenyl group hydroxylated at carbon 4. The canonical monomers are the three hydroxycinnamyl alcohols *p*-coumaryl alcohol (1.1 in Figure 1.6), coniferyl/guaiacyl alcohol (1.3) and sinapyl alcohol (1.5) (Ralph et al. 2004). In the lignin polymer, units derived from these alcohols are referred to as H, G and S units. In many plants esterified forms of these monolignols are also found incorporated into the polymer, e.g. coniferyl *p*-coumarate (3.7) in monocots (Smith et al., 2015). Intermediates

Figure 1.6: Phenylpropanoid metabolism. Adapted from Vanholme et al., (2019). The metabolites of phenylpropanoid metabolism are shown. The enzymes that convert the metabolites are labelled, acronyms are expanded in the text, apart from HCT (*p*-hydroxy cinnamoyl-CoA:quinate/shikimate *p*-hydroxycinnamoyl transferase), STS (stilbene synthase), THT (hydroxyl cinnamoyl-CoA:tyramine N-hydroxyl cinnamoyl transferase), PHT (hydroxyl cinnamoyl CoA:putrescine hydroxyl cinnamoyl transferase), CHS (chalcone synthase), CHI (chalcone isomerase), FNS (flavone synthase), F3'H (flavonoid 3'-hydroxylase), F5'H (flavonoid 5'-hydroxylase), HCALDH (hydroxycinnamaldehyde dehydrogenase), PMT (p-coumaroyl-CoA monolignol transferase) and FMT (feruloyl-CoA monolignol transferase). The names of metabolites that have been verified to incorporate into lignin, in normal or mutant plants, are surrounded by a red box and are numbered.

from hydroxycinnamyl alcohol synthesis can also be incorporated into lignins, and this is especially pronounced in phenylpropanoid synthesis mutants (Ralph et al., 2019). A product of flavonoid biosynthesis, tricin, is an important monomer in monocot lignin (Lan et al., 2015). Stilbene derivatives of hydroxycinnamyl alcohols (10.1-3) are also found in palm lignins (del Río et al., 2017). In addition hydroxycinnamides (7.1-2) can be found in some lignins (del Río et al., 2018). Lignins in any one plant do not contain all 35 identified monomers.

Hydroxycinnamyl alcohols are synthesised from phenylalanine and, in monocots, also tyrosine by the respective ammonia lyase (PAL/TAL) enzymes (Barros et al., 2016). RNAi downregulation of PAL and TAL results in reductions in lignin content (Sewalt et al., 1997; Barros et al., 2016). The cinnamic acid produced from phenylalanine is then hydroxylated at position 4 by cinnamate 4- hydroxylase (C4H) (Teutsch et al., 1993). The *c4h* mutant has reduced lignin content, and accumulates novel cinnamate esters (Schilmiller et al., 2009). The *p*-coumaric acid produced by C4H or TAL, can be acylated by coenzyme A by the enzyme 4-coumarate:CoA ligase (4CL) (Li et al., 2015). At this point the *p*-coumaroyl CoA can be diverted into flavonoid biosynthesis (Vanholme et al., 2019). Tricin, a monomer in monocot lignin, is synthesised from this pathway (Lan et al., 2015). In Arabidopsis, different 4CL isoenzymes have different roles, with the *4cl1* mutant having reduced lignin content, while *4cl3* has reduced anthocyanin (a flavonoid) content (Li et al., 2015).

p-Coumaric acid can also be hydroxylated by coumarate 3-hydroxylase (C3H) (Franke et al., 2002), producing caffeic acid, which is methylated by caffeoyl O-methyl transferase (COMT) to produce ferulic acid (Vanholme et al., 2019). The C3H mutant, reduced epidermal fluorescence 8 (ref8), has lignin composed mainly of *p*-coumaroyl alcohol (Bonawitz et al., 2014). *In vitro*, COMT can methylate caffeic acid (Zubieta et al., 2002). C3H can act on *p*-coumaroyl CoA, but COMT cannot act on the caffeoyl CoA that C3H produces, instead, caffeoyl CoA methyltransferase (CCoAMT) methylates caffeoyl CoA to feruloyl CoA (Guo et al., 2001). Mutants in COMT and CCoAOMT activity have reduced lignin content, and higher proportions of 5-hydroxyconiferyl alcohol (Guo et al., 2001). Like 4CL, methyltransferase isoforms may direct substrates into synthesis of specific phenylpropanoids. A sorghum CCoAOMT, when overexpressed, doesn't affect lignin composition or content but causes increases in ferulic acid content (Tetreault et al., 2018). The *p*-coumaroyl CoA can be converted to *p*-coumaryl shikimate and then

hydroxylated by C3H (Hoffmann et al., 2004; Shadle et al., 2007). The caffeoyl shikimate can then by converted to caffeic acid by caffeoyl shikimate esterase (CSE) and converted back to a CoA derivative by 4CL (Vanholme et al., 2013). The function of this pathway is unclear, though the *cse* mutant has significant reductions in lignin content.

p-Coumaroyl, caffeoyl and feruloyl CoA are all converted to hydroxycinnamaldehydes by cinnamoyl CoA reductase (CCR). The *irx4* mutant is mutated in CCR, and has a 50% reduction in lignin content (Jones et al., 2001). The hydroxycinnamaldehydes (apart from *p*-coumaraldehyde) can be hydroxylated and methylated by ferulate 5-hydroxylase (F5H) and COMT (Vanholme et al., 2019). The *f5h* mutant has G lignin only, due to the inability to produce 5-hydroxylconiferaldehyde or alcohol (Chapple et al., 1992). The aldehyde groups are reduced to a hydroxyl by promiscuous cinnamyl alcohol dehydrogenases (CAD), *cad* mutants have hydroxycinnamaldehyde lignin (Sibout et al., 2005; Zhao et al., 2013). The coniferyl CoA can be converted to the hydroxycinnamides 7.1 and 7.2 by enzymes from the BAHD family (Onkokesung et al., 2012). In addition, BAHD enzymes have been shown to esterify the hydroxycinnamyl alcohols with *p*-coumaoyl CoA (Sibout et al., 2016) and feruloyl CoA (Wilkerson et al., 2014). The biosynthetic enzymes for p-hydroxybenzoylation, and acetylation of hydroxycinnamyl alcohols have not yet been identified.

Substrate channelling and protein-protein interactions may play important roles in lignin monomer synthesis. Different isoenzymes of 4CL and CCoAOMT may have different substrate specificities (Zubieta et al., 2002), or are able to direct their substrates to different pathways (Li et al., 2015; Tetreault et al., 2018). The CCR and CAD enzymes have been shown to form a complex in Poplar (Yan et al., 2019), and it has been suggested that reductions in either CAD or CCR result in reciprocal reductions in the other enzymatic activity, however this is not supported by the relatively normal amount of lignin in *cad* mutant plants (Sibout et al., 2005). The C3H, C4H and F5H enzymes are all cytochrome P450s. All three enzymes have been found to localise to the cytosolic face of the ER (Gou et al., 2018), and this localisation is dependent on scaffold proteins.

Once synthesised in the cytosol, the 35 monomers are exported to the cell wall. Inhibitor studies have suggested that monolignol export over the plasma membrane is dependent on the activity of ATP-binding cassette (ABC) family transporters (Miao and Liu, 2010). So far, only one bona fide lignin monomer transporter has been found to have an effect on lignin synthesis, ATABCG29 (Alejandro et al., 2012). The *atabcg29* mutant has small

reductions in lignin content, and has *p*-coumaroyl transport activity when expressed in yeast. However, the plasticity of lignin synthesis, where many different monomers can be easily incorporated into the polymer (Vanholme et al., 2019) suggests that there are more promiscuous transport mechanisms, such as diffusion across the plasma membrane (Boija et al., 2007), though these two mechanisms are not mutually exclusive (Perkins et al., 2019).

1.2.3.2 Lignin polymerisation

In muro, the lignin monomers are oxidised into radicals by proteins or already formed radicals (Tobimatsu and Schuetz, 2019). Both laccases and peroxidases have been shown to oxidise lignin monomers. Laccases utilise oxygen to oxidise lignin monomers, while peroxidases utilise hydrogen peroxide. An Arabidopsis mutant, *laccase4*, has a 30% reduction in laccase activity and a 20% reduction in lignin content (Berthet et al., 2011). Removal of apoplastic hydrogen peroxide from a spruce cell culture system resulted in reduced extracellular lignin, suggesting peroxidase activity is important for lignin polymerisation (Laitinen et al., 2017).

Different oxidative proteins may have substrate preferences that affect the structure of the lignin they produce. The *lac4* mutant has a normal S:G ratio, but reduced lignin content, while *lac17* has an increased S:G ratio and reduced lignin content (Berthet et al., 2011). Overexpression of different Miscanthus laccases in Arabidopsis caused formation of lignins with subtly different compositions (He, 2019). These results suggest that laccases vary in their substrate specificity. Modelling of peroxidases with different substrates suggests that some are unable to oxidise sinapyl alcohol (Østergaard et al., 2000), while analysis of specific mutants suggests other peroxidases have a preference for sinapyl alcohol (Fernández-Pérez et al., 2015). Other purified peroxidises have more surface localised active sites which enable oxidiation of larger lignin oligomers (Shigeto et al., 2013).

Oxidative proteins localise differently and may have different functions in lignification. LACCASE4 is distributed throughout xylem secondary cell walls, but PEROXIDASE64 localises to the cell corners and middle lamella of fibres (Chou et al., 2018). The *in muro* localisation within the laccase/peroxidase family may vary, and laccase expression patterns suggest different laccases may form lignins in different tissues (Turlapati et al.,

33

2011). Laccases are immobile within secondary cell walls, but can freely diffuse through primary cell walls, suggesting they may be immobilised by binding to unique features of secondary cell walls (Chou et al., 2018). It has recently been shown that helicoidal deposition patterns of lignin and xylan still occur in cellulose synthesis mutants (Takenaka et al., 2018). Since some lignification occurs after polysaccharide synthesis (Smith et al., 2013), it is possible that xylan directs laccase localisation and hence lignin deposition. In the Casparian strip, PEROXIDASE64 appears to bind to <u>cas</u>parian strip domain <u>p</u>rotein 1 (CASP1) in the ER (Lee et al., 2013). The <u>r</u>espiratory <u>b</u>urst <u>o</u>xidase <u>h</u>omologue <u>f</u> (RBOHF) protein also accumulates with CASP1 in plasma membrane domains (Kalmbach et al., 2017). These protein interactions may lead to specific localisation of peroxidases in the wall, while the RBOHF localisation with CASP1 may lead to localised H_2O_2 supply, providing many levels of control over lignin deposition in the Casparian strip (Barbosa et al., 2019). These principles of control may apply in many lignifying tissues.

Lignin monomer radicals combine in a strictly chemical fashion (Tobimatsu and Schuetz, 2019). The lignin monomer radicals transition between different resonance structures that can then combine together in different ways (Ralph et al., 2019). The most common linkage formed is the β-O-4 linkage (Shi et al., 2016), but at least 6 linkages exist (Ralph et al., 2019). The β-O-4 linkage can form by the coupling of monomers or the extension of existing lignin oligomers. Some lignin monomers cannot form the β-O-4 linkage, such as tricin (Lan et al., 2015). These monomers appear to be restricted to existing as end units on lignin chains (Ralph et al., 2019) and may initiate lignin chains (Lan et al., 2015). The amount of monomers, such as *p*-hydroxybenzaldehyde or tricin, that cannot couple at the β-O-4 position may regulate the length of the lignin polymers (Eudes et al., 2012). The conventional monomer, sinapyl alcohol, can only form two linkages, β-β and β-O-4, thus plants that only make sinapyl alcohol have relatively simple lignins (Meyer et al., 1998). Thus, plants could manipulate the supply of different monomers to regulate the average length and linkage structure of lignins.

1.2.4 Cell Wall Proteins

Up to 2000 proteins are found in the plant cell wall (Jamet et al., 2006), though they form a small fraction of the cell wall by weight. Some proteins, like extensins have specific structural roles in the cell wall, and mutants in some extensins are lethal (Lamport et al.,

2011). Other proteins are largely involved in remodelling of the cell wall, and may be catalytic or non-catalytic; these will largely be discussed in Section 1.3.

1.2.4.1 Expansins, and Extensins

Expansins and extensins are both thought to be non-catalytic proteins that have important functions in the primary cell wall, though their structure and functions differ significantly (Cosgrove, 2015; Lamport et al., 2011).

Plant expansins are two-domain proteins (Cosgrove, 2015). There is a CBM domain that can bind cellohexaose and a double-psi β barrel (DPBB) with an unknown function. Though related to family GH45 endoglucanases, the DPBB domain has not yet been shown to have any catalytic function.

Extensins are hydroxy-proline rich glycoproteins (HRGPs) (Lamport et al., 2011). Some serine residues, near hydroxyproline residues, can be α-galactosylated by serine galactosyltransferase 1 from GT96 (Saito et al., 2014). In Arabidopsis, the hydroxyproline residues can be β-arabinosylated by three hydroxyproline O-β-arabinosyltransferases from family GT95 (Ogawa-Ohnishi et al., 2013). These arabinosyl residues can be β-1,2 arabinosylated by GT77 family members (Egelund et al., 2007). Another GT77 enzyme adds a third arabinosyl residue in the same linkage as the second (Gille et al., 2009). A GT47 family member adds an α-1,3 arabinosyl residue to the chain (Møller et al., 2017). Other GTs may be involved in adding another arabinosyl residue.

1.2.4.2 Arabinogalactan Proteins

The AGPs are also HRGPs, but instead of short arabino-oligosaccharides, AGPs are glycosylated with large complex polysaccharides (Lamport et al., 2011). The hydroxyproline residue is O-glycosylated with a β-1,3 galactan backbone with β-1,6 galactan side chains (Tryfona et al., 2012). The side chains can be substituted with α-1,3 arabinosyl residues, β-1,6 glucuronosyl residues and fucosyl residues.

AGPs are synthesised by GTs from multiple families. The β-1,3 galactan backbone is synthesised by enzymes from GT31 (Showalter and Basu, 2016). The O-galactosyltransferase (Basu et al., 2013) and β-1,3 galactosyltransferase (Qu et al., 2008) are encoded by separate GT31 enzymes. The β-1,6 galactosyltransferase that acts

on the β-1,3 main chain may be from family GT29 (Dilokpimol et al., 2014) and this chain is extended by a GT31 enzyme (Geshi et al., 2013). The glucuronic acid substitutions are added to the β-1,6 galactan side chain by GT14 enzymes (Knoch et al., 2013) and methylated by DUF579 methyltransferases (Temple et al., 2019). The fucosyl residues are added by GT37 enzymes (Liang et al., 2013; Tryfona et al., 2014).

1.3   The Properties of Plant Cell Walls and their Molecular Architecture

Primary cell walls undergo phases of extension and stiffening, and thus at different times must exhibit vastly different properties (Burton et al., 2010; Cosgrove, 2018). In addition to this, the interface of living plant tissue with the outside world is often the primary cell wall, and thus the primary cell wall must be somewhat resistant to enzymatic digestion (Bellincampi et al., 2014).

Thickened secondary cell walls are synthesised by relatively few cell types, such as xylem vessels, phloem and fibres, though by weight they make up the majority of plant biomass (Scheller and Ulvskov, 2010). Xylem vessels transport water to the aerial parts of the plant and thus must withstand negative pressures (Reiter, 2002). Along with fibres, the vessels must also support the weight of the plant, which can be up to 3,000,000 kg distributed over a height of above 50 m ("Sequoia Sempervirens (Coast Redwood) Description" n.d.). In addition, secondary cell walls must resist biotic degradation over the lifetime of the plant (Bellincampi et al., 2014).

It is clear that different types of cell wall must exhibit different properties, and even the same cell wall may vary in properties at different times in the lifetime of the plant (Burton et al., 2010). These properties are thought to arise from the molecular architecture of the cell wall. The molecular architecture of the secondary cell wall must contribute to mechanical strength and resistance to enzymatic degradation, while the primary cell wall must be plastic and extensible during growth and development but later mechanically strong enough to resist further expansion. Investigations of cell wall molecular architecture are hampered by the complex and heterogeneous nature of the cell wall. This aspect of cell wall biology remains extremely important for basic research in plant biology as well as sustainable biotechnology (Burton et al., 2010; McCann and Carpita, 2015).

This Section will consider the known molecular architecture of plant cell walls, focussing initially on interactions between polysaccharides, and some glycoproteins, and then on interactions between lignin and polysaccharides.

1.3.1 Cellulose interactions with polysaccharides

Cellulose microfibrils form the core of different cell wall types (Somerville, 2006). Each glucan chain is in a two-fold screw conformation, a flat ribbon where each glucosyl residue is rotated 180° relative to the previous one (Jarvis, 2018). Hemicelluloses share the equatorial geometry of the β-1,4 bonds in cellulose and thus the potential to fold onto cellulose as a two-fold screw (Scheller and Ulvskov, 2010). *In vitro*, xylans, GGMs and xyloglucans affect bacterial cellulose properties, and may bind to microfibrils (Tokoh et al., 2002; Whitney et al., 1998; Hackney et al., 1994).

The elegant structures of xylan have been proposed to enable a specific interaction on the hydrophilic face of cellulose (Busse-Wicher et al., 2014). As previously discussed, the xylans of eudicots and gymnosperms have been shown to have predominantly an even number of xylosyl residues between consecutive substitutions (Bromley et al., 2013; Busse-Wicher et al., 2014; 2016; Martínez-Abad et al., 2017). The even patterning of substitutions means that, when xylans form a two-fold screw, one face of the xylan has only unsubstituted hydroxyls, while the other is highly substituted (Busse-Wicher et al., 2014). The unsubstituted side of the xylan chain mimics a glucan chain in a cellulose microfibril, and could hydrogen bond to the hydrophilic faces of a microfibril (Jarvis, 2018). Randomly substituted xylans could not easily bind to the cellulose hydrophilic face as a two-fold screw due to the steric obstruction of the substitutions, though this is disputed and may depend on the shape of the cellulose microfibril (Martínez-Abad et al., 2017). Such a xylan would stably bind to the hydrophobic faces of cellulose, where even patterning would not be advantageous (Martínez-Abad et al., 2017). Solid-state NMR investigations of Arabidopsis and monocots have demonstrated that two-fold xylan exists in secondary cell walls (Simmons et al., 2016; Kang et al., 2019). The two-fold xylan is within 1 nm of the microfibril surface, and in the absence of cellulose, does not form (Simmons et al., 2016). The *esk1* mutant, which lacks evenly spaced xylan substitutions, has little two-fold xylan (Grantham et al., 2017). Therefore, the *esk1* mutant provides strong evidence for the importance of substitution patterning for cellulose-binding. In addition, *esk1* has collapsed xylem and a dwarf phenotype that can be suppressed by

additional xylem vessels or stem branches (Bensussan et al., 2015; Ramírez et al., 2018). Thus, the folding of xylan as a two-fold screw on cellulose could be important for the mechanical strength of the cell wall. The *esk1* phenotype may be at least partially attributable to the insolubility of this less substituted xylan, which may have effects on the cellulose-binding of xylan. Interestingly the xylan-cellulose interaction appears to be unimportant for restricting enzymatic digestion of the cell wall, *esk1* cell walls are no less recalcitrant than wild type (Lyczakowski et al., 2017).

In addition to xylan, another evenly patterned hemicellulose has been identified. The CSLA2- and MAGT1-syntheised GGM of Arabidopsis mucilage has a repeating backbone pattern of glucose-mannose, where around 60% of the mannoses are galactosylated (Yu et al., 2018). Due to the repeating backbone unit, this GGM has one face that mimics an unsubstituted glucan chain, in a similar fashion to the evenly substituted xylans. Molecular dynamics simulations suggest this GGM may behave similarly to the evenly substituted xylans, and have the ability to bind to the hydrophilic faces of cellulose. However, the axial position of the mannose carbon 2 hydroxyl destabilised interaction on certain cellulose faces. Due to the tissue location of this polysaccharide it was not possible to investigate its interaction with cellulose with solid-state NMR. In *csla2* mutant mucilage, the absence of GGM results in altered cellulose and pectin localisation (Yu et al., 2014). These effects could be due to a lack of GGM binding to cellulose (Yu et al., 2018). The CSLA2-syntheised GGM is notably different to the GGM structure reported in conifers (Hannuksela and Hervé du Penhoat, 2004). The native structures of spruce mannans are reported to bind poorly to cellulose *in vitro* (Hannuksela et al., 2002). Thus, GGM structures may be adapted in different tissues and species for different cellulose binding properties. However, the *in muro* sub-nanometre proximity of non-patterned GGM to cellulose has yet to be investigated with solid-state NMR, and this will partly be the subject of this thesis (chapter 6). Any cellulose-binding properties of GGM could be as important to conifer physiology as the cellulose-binding properties of xylan are to eudicots (Lefebvre et al., 2011). The structure of other CSL-synthesised polysaccharides such as mixed linkage glucan may also allow interaction with cellulose, but their chemical similarity to cellulose makes investigation of this extremely challenging.

The structure of xyloglucan is notably different to xylan and mucilage GGM. Although it is "patterned" in the sense that it has a repeating pattern of substitutions, these substitutions are on consecutive backbone residues (Scheller and Ulvskov, 2010), unlike

xylan. This would suggest that xyloglucan binds to cellulose (if at all) in a different way to xylan and mucilage GGM. Molecular dynamics simulations of xyloglucan tetra-oligosaccharides suggest that XXXG oligosaccharides can bind to the hydrophobic faces of cellulose, but more highly substituted oligosaccharides, such as XLLG, bind less well to cellulose (Zhao et al., 2014). Simulations of longer xylo-oligosacchrides suggest that the substitutions can stack onto the glucosyl backbone if it is in an extended two-fold conformation (Umemura and Yuguchi, 2005). It is unclear how simulations of oligosaccharide behaviour corresponds to *in muro* behaviour. Solid-state NMR investigations suggest there are some infrequent contacts between cellulose and xyloglucan (Dick-Pérez et al., 2011). AFM shows that, in *xxt1 xxt2* mutants, the cellulose microfibrils are more aligned (Xiao et al., 2016), suggesting xyloglucan may regulate cellulose microfibril bundling. FESEM and fluorescence microscopy suggests that xyloglucan binds extensively to cellulose, but the resolution of these techniques is an order of magnitude lower than solid-state NMR (Zheng et al., 2018). Therefore "binding" as detected by fluorescence microscopy is considerably different to "binding" in a solid-state NMR study.

The xyloglucan-cellulose contacts observed by solid-state NMR may be the target site of expansins (Wang et al., 2013). Expansins cause extension of the primary cell wall in response to lowered pH, but their mechanism of action is poorly understood (Cosgrove, 2018). Modulating xyloglucan-cellulose interactions could be an important part of this mechanism. Indeed, in the xyloglucan-deficient mutant *xxt1 xxt2*, the walls are more easily stretched mechanically but biologically they loosen more slowly, suggesting intrinsic cell-wall loosening mechanisms require xyloglucan, possibly through the expansin interaction (Park and Cosgrove, 2012). In maize, solid- state NMR suggests expansins bind highly substituted xylans as part of the mechanism of wall loosening (Wang et al.,, 2016a), which may be functionally equivalent to the xyloglucans in eudicot and gymnosperms cell walls.

Xyloglucans may be modified *in muro* to be covalently linked to each other or cellulose. A class of enzymes, the xyloglucan endotransglucosylase/hydrolase (XTH) proteins can hydrolyse the xyloglucan backbone, some can also transglycosylate the xyloglucan on to other xyloglucan chains, mixed linkage glucans or cellulose (Eklöf and Brumer, 2010; Hrmova et al., 2007). The XTHs can cause extension of primary cell walls, suggesting xyloglucan can act as a load bearing or cross-linking structure (Van Sandt et al., 2007).

In the G layer of tension wood, XTH activity has been suggested to be important for cross-linking the cellulose and xyloglucan rich G layer to the normal secondary cell walls that surround it (Nishikubo et al., 2007). Thus different XTHs may enhance wall extensibility or stiffness. It is unclear if xyloglucan is the only hemicellulose that is transglycosylated *in muro*.

Pectic polysaccharides may also bind to cellulose, though it is not clear how the profoundly different charges and glycosidic bond geometries of cellulose and pectic polysaccharides are compatible with close proximity to each other (Jarvis and Apperley, 1995). Solid-state NMR has suggested that there is close proximity between RG-I, HG and cellulose (Wang et al., 2015). A detailed study correlating mechanical properties with composition and polysaccharide interactions down the elongating stem suggested that the increased mechanical stiffness of the stem correlated with HG demethylesterification and enhanced interaction with cellulose (Phyo et al., 2017). The demethylesterification of pectin may lead to more calcium bridging of separate pectin molecules (Levesque-Tremblay et al., 2015). In addition to calcium based cross-linking of pectins, RG-II can form covalent cross-links through borate esters between apiosyl residues (O'Neill et al., 2001). These two methods of pectin cross-linking may be important for restricting enzymatic digestibility and enhancing wall stiffness (Biswal et al., 2018). The HRGP extensins may also play stiffening roles in primary cell walls (Lamport et al., 2011). They can be covalently cross-linked by extensin peroxidase, and may bind to pectins, though their exact function is not well understood.

1.3.2 Interactions between lignin and polysaccharides

The *in muro* location of lignin polymerisation enables cross-linking between lignin and polysaccharides (Terrett and Dupree, 2019). Such interactions between lignin and polysaccharides may be an important element of cell wall molecular architecture that contributes to recalcitrance to enzymatic digestion. Although some xylan:lignin cross-linking mechanisms are well established, others are more controversial, and the biological function of xylan:lignin cross-linking remains to be fully investigated (Buanafina, 2009; Terrett and Dupree, 2019; Giummarella et al., 2019).

The most well-established example of polysaccharide:lignin cross-linking is ferulate substitutions of xylan acting as lignin monomers (Vanholme et al., 2019), see Figure 1.7.

This means that ferulate substitutions on xylan can form an integral part of the lignin polymer (Ralph et al., 2019). The ferulate substitutions of xylan can undergo radical coupling reactions, just like monolignols, forming xylan:xylan and xylan:lignin cross-links (Ralph et al., 1994; Ralph et al., 1995; Jacquet et al., 1995). In *in vitro* systems, ferulate cross-linking leads to reduced enzymatic digestibility (Grabber et al., 1998). Importantly, monocot mutants that have reductions in ferulate substitutions have significant reductions in recalcitrance, suggesting that lignin:xylan cross-links are an important mechanism for cell wall recalcitrance (Chiniquy et al., 2012; de Souza et al., 2018-
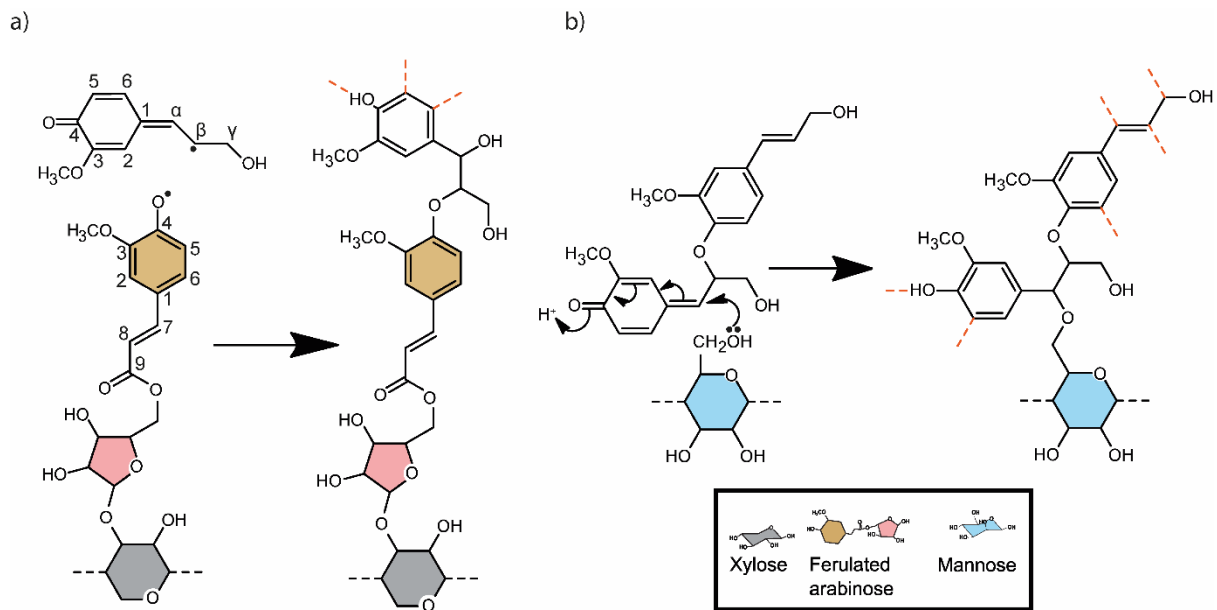
Figure 1.7: Two proposed mechanisms of polysaccharide:lignin cross-linking. Adapted from Terrett and Dupree (2019). a) Radical coupling of a monolignol to ferulated xylan. The backbone xylosyl residue is substituted by α-1,3-linked arabinosyl residue, which is esterified by ferulic acid. A β-O-4 linkage between the ferulate and a monolignol is shown, but other linkages exist. b) Re-aromatisation of the quinone methide intermediate by hemicellulose nucleophiles. The quinone methide intermediate of two monolignols forming a β-O-4 linkage is shown. The intermediate is re-aromatised by the carbon 6 hydroxyl of a mannosyl residue of GGM. A mannosyl residue ether-linked to a dilignol is shown, but other glycosyl residues could form the bond. Curly arrows show the movement of electrons. Further lignin polymerisation can occur at carbons marked by orange dotted lines.

; Jung and Phillips, 2010).These results demonstrate the importance of xylan:lignin cross-linking for limiting enzymatic digestibility of grass cell walls.

In addition to a role in recalcitrance to enzymatic digestion, xylan:lignin cross-linking may be important for non-covalent associations between xylan and lignin. Solid-state NMR shows that there is proximity between xylan and lignin, and some glucan chains in

cellulose (Kang et al., 2019). The cross-peaks indicate that xylan and some lignin are less than 1 nm away from each other in grass cell walls. These non-covalent interactions may be mediated by the covalent ferulate cross-links which may act as an anchor between the two polymers. It is unclear if the lignin stacks hydrophobically on polysaccharides via the aromatic rings (Houtman and Atalla, 1995), or if hydrogen bonding is the primary driver of such non-covalent interactions, though the latter has been suggested in (Kang et al., 2019). Thus, lignin coating of xylan, facilitated by ferulate cross-links to lignin, may be a biologically important aspect of the molecular architecture of grass cell walls. It has not been investigated if covalent cross-linking is important for these non-covalent interactions between xylan and lignin.

Outside of grasses, ferulate substitutions of xylan have not been reported, however polysaccharide nucleophiles can react with lignification intermediates to form covalent cross-linkages (Terrett and Dupree, 2019). During the formation of β-O-4 linkages, a quinone methide intermediate is formed, see Figure 1.7 (Giummarella et al., 2019). To re-aromatise the quinone methide intermediate (i.e. reform the aromatic rings found in the mature lignin polymer) a nucleophile must attack the quinone methide intermediate at the α-position of the β-O-4 linkage (Mottiar et al., 2016). One of the most prevalent nucleophiles during lignification is water. Re-aromatisation by water results in the formation of an α-hydroxyl in the β-O-4 bond (Ralph et al., 2004). However other nucleophiles exist in wood, such as polysaccharide hydroxyls or carboxylate ions, and these could act as nucleophiles in re-aromatisation of quinone intermediates (Terrett and Dupree, 2019). Analysis of such lignin-carbohydrate complexes (LCCs) is extremely difficult due to low amounts of LCCs and the requirement for complex NMR experiments to convincingly verify their existence (Giummarella et al., 2019), thus the existence of many proposed LCC linkages remains controversial. Despite this complexity this method of lignin-polysaccharide cross-linking has been verified in Japanese red pine (Nishimura et al., 2018). Multidimensional NMR experiments of a mannanase-digested sample showed that the carbon 6 of a mannosyl residue in GGM was etherified to the α-position of a β-O-4 linkage.

It remains to be discovered if other polysaccharides can re-aromatise quinone methide intermediates, and if other linkages, such as a lignin-glucuronate ester connecting xylan and lignin, may form, as has been proposed (Das et al., 1984). Xylan has been shown to largely exist in the two-fold screw conformation on the cellulose microfibril (Simmons et

al., 2016). It has been proposed that this occurs on the hydrophilic face of cellulose (Busse-Wicher et al., 2014), and supporting *in vivo* evidence has been found (Grantham et al., 2017). In this model, most xylan backbone hydroxyls are concealed from the matrix due to acetylation on the matrix-face of the xylan chain. Due to its pK_a (Wang et al., 1991), the carboxylic acid of glucuronic acid is likely to exist as a carboxylate ion at the pH of wood (Campbell and Bryant, 1941), and is thus a strong nucleophile. Therefore glucuronic acid substitutions of xylan could act as nucleophiles that re-aromatise some quinone methide intermediates of lignification. This would result in the formation of lignin-glucuronate esters.

There is some indirect evidence for the presence of lignin-glucuronate esters. Esterases from family CE15 have been shown to remove phenolic esters from glucuronic acid *in vitro* (d'Errico et al., 2015). When digesting a lignin and polysaccharide rich extract from wood, such esterases boost the amount of acidic oligosaccharides released by xylanases (Mosbech et al., 2018). CE15 esterases also decrease the molar mass of water-extracted lignin and polysaccharides (Arnling Bååth et al., 2016). CE15 esterases also boost the saccharification of maize lignocellulose (d'Errico et al., 2016; Arnling Bååth et al., 2018). There is not yet convincing solution NMR verification of the existence of these structures (Nishimura et al., 2018). These studies suggest that glucuronic acid may be esterified to lignin *in vivo*, but the specific identification of this linkage and its biological function have not been investigated (Giummarella et al., 2019; Terrett and Dupree, 2019). It was recently reported that glucuronic acid substitutions are crucial for cell wall recalcitrance (Lyczakowski et al., 2017). The greatly reduced recalcitrance of the *gux1 gux2* mutant may suggest that glucuronic acid plays a cross-linking function in the cell wall, but this has not yet been investigated. The role of glucuronic acid substitutions in cross-linking xylan and lignin in Arabidopsis will be investigated in this thesis, facilitated by investigation of the *gux1 gux2* mutant.

Such xylan:lignin cross-linking could be prevented by certain mutations in lignin synthesis, as some lignin monomers are incompatible with this method of cross-linking (Terrett and Dupree, 2019). When caffeoyl alcohol and 5-hydroxyguaicylalcohol form β-O-4, the hydroxyl group on the aromatic ring can act as the nucleophile to re-aromatise the quinone methide intermediate, forming benzodioxane linkages, see Figure 1.8 (Chen et al., 2012; Terrett and Dupree, 2019). Such intramolecular reactions are quicker than intermolecular ones. In COMT-downregulated plants, or tissues with pure caffeoyl alcohol

lignins, the lignins produced are greatly enriched in benzodioxane linkages (Ralph et al., 2001; Chen et al., 2012). <u>Dib</u>enzo<u>diox</u>ocin (DBDOX) linkages may have a similar effect. DBDOX linkages form when a 5-5 linked pair of monolignols then β-O-4 couples to another monolignol (Ralph et al., 2004). The quinone methide is then re-aromatised by the 4-hydroxyl of one of the 5-5 linked monolignols rather than an external nucleophile. This type of linkage appears to be more common in certain mutants, such as the C3H mutant *ref8 med5a/b*, which produces more *p*-coumaroyl alcohol than wild type (Bonawitz et al., 2014; Shi et al., 2016).
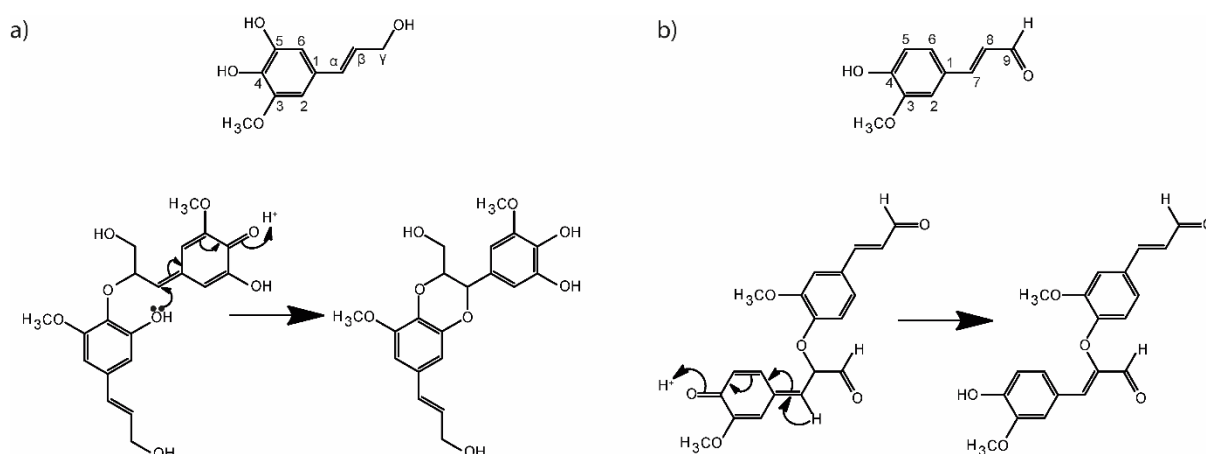
Figure 1.8: Polymerisation of some lignin monomers prevents lignin:polysaccharide cross-linking. Adapted from Terrett and Dupree, (2019). a) The quinone methide intermediate of two 5-hydroxyguaicyl monomers that are β-O-4 coupling is shown. The hydroxyl group on the ring acts as the nucleophile that re-aromatises the quinone methide intermediate. b) The quinone methide intermediate of two coniferaldehyde monomers that are 8-O-4 coupling is shown. The quinone methide intermediate is re-aromatised by elimination of a proton.

Some lignin monomers do not re-aromatise via nucleophilic attack, and thus prevent this method of cross-linking to polysaccharides (Terrett and Dupree, 2019). The hydroxycinnamaldehyde monomers found in *cad* mutants can couple to form 8-O-4 linkages (equivalent to β-O-4 linkages between monolignols) (Zhao et al., 2013). The presence of the γ aldehyde group makes the 7 proton more acidic than in coupling of monolignols (Kim et al., 2003), and thus elimination of this proton is quicker than intermolecular nucleophilic attack, see Figure 1.8. Supporting the functional relevance of these reported mechanisms of lignin:polysaccharide cross-linking, *cad* (Fu et al., 2011), *comt* (Vanholme et al., 2010) and *c3h* mutants (Bonawitz et al., 2014) have been reported to have decreased recalcitrance to enzymatic digestion.

1.4   The use of carbohydrate active enzymes (CaZYmes) in this thesis

As discussed above, the structures of polysaccharides and lignin are complex, containing linkages between many different sugars and lignin monomers, and even between the polymers (Scheller and Ulvskov, 2010; Ralph et al., 2019; Terrett and Dupree, 2019). Yet as the largest fraction of global carbon (Bar-On et al., 2018), lignocellulose is a major energy reserve utilised for energy by many organisms, especially fungi and bacteria (Andlar et al., 2018; Ndeh and Gilbert, 2018). The depolymerisation of the cell wall requires glycosyl hydrolases (GH), carbohydrate esterases (CE), peroxidases and many auxiliary activity (AA) families, which includes the lytic polysaccharide monooxygenases (LPMOs) (Andlar et al., 2018). Enzymes are classified into different families in the CaZY database based on sequence analysis (Coutinho et al., 2003). The structural complexity of cell wall components is matched or exceeded by the number of depolymerisation enzymes (Ndeh et al., 2017). In this Section, the application of such enzymes to analyse polysaccharide structures and conduct saccharification experiments will be briefly described.

1.4.1   Glycosyl hydrolases and carbohydrate esterases

Glycosyl hydrolases and carbohydrate esterases have been crucial tools to analyse the structure of polysaccharides (Scheller and Ulvskov, 2010). The products of hydrolysis can be separated and visualised by techniques such as polysaccharide analysis by carbohydrate electrophoresis (PACE), DNA-sequencer assisted saccharide analysis in high-throughput (DASH) (Li et al., 2013; Tryfona et al., 2019) or mass spectrometry (MS) (Tryfona et al., 2012) or chromatographic techniques (Šuchová et al., 2018).

In the case of xylan, most structural studies have been performed with endo-xylanases from families GH10, 11 and 30. Bacterial xylanases from subfamily GH30_8 recognise glucuronic acid substitutions, and cleave the xylan backbone with the substitution in the -2 position (Urbániková et al., 2011). The substrate specificity of these enzymes enabled the identification of glucuronic acid substitutions with even numbers of xylosyl residues between consecutive substitutions (Bromley et al., 2013; Busse-Wicher et al., 2016). GH11 xylanases can cut the xylan backbone when there are three consecutive unsubstituted xylosyl residues, with one residue in the +1 subsite (Vardakou et al., 2008). GH10 xylanases can accommodate substitutions in the +1 subsite, but require two

unsubstituted residues in the -1 and -2 position (a 3-linked substitution in the -2 position is tolerated) (Kolenová et al., 2006). These highly active xylanases are useful to profile substitution patterns, especially complex xylans (Tryfona et al., 2019).

The identity of xylan substitutions can be probed with GHs and CEs. Arabinosyl substitutions are removed from the xylan backbone by enzymes from families GH43, GH51 and GH62 (Rogowski et al., 2015; Sakamoto et al., 2011). GH62 enzymes can only remove arabinosyl substitutions from singly substituted xylosyl residues (Kellett et al., 1990), while some GH43 enzymes remove arabinose from doubly substituted xylosyl residues (Rogowski et al., 2015). Glucuronic acid substitutions can be removed by enzymes from GH67 and GH115 (Rogowski et al., 2015; Martínez et al., 2016; McKee et al., 2016). GH67 glucuronidases remove glucuronic acid from terminal xylosyl residues only, while GH115 glucuronidases have no such specificity (Rogowski et al., 2015). Acetyl esterases remove acetate groups the backbone in a regiospecific manner (Neumüller et al., 2015).

1.4.2   Enzymatic saccharification

Cell wall recalcitrance to enzymatic digestion is primarily assayed by applying saccharification enzyme mixtures to cell wall material and analysing the weight lost, or amount of saccharides released by the enzymatic mixture (McCann and Carpita, 2015; Holwerda et al., 2019).

In this thesis the saccharification mixture used was Cellic® CTec2, produced by Novozymes. This enzyme mixture has been reported to contain glucanase, xylanases, glucosidase and xylosidase activity (Yarbrough et al., 2015). The mixture has also been reported to contain glucuronoyl esterase activity (d'Errico et al., 2016). In comparison to previous commercial enzyme mixtures, CTec2 contains AA9 LPMOs (Novozymes, 2017). These enzymes oxidise polysaccharide substrates, and are believed to be important for generating new target sites for glycosyl hydrolases (Quinlan et al., 2011). Beyond what has been reported in publications, the exact enzyme families, and the ratios of each enzyme in CTec2 are a commercial secret and will not be discussed further.

1.5 Research Aims

As has been discussed, the structure of cell wall polysaccharides is well understood (Scheller and Ulvskov, 2010). Many of the genes that contribute to polysaccharide synthesis have been identified, see Section 1.2, though significant gaps remain in our understanding. As discussed in Section 1.3, one aspect of cell wall biology that is relatively poorly understood is molecular architecture; how the different cell wall components interact with each other *in muro* and how this leads to higher order properties such as mechanical strength and recalcitrance to enzymatic digestion (Burton et al., 2010). Understanding the molecular architecture of the cell wall is of crucial importance for designing new strategies to chemically or genetically modify plant biomass for better utilisation in a sustainable economy (McCann and Carpita, 2015).

This study was conceived with the objective of investigating three key aspects of cell wall molecular architecture:

- What is the molecular architecture function of glucuronic acid substitutions on xylan? The *gux1 gux2* mutant has a severe reduction in cell wall recalcitrance to enzymatic digestion. Is this related to a change in the molecular architecture of the cell wall? In particular, can this phenotype be explained by the previously proposed lignin-glucuronate ester? These questions are investigated in chapters 3 and 4;
- The ferulate substitutions of monocot xylan are an important xylan substitution for cross-linking xylan and lignin. The synthesis of these substitutions is complex and incompletely understood. Can synthetic biology be used to test our model of ferulate susbstitution synthesis? This is investigated in chapter 5;
- What are the molecular architecture functions of the different hemicelluloses in conifers? Can solid-state NMR be used to investigate the cellulose and lignin-binding properties of xylan and galactoglucomannan in conifers? These questions are investigated in chapter 6 of this thesis.

Chapter 2: Materials and Methods

2.1 Plant Material

2.1.1 Plant lines

All *Arabidopsis thaliana* plants used were of the Columbia-0 ecotype. The *cad2-1 cad6-1* mutants were received from Dr Ruben Vanholme, but their characterisation is currently unpublished. *cad2-1* is the SAIL_1265_A06 insertional line of AT3g19450 and *cad6-1* is the SALK_040062 insertional line of AT4g34230. The *lac4-2* and *4cl1-1* mutants were received from Professor Wout Boerjan, described in (Van Acker et al., 2013). The *ref8 med5a/b* mutant was received from Professor Clint Chapple, described in (Bonawitz et al., 2014). The *gux1-2 gux2-1* was described in (Mortimer et al., 2010) and received from Dr Marta Busse-Wicher. The *esk1-5 kak-8* mutant was provided by Dr Nick J Grantham and described in (Bensussan et al., 2015). The *atr2* mutant was the SALK_026053 insertional line of AT4g30210, purchased from NASC.

2.1.2 Seed preparation and growth on soil

For soil growth, seeds were put in cold water in Eppendorf tubes, wrapped in foil and stratified for 48 hours at 4ºC. Seeds were then sown onto soil composed of Levingtons M3 compost and vermiculite. Plants were grown in a 16 hour light and 8 hour dark cycle at 60% humidity, 21ºC, 100 µmol m^{-2} s^{-1}. The plants were grown for 6-8 weeks and watered regularly when required with de-ionised water.

For hydroponic growth the seeds were sterilised in 80% Ethanol (1% Tween-20 and 1% bleach) for 5 minutes. They were then washed with 65%, 80% and 100% ethanol and dried in a laminar flow hood. The sterile seeds were sown onto 0.8% agar plates with ½ strength Murashige and Skoog salts (Sigma Aldrich) and 1% sucrose, adjusted to pH 5.8. The plates were wrapped in foil and stratified for 48 hours at 4ºC. The plants were then grown on plates for two weeks in the conditions above.

2.1.3 Hydroponic growth

The two week old plantlets were transplanted into foil-covered rockwool that had been leached of nutrients in distilled water. The plants were grown in the light and temperature conditions above. Plants were watered up to twice a week with hydroponic solution until they were 6-8 weeks old, and the first siliques had turned brown. The solution was prepared to the final concentrations shown in table 2.2. The stock solutions were produced with Sigma Aldrich reagents and prepared in MilliQ water.

Salt	Concentration (mM)
$MgSO_4$	2
$Ca(NO_3)$	1.0/0.5/0.2
FeEDTA	0.05
KNO_3	5
$K_2HPO_4+KH_2PO_4$ pH=5.5	2.5
H_3BO_3	0.07
$MnCl_2$	0.014
$CuCo_4$	0.0005
$ZnSO_4$	0.001
$NaMoO_4$	0.0002
NaCl	0.01
$CoCl_2$	0.0001

Table 2.1 Concentrations of salts in hydroponic growth medium.

2.1.4 AIR and dried biomass wet milling

The basal fifth of the stem (approximately 6cm) was harvested when the first siliques started to brown and used for AIR preparation. The stems were boiled at 70°C in 99.8% ethanol for half an hour. The stems were then ball-milled in the ethanol for 10 minute periods, at 25 rpm in a Retsch MM 400 until a fine powder. The pellet was spun in a 50 ml Falcon tube, the supernatant was disposed of and washed with 99.8% ethanol. The spinning was repeated and the supernatant was disposed of for each of the following washes. The washes were a 2 hours 2:3 MeOH/CH_3Cl wash, then an overnight wash with the same mixture. Then the pellet was washed with 99.8%, 65%, 80% and 99.8% ethanol successively. The pellet was then dried at 50°C.

For the ball-milled dried biomass, the same stems were dried for two weeks at room temperature. 8 mg of dried stem was ball-milled in 8 ml of 0.1 M ammonium acetate pH 5.0 buffer for three 10 minute periods, at 25 rpm in a Retsch MM 400, with a cooling period of 10 minutes between each milling period.

2.2 Saemann hydrolysis

0.5 mg aliquots of ball-milled dried biomass were hydrolysed at room temperature for 3 hours with, 125 µl of 72% H_2SO_4 (Sigma Aldrich, prepared gravimetrically with MilliQ water). Samples were then spiked with 10 mmol myo-inositol and then diluted with MilliQ water to a total volume of 1.5 ml, and hydrolysed for 2.5 hours at 100°C. The samples were then cooled on ice. Half the sample was put into a 15 ml Falcon tube and diluted with an equal volume of MilliQ water. The acid was neutralised with barium carbonate and the pH was tested with pH paper. The samples were spun for 10 minutes at 3000X g, and the supernatant was then used for quantification of monosaccharides by HPAEC-PAD, using a CarboPac PA20, as described in (Currie and Perry 2006). Standards were run to enable calculation of response curves for each biological replicate.

2.3 Enzyme preparation and hydrolysis

2.3.1 Desalting and Dilution of CTec2 saccharification mixture

A Sephadex 25 PD-10 column was used as per the manufacturers instructions, to remove excess salt and mono/oligosaccharides from the Cellic® CTec2 (Novozymes) mixture. Hereafter the product will be referred to as CTec2. Column buffer was exchanged using 25 ml 0.1 M ammonium acetate pH 5.0 buffer. 350 µl of CTec2 was added to 2.15 ml of 0.1 M ammonium acetate pH 5.0 buffer. The 2.5 ml volume was passed through the PD-10 column and eluted with 3.5 ml of 0.1 M ammonium acetate pH 5.0 buffer. The desalted 1:10 CTec2 solution was then stored at 4°C for no longer than 1 month.

2.3.2 CTec2 Saccharification

20 µl of the 1:10 CTec2 dilution was added to 0.5 ml aliquots of the 1 mg ml^{-1} ball-milled dried biomass in 1.5 ml Eppendorf tubes. The tubes were placed in an Eppendorf Thermomixer at 45°C for 24 hours. Tubes were shaken for 30 s and left to settle for 4

minutes in a repeating cycle for this 24 hour period. The tubes were vortexed and then spun at 15,000 rpm in a 4°C centrifuge. The concentrations of glucose and xylose in the supernatant were measured in cuvettes using the Megazyme D-Glucose Assay Kit and Megazymes D-Xylose Assay Kit (Megazyme catalogue codes K-GLUHK-220A and K-XYLOSE respectively), as per the manufacturers instructions. Negative controls of CTec2 alone, and non-saccharified biomass were used to adjust the readings for background monosaccharides from either source. For each assay a standard curve with pure xylose or glucose solutions was produced.

2.3.3 Purified enzymes and enzymatic hydrolysis

The purified recombinant enzymes were taken from different sources. See Table 2.2 for organism the enzymes originated from, and the source of those enzymes.

Name	CaZY family	Organism	Source	Concentration (mg ml^{-1})
GH11A β-1,4 xylanase	GH11	*Neocallimastix patriciarum*	Prof Harry Gilbert	5.4
β-1,4 xylanase	GH10	*Aspergillus aculeatus*	Novozymes	0.5
α-glucuronidase	GH115	Unknown	Novozymes	17.8
Acetyl xylan esterase	CE4	*Cellvibrio japonicus*	nzytech	0.3
α-arabinofuranosidase	GH62	*Penicillum aurantiogriseum*	Novozymes	3.8
Endoglucanase	GH45	*Humicola insolens*	Novozymes	2.3
Cellobiohydrolase	GH7	*Trichoderma reesei*	Novozymes	7.1
β-glucosidase	GH1	*Trichoderma harzianum*	Dr Clelton Dos Santos	1.0

Table 2.2 CaZYmes used in this thesis.

2.3.3.1 Hydrolysis for PACE in chapter 5

Xylan was alkali-extracted from 1 mg AIR for 1 hour with 20 μl 4 M NaOH. This was neutralised with 80 μl of 1 M HCl. The neutralised sample was suspended in 1 ml pH 5.5 0.1 M ammonium acetate. The xylan was digested with 6 μl of GH11 β-1,4 xylanase and 3 μl GH62 α-arabinofuranosidase overnight at room temperature. The samples were then analysed using PACE (2.3).

For the mild TFA experiment, 1 mg AIR was suspended in 400 μl of 0.05 M TFA. It was then hydrolysed at 100°C for two hours. The tubes were vortexed and then spun at 15,000 rpm in a 4°C centrifuge. The supernatant was removed and then the pellet was

washed with water. The tubes were vortexed and then spun at 15,000 rpm in a 4°C centrifuge. The wash supernatant was added to the supernatant and dried overnight before being analysed using PACE.

2.3.3.2 Enzyme hydrolysis for PACE and solution NMR in chapter 4

500 mg of AIR was suspended in 25 ml 0.5% ammonium oxalate and depectinated by heating to 85°C for 2 hours. The mixture was spun at 3700 rpm for 10 minutes and the supernatant disposed of. The pellet was washed in 50ml MilliQ water. To delignify the AIR, 40ml of 11% peracetic acid was added to the pellet and it was shaken at 750 rpm at 85°C for 30 minutes. The mixture was spun at 3700 rpm for 10 minutes and the supernatant disposed of. This process was repeated with 50 ml washes of acetone, and MilliQ water until the supernatant was pH 7. The final wash was with 100% ethanol. To extract xylan and associated lignin, 25 ml of DMSO was added and the samples were heated to 80°C for 24 hours. The mixture was spun at 5000 rpm for 10 minutes and the supernatant was reserved. The DMSO extraction was repeated and the supernatants were pooled. The extracted polysaccharides were then exchanged in 20 mM pH 7.0 ammonium acetate buffer using PD-10 columns. 3.5 ml of this was digested with 15 μl of acetyl xylan esterase and 50 μl of AaGH10. The sample was digested for 24 hours at 45°C, with 750 rpm shaking. The sample was then purified by solid-phase extraction using a C-18 column. Using syringes to apply the solvents, the column (Sep-Pak C18 Classic Cartridge WAT051910) was washed in 10 ml of 100% ethanol, followed by 20 ml MilliQ water and 2 ml of 100% acetonitrile. The column was then equilibrated with 10 ml of 20 mM pH 7.0 ammonium acetate buffer. Then 3.5 ml of sample was loaded onto the column. The column was then washed with 10 ml MilliQ water, 5 ml 10% ethanol, 5 ml of 20% ethanol and 10 ml of 100% ethanol; each fraction was collected in separate 15 ml Falcon tubes. The samples were then analysed by PACE (2.4) or solution NMR (2.6).

2.3.3.3 Xylanase and glucanase accessibility

0.5 mg of ball-milled dried biomass was suspended in 0.5 ml pH 6.0 0.1 M ammonium actetate buffer. 10 μl of AaGH10, 6 μl of acetyl xylan esterase and 10 μl GH115 were added to the solution to assess xylan accessibility. 10 μl GH45 endoglucanase, 10 μl GH7 cellobiohydrolase and 5 μl β-glucosidase were added to assess cellulose

accessibility. The samples were incubated overnight at 45°C for 24 hours. The tubes were vortexed and then spun at 15,000 rpm in a 4°C centrifuge for ten minutes. To separate the oligosaccharides from the rest of the biomass, the supernatant was removed and placed in a new tube. 0.5 ml pH 5.5 0.1 M ammonium acetate buffer was used to wash the pellet. The samples were re-vortexed and the wash supernatant was added to the digest supernatant. 20 µl of the 1:10 CTec2 dilution was then added to the 1 ml supernatant volume. The tubes were placed in an Eppendorf Thermomixer at 45°C for 24 hours. Tubes were shaken for 30 s and left to settle for 4 minutes in a repeating cycle for this 24 hour period. The tubes were vortexed and then spun at 15,000 rpm in a 4°C centrifuge. The concentration of glucose and xylose in the supernatant was measured in cuvettes using the Megazymes D-Glucose Assay Kit and Megazymes D-Xylose Assay Kit, as per instructions.

2.4   PACE

Hydrolysed samples were vacuum dried and fluorescently derivatised with 5 µl 0.2 M 8-aminonaphthalene-1,3,6-trisulphonic acid (ANTS) in 10 µl of a buffer containing 20:3:17 of DMSO, acetic acid and water and 5 µl of 0.2 M picoline borane overnight at 37ºC. The samples were vacuum dried and resuspended in 3 M urea. 2.5 µl of each sample was loaded into the PACE gel wells. The PACE gels were composed of a 10% polyacrylamide gel in a 0.1 M Tris-hydroxymethyl aminomethane borate pH 8.2 buffer. The gels were 24 cm wide by 0.025 cm deep. Samples were electrophoresed at 1000 V for 45 minutes in a Hoefer SE 660 series vertical gel, cooled to 10ºC.

The gels were imaged using a G-box camera with transillumination at wavelength of 365 mm.

2.5   Genetic transformation of Arabidopsis

Genetic constructs were made using the Goldengate modular cloning method, the specific constructs made are shown in the relevant chapter (Patron et al., 2015). In Goldengate cloning, L0 parts, such as promoters, CDS, terminators and tags are assembled into L1 transcriptional units. The L1 transcriptional units are assembled into L2 binary vectors which can be transformed into *Agrobacterium tumefaciens* to enable plant transformation. The Goldengate method uses type IIS restriction enzymes that cut

DNA leaving four basepair overhangs a set distance from their recognition site (Szybalski et al., 1991). Careful design of the overhangs enables specific and ordered assembly of the cut DNA parts into the correct order.

The DNA parts, restriction enzyme and T4 DNA ligase were included in a one pot reaction mixture. Cycles of restriction and ligation enable production of the construct. Specifically, the mixture was incubated in a thermocycler with the following program 25 cycles of 37 °C for 3 minutes (restriction) followed by 16 °C for 4 minutes (ligation), followed by enzyme denaturation at 50 °C, then 80 °C for five minutes each.

The mixture was then transformed into NEB DH5α *E.coli* competent cells using a standard heat shock protocol (Froger and Hall, 2007). The cells were plated onto LB agar plates with the relevant selection components (ampicillin, IPTG, X-Gal for L1 assemblies, kanamycin for L2 assemblies). The plates were incubated at 37°C overnight. Cultures were grown from single colonies and plasmids were recovered using Qiagen Miniprep kits.

The L2 plasmid was used to transform GV3101 *Agrobacterium tumefaciens* using a standard heat shock protocol (Cui et al., 1995). The cells were plated on LB agar containing kanamycin and gentamycin and incubated at 30°C for two days.

To transform 4 week old Arabidopsis, the floral dip method was used (Clough and Bent, 1998). Single GV3101 colonies were picked and grown in an overnight LB culture with the relevant antibiotics. The overnight culture was used to inoculate a larger culture which was grown to an OD_{600} of approximately 0.6. The cells were pelleted at 3000 rpm and resuspended in the dipping solution (5% sucrose, 10 mM $MgCl_2$, 0.001% Silwett® L-77). The flowers were submerged in the dipping solution for 1 minute. The dipped plants were bagged for two days. The bags were removed and the plants were left to grow and senesce. Transformed seeds were selected by fluorescence from the OLEOSIN-GFP fusion protein (Shimada et al., 2010).

2.6 Solution-state NMR

Solution state NMR analysis was performed on the 20% ethanol fraction of digested LCC in chapter 4.

Experiments were performed on a Bruker DRX600 spectrometer. The chemical shifts were measured relative to acetone (δ^1H = 2.225 ppm, $\delta^{13}C$ = 31.07 ppm). The [^1H, ^{13}C] HSQC and 2D and 3D HSQC-TOCSY experiments were performed according to established protocols (Cavanagh et al., 2007). The ^1H mixing time of the HSQC-TOCSY experiments was 100 ms. The data were processed with AZARA 2.7 (copyright 1993-2019, Wayne Boucher and Department of Biochemistry, University of Cambridge), and chemical shift assignment was performed with ANALYSIS 2.4.2 (Collaborative Computational Project for NMR Analysis, CCPN) (Vranken et al., 2005).

2.7   Solid-state NMR

^{13}C-enriched spruce and Canadian poplar were purchased from IsoLife.

To enrich the Arabidopsis plants, they were prepared and grown as described in Section 2.1.2-3. The plants in rockwool were placed in a sealed chamber with a clear lid (Simmons et al., 2016). The air was scrubbed of $^{12}CO_2$ and re-mixed with $^{13}CO_2$ in a ratio of 2000:1. The plantlets were then grown in this atmosphere with the same light levels as in Section 2.1.2. The temperature was 25°C. The plants were watered weekly with the hydroponics medium described in 2.1.3 with a Ca^{2+} concentration of 1 mM. The plants were harvested after 6 weeks of growth. The basal 3 cm of stem were harvested and frozen in liquid nitrogen. They were then stored at -80°C until the experiments were run.

35-50 mg basal stems were loaded into a 3.2 mm MAS probe. The CP-INADEQUATE and CP-PDSD experiments were performed on a Bruker (Karlsruhe, Germany) 850 MHz Advance III solid-state NMR spectrometer by Prof Ray Dupree at the University of Warwick. The MAS spinning speed was 12.5 kHz and the sample temperature was between 25-30°C. The parameters of the described experiments were similar to (Simmons et al., 2016; Grantham et al., 2017).

The ^{13}C chemical shift was determined using the carbonyl peak at 177.8 ppm of L-alanine as an external reference with respect to tetramethylsilane (TMS). The spectra were obtained by Fourier transformation into 4 K (F_2) × 2 K (F_1) points with exponential line broadening of 50 Hz in F_2 and 90°-shifted squared sine bell processing in F_1. The data was processed in TopSpin 4.0.

2.8   Confocal Microscopy

Stems of four week old plants were cut at the first internode. The stems were hand-sectioned, stem slices between 5-10 µm thick were placed on a microscope slide, wetted and covered with a cover slip. The sections were imaged by confocal laser scanning microscopy using a Leica TCS SP8 confocal microscope. For the eGFP, the excitation laser had a wavelength of 488 nm and transmission occurred in the range 500-520 nm. For mCherry the excitation laser wavelength was 587 nm and the transmission occurred in the range 600-620 nm. Each channel was imaged sequentially to reduce background signal. The stacks were examined in ImageJ and the brightest slice for each genotype was selected and converted to a TIFF file. The TIFF files were contrast enhanced in Adobe Photoshop, with the same settings for each genotype.

2.9   Thermoporometry by differential scanning calorimetry

These experiments were performed by Dr Carlos E. Driemeier. The ball-milled dried biomass sample were conditioned for 24 hours in de-ionised water. The sample was transferred to aluminium pans and sealed. The mass of the empty and filled pan was determined gravimetrically. At the end of the experiment the samples were dried at 105°C and their mass was determined gravimetrically. These values were used to determine the dry sample mass.

The sample was frozen to -70ºC. The heat flow required to heat the sample to successive isotherms of −60.0, −50.0, −40.0, −30.0, −20.0, −15.0, −10.0, −6.0, −4.0, −2.0, −1.5, −1.1, −0.8, −0.5, −0.2, −0.1, 0.5, and 5.0 °C was measured. For each temperature transition the sample was heated at 1ºC min^{-1} (Driemeier et al., 2012).

The theoretical principle of the experiment is that the melting temperature of ice is suppressed in small pores, and the temperature at which ice in a porous material melts can be used to infer the amount of ice in pores of different sizes.

The temperature suppression (*ΔT*) is related to the ice water interface area (*S*) and ice volume (V) by the equation:

$$\Delta T = -K_c \frac{\delta S}{\delta V}$$

K_c is a "simple function of enthalpy and temperature of the bulk phase transition, liquid density, and ice-water interface energy" (Driemeier et al., 2012).

The total pore area (A) at a given isotherm is calculated from:

$$A = \int \delta M \frac{-\Delta T}{K_c \cdot \rho_{ice}}$$

Where M is the mass of ice, and ρ_{ice} is the ice density.

The pore diameter (x) can be calculated using the following equation, which assumes the pores are cylindrical:

$$x = -2 \cdot \frac{K_c}{\Delta T}$$

The heat flow data was processed in an Excel document designed by Dr Carlos E Driemeier, using the theoretical equations above, to produce the data shown in chapter 3.

2.10 Western Blotting

The section of stem between the second and third internode was harvested from four week old T_3 homozygous plants, with matching plants of the same genetic background as negative controls. The stem sections were ground with a plastic pestle in Lamelly buffer, containing DTT and mercaptoethanol to act as reducing agents. For the high efficiency extraction stems were flash frozen in liquid nitrogen and ball-milled for 1 minute, the smashed stems were suspended in Lamelly buffer containing DTT and mercaptoethanol. The extracted proteins were denatured at 70°C for 10 minutes. The samples were loaded into a pre-prepared 12% acrylamide Mini-Protean® TGX Stain-Free™ gel and separated by gel electrophoresis at 200V for 40 minutes. The gels were Stain-Free imaged using a BioRad ChemiDoc imaging system. The proteins were transferred to a nitrocellulose membrane with an iBlot system. The membranes were blocked overnight in 5% semi-skimmed milk in 1X TBS. They were then incubated with a primary antibody for 2 hours, 1:2500 rabbit anti-Myc, (A-14, Santa-Cruz Biotech) 1:5000 mouse anti-FLAG (M2, Sigma-Aldrich), 1:5000 rabbit anti-HA (Y-11, Santa-Cruz Biotech) or 1:2500 rabbit anti-GFP (ab6556, Abcam). The blots were then washed with 1X TBST for 10 minutes, three times. The membrane was then blocked again in 5% semi-skimmed

milk in 1X TBS. The secondary antibody, horseradish peroxidase-linked (HRP-linked), against the relevant species was then incubated for 1 hour, and the excess was removed with 5 washes of 1X TBS for 5 minutes. The blot was then developed with Amersham ECL prime HRP substrate (GE-Lifesciences), and imaged.

2.11 Thesis Preparation and Statistical Analysis

The thesis was written in Microsoft Word. Figures were produced in Microsoft Excel, ImageJ, Adobe Illustrator and Photoshop. Statistical analysis was performed in R, the exact statistical tests performed are stated in Figure legends.

Chapter 3: Investigating the recalcitrance phenotype of the *gux* mutants

3.1 Introduction

The majority of plant cells are surrounded by a cell wall composed of polysaccharides, polyphenolics and some proteins (Scheller and Ulvskov, 2010). The secondary cell walls that make up the highest proportion of global plant biomass are often referred to as lignocellulose (Bar-On et al., 2018). Cell walls control the growth and development of plant cells (Burton et al., 2010), and act as a defensive barrier to pathogens, by preventing entry of the pathogen into the cell (Bellincampi et al., 2014). Cell walls contribute to disease resistance by being recalcitrant to enzymatic digestion, and by this same property prevent economic utilisation of lignocellulosic materials for biofuels (McCann and Carpita, 2015). The molecular basis of cell wall recalcitrance is not well understood but is thought to relate to cell wall composition and cross-linking (Carmona et al., 2014; McCann and Carpita, 2015; Terrett and Dupree, 2019). Understanding and modulating the recalcitrance of plant cell walls could be important for improving resistance to plant pathogens and improving the economic viability of second generation biofuels (Pauly and Keegstra, 2008).

Glucuronic acid substitutions of xylan are very important for lignocellulose recalcitrance. The Arabidopsis *gux1 gux2* mutant lacks glucuronic acid substitutions on secondary cell wall xylan (Mortimer et al., 2010). The *gux1 gux2* mutant, when digested with the commercial saccharification enzyme mixture CTec2, releases 7 x more Xyl, primarily from xylan, and 1.5 x more Glc, primarily from cellulose, demonstrating a significant drop in the recalcitrance of the cell wall (Lyczakowski et al., 2017). This is a sizable decrease in recalcitrance, especially compared to other Arabidopsis xylan branching mutants which have negligible effects on saccharification (Xiong et al., 2013; Bensussan et al., 2015; Lyczakowski et al., 2017). There are some lignin mutants with even larger drops in recalcitrance, though they often lead to dramatic dwarfing not seen in *gux1 gux2* (Van Acker et al., 2013; Meester et al., 2018). There could be many mechanisms by which xylan glucuronic acid substitutions affect cell wall recalcitrance, as there are multiple ways in which lignocellulose prevents or inhibits enzymatic digestion.

The complexity of lignocellulose is one mode of recalcitrance. Lignocellulose is primarily degraded by glycosyl hydrolases (GHs) and lytic polysaccharide mono-oxygenases

(LPMOs) (Andlar et al., 2018). The structural complexity of polysaccharides increases cell wall recalcitrance, as many different GH families are required to depolymerise complex polysaccharides (Ndeh et al., 2017). Undigested polysaccharides or lignin in lignocellulose can prevent digestion of the whole composite (Vermaas et al., 2015). Adding additional enzyme families to saccharification mixtures can reduce recalcitrance by digesting polysaccharides that previously were left intact (Vaaje-Kolstad et al., 2010; von Freiesleben et al., 2018). In the case of the *gux1 gux2* mutant, the addition of GH115 does not affect saccharification yields (Lyczakowski et al., 2017), suggesting there is not a lack of glucuronidase activity.

The lignocellulose, or breakdown products of lignocellulose can also inhibit enzyme activity. The products of some hydrolytic enzymes can act as inhibitors for others, for instance xylo-oligosacchrides produced by xylanases can inhibit cellulases and xylanases (Xue et al., 2015; Qing and Wyman, 2011). Lignin-derived chemicals such as vanillin can reduce enzymatic activity, partially by precipitating hydrolytic proteins (Qin et al., 2016). In addition, hydrolytic enzymes, especially high molecular weight β-glucosidases, can non-productively bind to the lignin polymer, preventing polysaccharide breakdown (Yarbrough et al., 2015).

Morphological characteristics of lignocellulose can affect recalcitrance as well. The surface area of lignocellulose particles may affect the rate of digestion (Bragatto et al., 2012). A larger particle size could cause increased recalcitrance and may reflect an increased resistance to mechanical breakdown. A more significant morphological characteristic that affects recalcitrance is the porosity of the cell wall (Meng and Ragauskas, 2014; Tavares et al., 2015; Bragatto et al., 2012). The porosity of the cell wall is thought to enable or restrict enzyme penetration into the cell wall (Tavares et al., 2015). Pores that are less than 10nm in diameter are thought to restrict enzyme access, while larger, or very well hydrated pores enable enzyme penetration into lignocellulose and facilitate digestion (Meng and Ragauskas, 2014; Pihlajaniemi et al., 2016; Rondeau-Mouro et al., 2008). Increased porosity of cell walls both before and after pretreatments correlates well with enzymatic digestibility (Crowe et al., 2017; Herbaut et al., 2018; Herbaut et al., 2018).

There are therefore many factors that must be considered when investigating how glucuronic acid substitutions of xylan strongly affect the recalcitrance of the secondary

cell wall. The aim of this chapter is to understand the origin of the significant reduction in biomass recalcitrance in the *gux1 gux2* mutant. The specific aims are:

- To investigate the role of enzyme inhibition in the recalcitrance phenotype of *gux* mutants
- To investigate the role of lignocellulose morphological characteristics in the recalcitrance phenotype of the *gux* mutants.

3.2   Results

3.2.1 Enzyme inhibition and non-productive binding are not important factors in the recalcitrance differences between wild type and *gux1 gux2*

Previously, enzyme inhibition by xylo-oligosaccharides has been proposed to play an important role in recalcitrance by inhibiting the enzymes of saccharification mixtures (Xue et al., 2015). It could be the case that glucuronic acid-containing xylo-oligosaccharides accumulate in the wild type during saccharification, and that these oligosaccharides prevent further saccharification by inhibiting CTec2 enzymes.

If enzyme inhibition occurs in wild type, it would be expected that the recalcitrance of the cell wall varies with the amount of glucuronic acid substitutions on xylan. The single *gux* mutants, *gux1* and *gux2*, have 30% and 70% of wild type xylan glucuronidation, respectively (Mortimer et al., 2010). To test if cell wall recalcitrance is proportional to the degree of xylan glucuronidation, wild type, the single *gux* mutants and the *gux1 gux2* double mutant were saccharified for 24 hours using CTec2, as previously described in (Lyczakowski et al., 2017) and Chapter 2. The Xyl and Glc released after saccharification of wild type, *gux1*, *gux2*, and *gux1 gux2* were measured using the appropriate Megazyme kits, the data are shown in Figure 3.1a. The mean sugar released from wild type was 21 µg Xyl mg^{-1} dried biomass and 129 µg Glc mg^{-1} dried biomass, while from *gux1 gux2* the mean sugar released was 4.2 x the amount of Xyl and 1.2 x the amount of Glc, similar to previously reported results (Lyczakowski et al., 2017). The single gux mutants released similar amounts of Xyl and Glc to wild type. One way ANOVA analysis showed that the small differences between wild type and the single *gux* mutants were not statistically significant. The differences in Xyl-release between wild type and *gux1 gux2* were significant, but the differences in Glc-release were not.
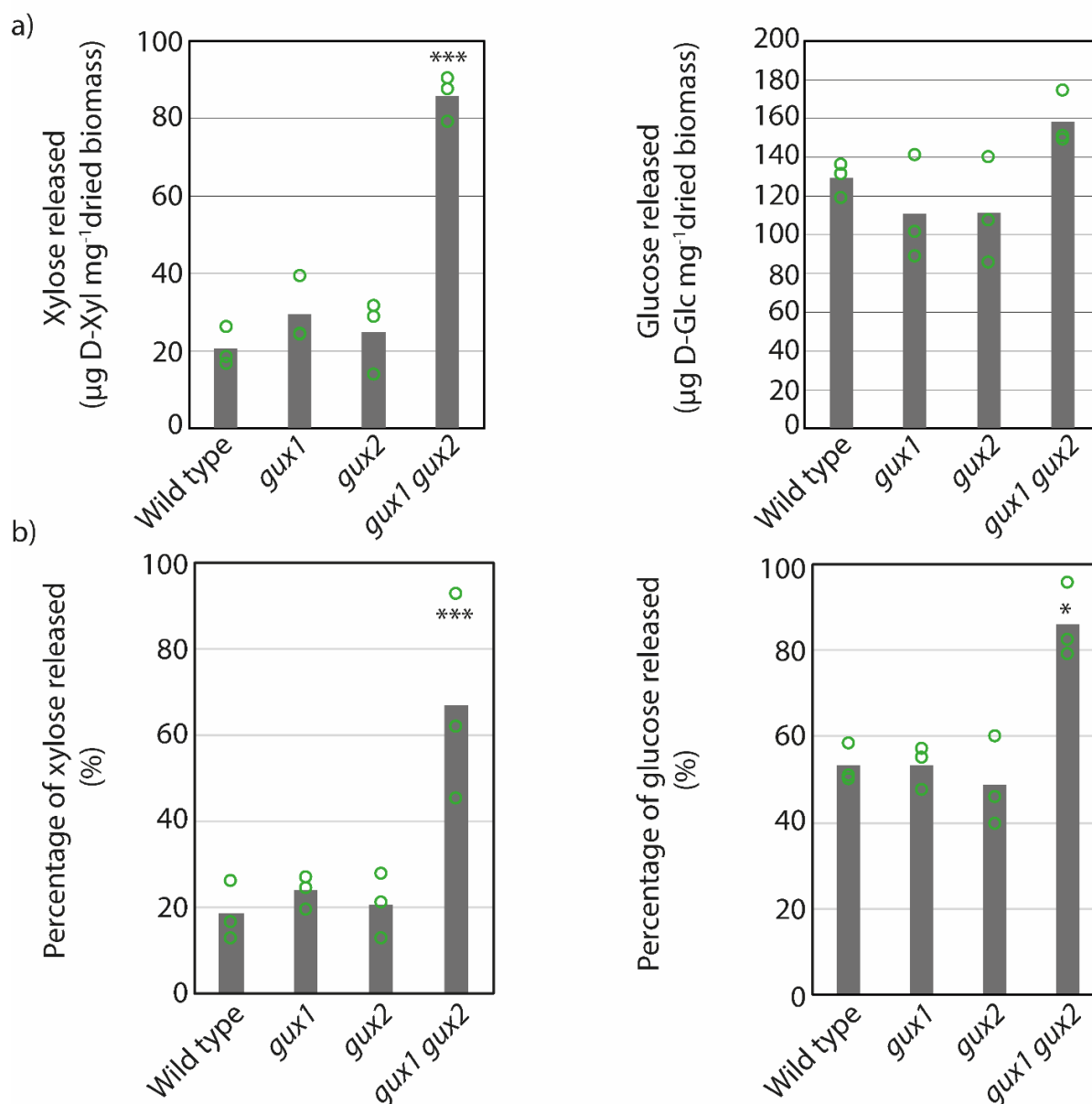
Figure 3.1: There is no proportionality between xylan glucuronidation and recalcitrance. a) Ball-milled dried biomass at 1mg ml^{-1} was saccharified with CTec2 for 24 hours at 45°C. The Xyl and Glc released was measured using Megazyme enzymatic assays. b) The total Xyl and Glc released from Saemann hydrolysis of the dried biomass aliquots was measured by HPAEC-PAD, and the data from a) is expressed as a percentage of this total. Statistical significance was assessed by one way ANOVA with pairwise comparison of means. Significant p values between wild type and the mutants are shown. *=p<0.05, **p<0.01, ***p<0.001. Each data point is the mean of three technical replicates of saccharification of material from one biological replicate. Each biological replicate used material from at least 24 plants.

Table 3.1: Amount of xylose and glucose released from Saemann hydrolysis

	WT	gux1	gux2	gux1 gux2
µg Xyl mg^{-1} dried biomass	114 (±24)	125 (±35)	119 (±14)	140 (±49)
µg Glc mg^{-1} dried biomass	244 (±25)	207 (±43)	227 (±11)	186 (±28)

It has previously been reported that mutation of GUX enzymes affects the monosaccharide composition of the cell wall, possibly due to alterations in metabolic flux due to excess UDP-GlcA in the Golgi (Chong et al., 2015). Therefore, different total amounts of Xyl and Glc in the cell wall may be affecting the apparent recalcitrance of the different mutants. To test this idea, the total amount of Xyl and Glc were measured by HPAEC-PAD after Saemann hydrolysis, and the data presented in Figure 3.1a was re-expressed as a percentage of the total Xyl and Glc in each genotype (Figure 3.1b). The raw data of total xylose and glucose content is shown in table 3.1. There were small statistically insignificant differences in the mean percentage of Xyl and Glc released from wild type and the single *gux* mutants, around 20% and 53% respectively. In the *gux1 gux2* mutant, the mean percentage of sugars released was 66% of Xyl and 86% of Glc. The increases in Glc and Xyl released were both statistically significant compared to wild type, after one-way ANOVA analysis. These results show that there is no proportionality between enzymatic recalcitrance and xylan glucuronidation, instead more than 70% of glucuronidation must be genetically removed to have an effect on recalcitrance. This suggests that accumulation of CTec2-inhibiting glucuronidated xylo-oligosaccharides is not a significant factor in the recalcitrance differences between wild type and *gux1 gux2*.

To further investigate if saccharification products of wild type inhibit the CTec2 saccharification enzymes, a 50:50 mixture of wild type and *gux1 gux2* dried biomass was saccharified and the Xyl and Glc released was compared to the amount released from saccharification of wild type and *gux1 gux2* biomass alone. If saccharification products from the wild type inhibit saccharification, it would be expected that the saccharification of *gux1 gux2* would be inhibited by the presence of wild type biomass. In this case, the Xyl and Glc release would be lower than the average of that released from the wild type and *gux1 gux2*. If no inhibition occurs one would expect the release of Xyl and Glc to be the average of that released from wild type and *gux1 gux2*. After 24 hours of hydrolysis, the Xyl and Glc from the three samples was measured using the relevant Megazyme assays (Figure 3.2). The mean sugar release from wild type and *gux1 gux2* was similar to that reported in Figure 3.1a. The calculated average of the wild type and *gux1 gux2* numbers is 55 µg Xyl and 134 µg Glc mg^{-1} dried biomass. The mean sugar released from the 50:50 mixture of wild type and *gux1 gux2* biomass was 54 µg Xyl and 134 µg Glc mg^{-1} dried biomass. A paired t-test between the calculated average and the mean sugar released from the 50:50 mixture showed that they were not significantly statistically different (Xyl p value = 0.17, Glc p value = 0.52). This further suggests that there is no
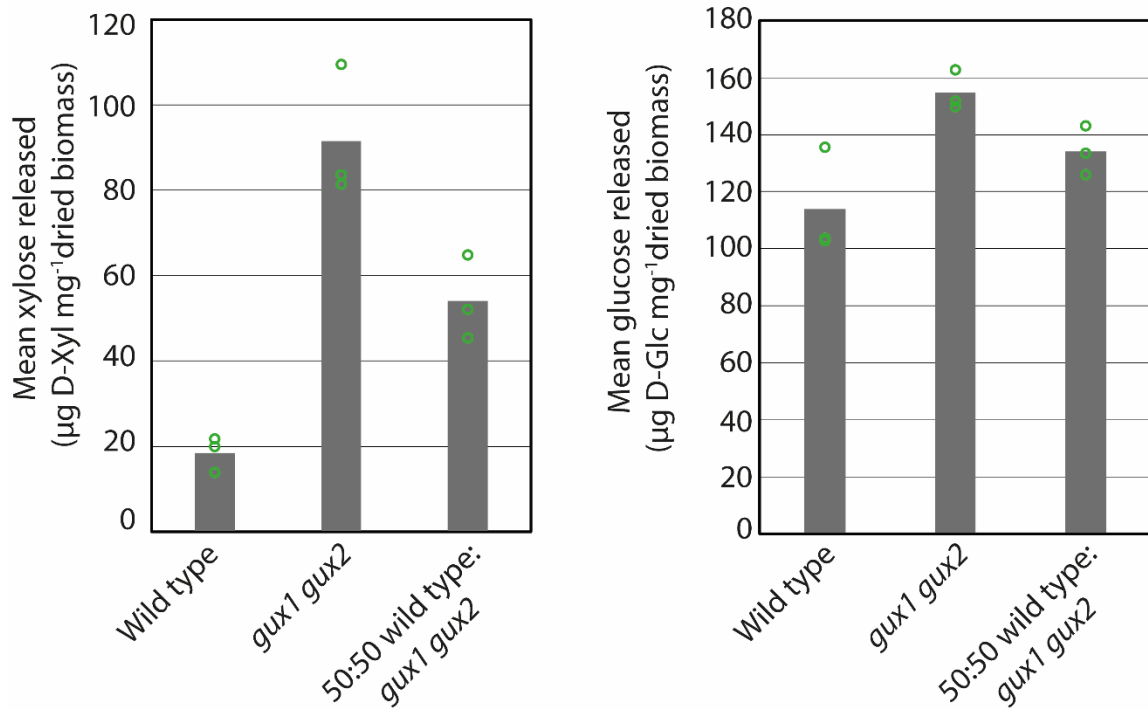
Figure 3.2: Wild type saccharification products do not inhibit saccharification of *gux1 gux2*. Wild type, *gux1 gux2* or a 50:50 mixture ball-milled dried biomass at 1mg ml^{-1} was saccharified with CTec2 for 24 hours at 45ºC. The Xyl and Glc released was measured using Megazyme enzymatic assays. Each data point is the mean of three technical replicates of saccharification of material from one biological replicate. Each biological replicate used material from at least 24 plants.

inhibition of CTec2 by wild type soluble saccharification products.

Next, the role of non-productive enzyme binding was investigated as a mechanism of wild type recalcitrance. Wild type ball-milled dried biomass was saccharified with CTec2 as previously described, and a CTec2 control with no biomass was also incubated for the same time period. The supernatant of each reaction was then removed and used to saccharify fresh wild type or *gux1 gux2* dried biomass for 24 hours. If significant non-productive enzyme binding had occurred, it would be expected that the saccharification yield from CTec2 that had been previously used to saccharify wild type would be significantly lower than from the saccharification with CTec2 that had not already been applied to biomass. The Xyl from the first and the second saccharification reactions was measured with the Megazyme kit to enable calculation of the Xyl released during the second saccharification. The difference in the Xyl released from wild type and *gux1 gux2* biomass was statistically insignificant whether CTec2 had been previously incubated with biomass or an empty tube; the data is shown in Figure 3.3. This suggests that non-productive enzyme binding does not significantly contribute to the recalcitrance differences between wild type and *gux1 gux2*.

3.2.2 There is an increase in the accessibility of xylan and cellulose to enzymes in the *gux1 gux2* mutant

It has been reported that the size of lignocellulose particles significantly affects the saccharification yield, due to the effect of surface area to volume ratio (Bragatto et al., 2012). Although the ball-milling of each genotype was matched in intensity, there could still be differences in the resultant particle size, which could cause genotype dependent differences in saccharification yield. To test this idea, dried stems of wild type and *gux1 gux2* were chopped into approximately 1 mm sections and saccharified using CTec2. The Xyl and Glc released was measured using Megazyme kits (Figure 3.4). Both the wild type and *gux1 gux2* released less Xyl and Glc than saccharification of ball-milled material (Figure 3.1 and Figure 3.4), suggesting ball-milling is an important way of reducing recalcitrance. The *gux1 gux2* mutant released 6x more Xyl than wild type and a similar amount of Glc. This suggests the effect of ball-milling on lignocellulose particle size is not a significant factor in the recalcitrance differences between wild type and *gux1 gux2*. This analysis would benefit from further biological replication, but this was not performed as the experiment used a large amount of biomass. A statistical test was not performed as
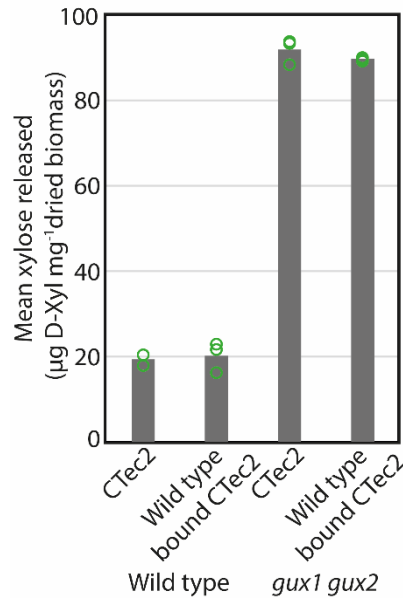
Figure 3.3: Non-productive enzyme binding is not a significant contributor to wild type recalcitrance. Wild type dried biomass at 1mg ml^{-1} and a no biomass control was saccharified with CTec2 for 24 hours at 45ºC. The supernatant of each was used to saccharify wild type and *gux1 gux2* dried biomass at 1mg ml^{-1} for 24 hours at 45ºC. The Xyl released from the initial saccharification, and from the second saccharification was measured using the Megazyme kit. The Xyl released from the initial saccharification was subtracted from the second saccharification. Each data point is the mean of three technical replicates of saccharification of material from one biological replicate. Each biological replicate used material from at least 24 plants.
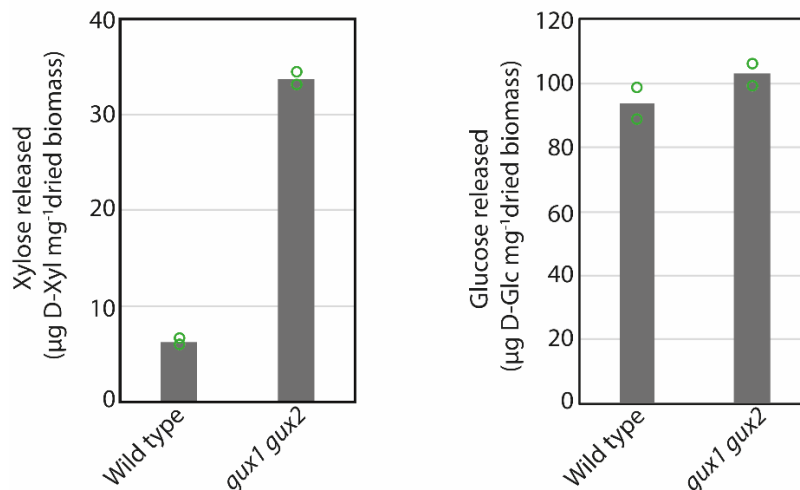
Figure 3.4: The *gux1 gux2* saccharification phenotype is not fully dependent on ball-milling. Wild type and *gux1 gux2* dried stems were chopped into roughly 1 mm sections and saccharified with CTec2 for 24 hours at 45ºC. The Xyl and Glc released was measured using the Megazyme kits. Each data point is one technical replicate of a single biological replicate of at least 24 plants.

66

there was only one biological replicate.

The porosity of cell walls has been reported to be a significant factor in the recalcitrance of cell walls (Meng and Ragauskas, 2014; Pihlajaniemi et al., 2016). This is because an increase in the total volume of nanometre pores with a diameter above 10nm may enable enzymes to penetrate lignocellulose and access more substrate. To assess the porosity of the cell walls in different mutants, thermoporometry by differential scanning calorimetry (TP-DSC) was performed (Driemeier et al., 2012). TP-DSC experiments were performed by Dr Carlos Driemeier, and he processed the data to produce freezing bound water (FBW) and accessible pore area measurements. TP-DSC measures the amount of heat required to melt ice in a material. The melting point of ice in pores is suppressed and the temperature of melting is proportional to the diameter of the pore, thus the amount of heat absorbed at different temperatures can be used to assess the volume of ice in different diameter pores, (see chapter 2 for further details). Figure 3.5a shows the amount of FBW in different diameter pores in wild type, *gux1*, *gux2* and *gux1 gux2* ball-milled dried biomass. Below 10nm, the FBW of ice in the different mutants and wild type is similar. Above 10 nm, the enzyme-accessible pore space, more FBW is found in *gux1 gux2* than the other samples, and this trend is more pronounced at larger pore diameters. Figure 3.5b shows the accessible pore area, the pore area per gram of biomass that occurs in pores with a diameter of 10 nm or more. The wild type has an accessible pore area of 21.5 m^2 g^{-1}, while *gux1* and *gux2* have 22.1 and 22.2 m^2 g^{-1} respectively, a statistically insignificant difference. One way ANOVA analysis showed that the accessible pore area of 26.6 m^2 g^{-1} in the *gux1 gux2* mutant is a statistically significant increase over wild type.

If the increase in porosity of *gux1 gux2* leads to an increase in enzyme accessibility, one would expect that simple enzyme mixtures, with activities specific for digestion of one polysaccharide, would release more sugar from *gux1 gux2* than wild type. To test if a xylanase can release significant amounts of Xyl from *gux1 gux2*, an *Aspergillus aculeatus* GH10 xylanase (*Aa*GH10) at different concentrations was applied to *gux1 gux2* biomass at 45°C for 24 hours. The supernatant was then hydrolysed into monomeric Xyl by saccharification with CTec2 and the Xyl concentrations were measured using the Megazyme kit (Figure 3.6). As the concentration of *Aa*GH10 increased from 5 ng ml^{-1} to 10 ng ml^{-1} the Xyl released increased from 6 µg Xyl mg^{-1} dried biomass to 13 µg Xyl mg^{-1} dried biomass. After this point, doubling the concentration of *Aa*GH10 lead to no
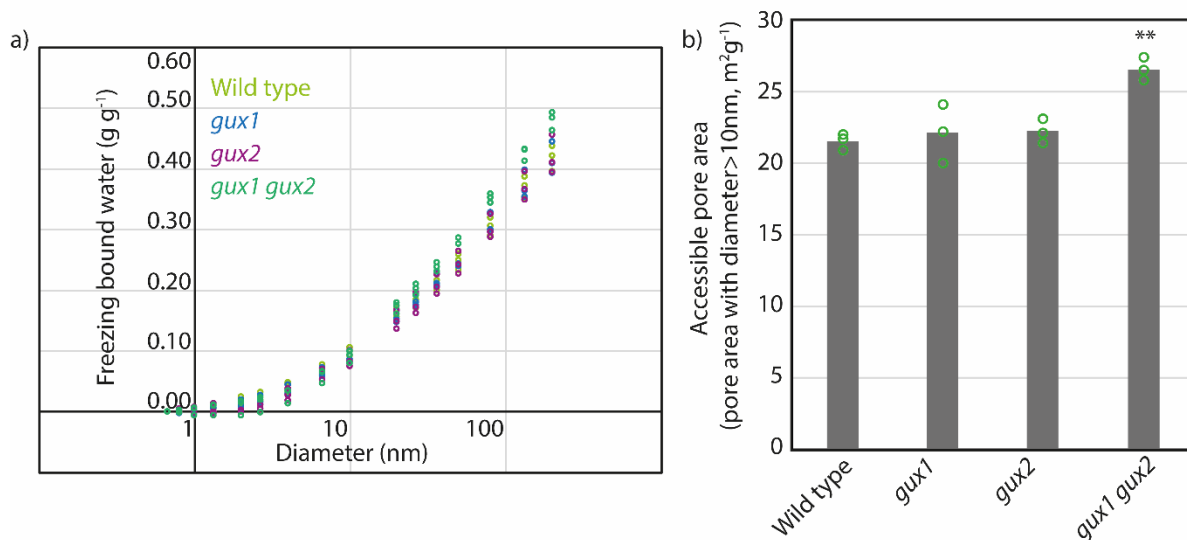
Figure 3.5: The *gux1 gux2* mutant has increased cell wall porosity. Thermoporometry by differential scanning calorimetry of ball-milled dried biomass was used to calculate a) the freezing bound water in pores of different diameter, and b) the pore area that is above 10 nm, and therefore accessible to enzymes. Statistical significance in b) was assessed by one way ANOVA with pairwise comparison of means, only significant p values between wild type and the mutants are shown. *=$p<0.05$, **$p<0.01$, ***$p<0.001$. Each data point is the mean of three technical replicates of material from one biological replicate. Each biological replicate used material from at least 24 plants.

Figure 3.6: *AaGH10* alone digests significant amounts of xylan. *Aa*GH10 at different concentrations was applied to *gux1 gux2* ball-milled dried biomass at 1mg ml^{-1} at 45°C for 24 hours. The oligosaccharides in the supernatant were then saccharified with CTec2. The Xyl released was measured using the Megazyme kits. Each data point is one technical replicate of a single biological replicate of at least 24 plants.

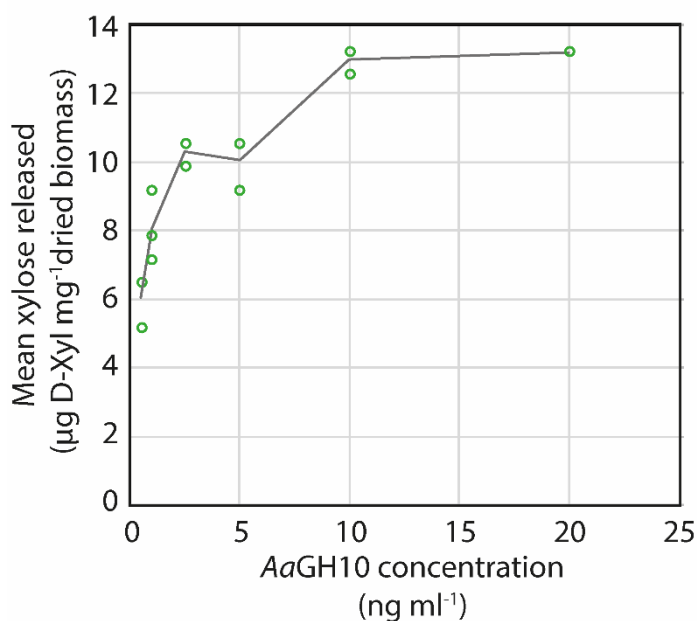significant increase in Xyl release. These results suggest that the porosity of the *gux1 gux2* cell wall allows substantial digestion of xylan without degradation of any other parts of lignocellulose. The Xyl released was over 6.8x less than when *gux1 gux2* was saccharified with CTec2 (Figure 3.1). The results further reveal the optimised amount of xylanase required for this accessibility experiment.

When performing the xylan enzyme accessibility experiment on wild type, very little Xyl was released, data not shown. The glucuronic acid substitutions found on wild type xylan may reduce the enzyme accessibility of wild type xylan to *Aa*GH10. To resolve this issue, an acetyl esterase and GH115 enzyme were included with *Aa*GH10 in enzyme accessibility experiments on wild type. To compare the effect of *gux* mutation on the enzyme accessibility of xylan, wild type, *gux1*, *gux2* and *gux1 gux2* ball-milled dried biomass was incubated with *Aa*GH10, GH115 and an acetyl esterase for 24 hours at 45°C. The supernatant of this digestion was then saccharified into monomeric Xyl by CTec2 and the Xyl released was measured using the Megazyme kit, the data is shown in Figure 3.7a. The Xyl released from the wild type was 3.3 µg Xyl mg^{-1} dried biomass, and from *gux1 gux2* was 33.5 µg Xyl mg^{-1} dried biomass. The Xyl released from *gux1* was 6.1 µg Xyl mg^{-1} dried biomass and from *gux2* was 4.8 µg Xyl mg^{-1} dried biomass. Only *gux1 gux2* released significantly more Xyl than wild type. To account for the different amounts of Xyl in each mutant, the Xyl released was expressed as a percentage of the total Xyl released from Saemann hydrolysis of dried biomass (Figure 3.7b). The percentage of Xyl released from wild type, *gux1* and *gux2* was 2.9, 4.7 and 4.0% respectively. Despite the additional Xyl in *gux1 gux2*, the percentage of Xyl released was 24.4 %. These results suggest that the significantly greater porosity of the *gux1 gux2* cell wall significantly increases the accessibility of xylan to xylanases.

To test if there is an increase in the enzyme accessibility of cellulose as well, a similar experiment with cellulases was performed. A GH45 endoglucanase, GH7 cellobiohydrolase and engineered GH1 β-glucosidase (Santos et al., 2019) was applied to ball-milled dried biomass for 24 hours and 45°C. The supernatant of this digestion was then removed and saccharified with CTec2 and the Glc released was measured with the Megazyme kit (Figure 3.8a). The initial optimisation experiments (enzyme amounts and combinations) were performed by me (data not shown), but the data shown here was produced by Dr Clelton A. Santos. The amount of Glc released from wild type, *gux1*, *gux2*, *gux1 gux2* was 15.5, 16.4, 23.0 and 21.4 µg Glc mg^{-1} dried biomass. No differences were
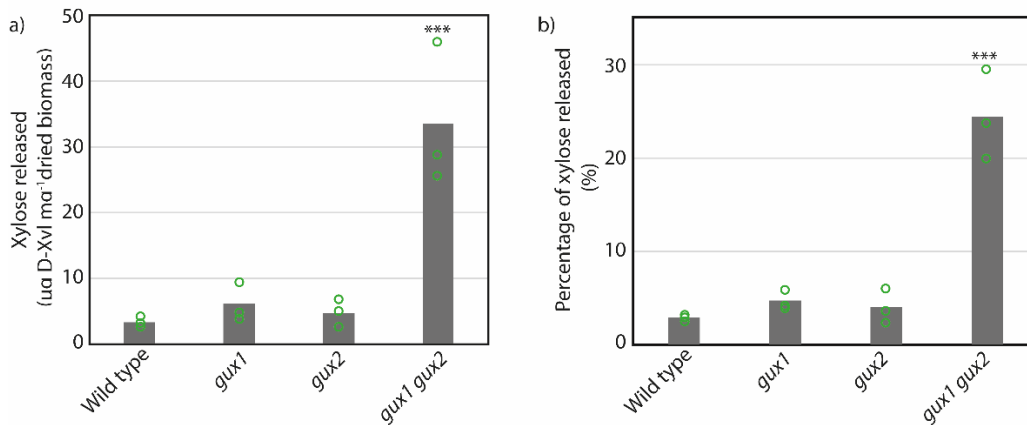
Figure 3.7: *gux1 gux2* xylan is more accessible to xylanases a) Ball-milled dried biomass at 1 mg ml^{-1} digested with *Aa*GH10, acetyl esterase and GH115 α-glucuronidase for 24 hours at 45°C. The oligosaccharides in the supernatant were removed and saccharified with CTec2. The Xyl released was measured using the Megazyme enzymatic assay. b) The data in a) is expressed as a percentage of the total Xyl released from Saemann hydrolysis of the dried biomass aliquots, measured by HPAEC-PAD. Statistical significance was assessed by one way ANOVA with pairwise comparison of means, only significant p values between wild type and the mutants are shown. *=p<0.05, **p<0.01, ***p<0.001. Each data point is the mean of three technical replicates of saccharification of material from one biological replicate. Each biological replicate used material from at least 24 plants.

Figure 3.8: There are no significant differences in cellulose accessibility. a) Ball-milled dried biomass at 1 mg ml^{-1} digested with GH45 endoglucanase, GH7 cellobiohydrolase and GH1 β-glucosidase for 24 hours at 45°C. The oligosaccharides in the supernatant were removed and saccharified with CTec2. The Glc released was measured using the Megazyme enzymatic assay. b) The data in a) is expressed as a percentage of the total Glc released from Saemann hydrolysis of the dried biomass aliquots, measured by HPAEC-PAD. Statistical significance was assessed by one way ANOVA with pairwise comparison of means, only significant p values between wild type and the mutants are shown. *=p<0.05, **p<0.01, ***p<0.001. Each data point is the mean of three technical replicates of saccharification of material from one biological replicate. Each biological replicate used material from at least 24 plants.

statistically significant. To account for the differences in Glc released, the data was expressed as a percentage of the Glc released from Saemann hydrolysis. The percentage of Glc released from wild type, *gux1*, *gux2* and *gux1 gux2* was 6.4, 8.1, 10.2 and 11.6 %. Despite the greater percentage Glc release from *gux2* and *gux1 gux2*, the differences were not statistically significant. These results suggest that the increase in enzyme accessibility is specific to xylan, and does not extend to cellulose.

3.3   Discussion

Previously it was shown that the *gux1 gux2* mutant has a notably large reduction in recalcitrance and it was demonstrated that this was not due to a lack of glucuronidase activity in the saccharification enzyme mixture (Lyczakowski et al., 2017). Multiple mechanisms for the recalcitrance of biomass to enzymatic digestion have been proposed. These include various mechanisms of enzyme inhibition and specific morphological characteristics of lignocellulose (Meng and Ragauskas, 2014). In this chapter the extent to which these different factors contribute to the reduction of recalcitrance observed in *gux1 gux2* was evaluated. Notably it was found that the morphological characteristic of cell wall porosity was significantly increased in *gux1 gux2* and that this likely leads to increased penetration of xylanases into lignocellulose and hence more extensive digestion. This may be caused by nanoscale changes in the interactions between different cell wall components.

3.3.1   Enzyme inhibition is not a significant factor in the recalcitrance differences between wild type and *gux1 gux2*

Previously it has been reported that enzyme inhibition is a significant factor that reduced saccharification yields and enzyme activity, and this could be a cause of the reduced recalcitrance in the *gux1 gux2* mutant (Qing and Wyman., 2011). In the context of saccharification, enzyme inhibition can occur in two main ways: saccharification products can accumulate competitively inhibit enzyme activity (Qing and Wyman., 2011; Xue et al., 2015; Qin et al., 2016) or enzymes can non-productively bind to non-substrate materials (Yarbrough et al., 2015), such as glycosyl hydrolases binding to lignin.

The former could be a significant factor in the reduced recalcitrance of the *gux1 gux2* mutant. Many studies have shown xylo-oligomers can accumulate and inhibit glycosyl

hydroxylases (Xue et al., 2015; Kormelink and Voragen, 1993), especially cellobiohydrolases and β-glucosidases (Kont et al., 2013; Santos et al., 2019). The extracted xylan of the *gux1 gux2* mutant can be cut into smaller oligosaccharides than wild type (Mortimer et al., 2010), and they are also neutral, due to the lack of glucuronic acid branches. Thus, in comparison to wild type, the smaller, more neutral oligosaccharides released from *gux1 gux2* may exhibit less of an inhibitory effect on CTec2 during saccharification. Previously, to investigate this idea, an α-glucuronidase was used to supplement CTec2, to try and boost wild type saccharification to *gux1 gux2* levels (Lyczakowski et al., 2017) but no significant boosting effect was observed, suggesting enzyme inhibition is not a significant factor in the reduced recalcitrance of *gux1 gux2*.

To explore the role of enzyme inhibition further, in this chapter experiments were performed to investigate if there was a proportionality between recalcitrance and the amount of glucuronic acid decoration, by saccharifying single *gux* mutants. However the single *gux* mutants, with varying amounts of glucuronic acid on their xylan, had no differences in the recalcitrance of their cell walls. In combination with other experiments described in Section 3.2.1 this suggests that enzyme inhibition by accumulated oligosaccharides or non-productive binding of enzymes does not play a significant role in the differences between the recalcitrance of wild type and *gux1 gux2*.

Interestingly, the unique glucuronic acid patterns do not seem to play a role in the recalcitrance of the cell wall. The GUX1 enzyme adds glucuronic acid to xylan only with an even number of backbone xylosyl residues between consecutive glucuronic acid residues, while GUX2 adds glucuronic acid substitutions more closely together with no preference for an even number of xylosyl residues between consecutive substitutions (Bromley et al., 2013). Thus, the *gux1* mutant only has the tightly clustered substitutions synthesised by GUX2, while the *gux2* mutant only has the evenly patterned substitutions generated by GUX1. Both single mutants have wild type recalcitrance levels, showing that it is the presence and not the patterning of glucuronic acid that is important for cell wall recalcitrance. Supporting this is the *esk1* mutant, which has an altered pattern of glucuronic acid substitutions. The substitutions are more frequent than in wild type and the even patterning of glucuronic acid substitutions is lost. When saccharified with a similar protocol to that used here, the *esk1* mutant has the same recalcitrance as wild type (Lyczakowski et al., 2017). Interestingly the *esk1* mutant has no cellulose-bound

xylan, instead the xylan is found in the three-fold conformation, rather than bound to cellulose in the flat two-fold conformation (Grantham et al., 2017). This suggests that xylan-binding to cellulose is not an important mechanism of recalcitrance, in contrast to previous results suggesting that xylan-binding to cellulose may inhibit cellulose hydrolysis (Kont et al., 2013).

3.3.2 The increase in *gux1 gux2* cell wall porosity likely enhances enzyme accessibility and contributes to the reduction in recalcitrance

Since there seemed to be no effect of enzyme inhibition in the recalcitrance differences of wild type and the *gux1 gux2* mutant, it was hypothesised that the glucuronic acid substitutions of xylan may affect the morphological characteristics of lignocellulose that affect recalcitrance (Meng and Ragauskas, 2014).

Previously, two significant morphological characteristics of lignocellulose that affect recalcitrance have been identified. The first is the size of lignocellulose particles (Bragatto et al., 2012), which affects the total surface area accessible to enzymes (Pihlajaniemi et al., 2016; Pihlajaniemi et al., 2016). The second is the porosity of the cell wall, which also affects the total surface area of lignocellulose accessible to enzymes, larger pores generate enzyme accessible spaces inside lignocellulose particles, accelerating the rate of deconstruction (Bragatto et al., 2012; Meng and Ragauskas, 2014; Pihlajaniemi et al., 2016; Herbaut et al., 2018). In Section 3.2.2 it was found that the size of lignocellulose particles did not strongly affect the recalcitrance differences between wild type and *gux1 gux2*, as dried biomass that had been chopped rather than ball-milled retained large differences in the Xyl and Glc released between the genotypes. The monosaccharide yield was much lower from chopped material than ball-milled, supporting the effect of particle size on recalcitrance in both genotypes. Interestingly, the ball-milling effect is much more significant for the release of Xyl than Glc. This could be due to the localisation of xylan on the cellulose microfibril surface (Simmons et al., 2016).

To evaluate the porosity of the cell wall, TP-DSC was performed on wild type and *gux* mutant ball-milled dried biomass. It was found the single *gux* mutants and wild type had a similar volume of different size pores and a similar accessible pore area. However, the *gux1 gux2* mutant had a significantly greater pore volume in pores of a diameter of greater than 10nm and a 23.7% increased accessible pore area. By applying simple

73

xylanase mixtures to dried biomass and measuring the Xyl released, it was shown that there was a greatly increased enzyme accessibility of xylan in the *gux1 gux2* mutant. In wild type and the single gux mutants, less than 4% of xylan could be digested, while in *gux1 gux2*, over 24% of xylan could be digested. Thus the increased accessible pore area correlates well with the increase in xylan accessibility.

There are many methods in addition to TP-DSC that can assess porosity of cell walls. Solute exclusion can be used to measure porosity: dextrans of different molecular weights will only penetrate biomass with sufficiently large pores and thereby be diluted (Ishizawa et al., 2007). Changes in the dextran concentration of different molecular weights will reflect the volumes of pores of different sizes. Vapour sorption can also be used to assess porosity characteristics of lignocellulose (Driemeier et al., 2012). In addition the T_2 relaxation time of water protons can be measured (Rondeau-Mouro et al., 2008). This relaxation time reflects the mobility of water, which can be affected by the size of the pores that the water can penetrate. The T_2 relaxation time measurements benefit from being able to use native materials at high humidity, and may reflect "native" porosity without any effect of freezing or drying. The conclusions about the increase in cell wall porosity of the *gux1 gux2* mutant would benefit from verification with another of these methods. However, the porosity is a method to assess if the lignocellulose is likely to be more accessible to enzymes, and whether this difference in accessibility explains recalcitrance differences. The assay of xylan accessibility in Section 3.2.2 provides evidence that there is a functional increase in the accessibility of the xylan to enzymes in *gux1 gux2*; it is likely that this is mediated by the porosity changes measured by TP-DSC.

The increase in cell wall porosity in the *gux1 gux2* mutant may not fully account for the reduction in recalcitrance to enzymatic digestion. Dr Carlos Driemeier, who performed the TP-DSC experiments, thought the porosity differences were too low to fully account for the dramatically reduced recalcitrance of the *gux1 gux2* plant (Driemeier, personal communication). Supporting this, a recent study on differently pretreated wheat straw showed that a larger increase in accessible pore area than in Figure 3.5 contributed to smaller increased monosaccharide release during saccharification (de Assis et al., 2018). Thus glucuronic acid may contribute to cell wall recalcitrance by other mechanisms in addition to restricting the porosity of the cell wall.

Differences in porosity could be linked to alterations in composition and interactions between cell wall components. In studies relating composition to enzyme accessibility, it

has been shown that lignin removal from lignocellulose dramatically enhances the digestion of lignocellulose polysaccharides (Ding et al., 2012). Studies investigating lignin removal by pretreatments have shown that lignin removal can effectively enhance the porosity of cell walls, which may lead to enhanced enzyme accessibility (Crowe et al., 2017; Herbaut et al., 2018). In the case of the *gux1 gux2* mutant, the interactions between cell wall components (cell wall molecular architecture) may be altered and could cause the differences in porosity and enzyme accessibility. Supporting the idea that the molecular architecture of the wall is altered, both lignin and xylan were found to be more extractible from the *gux1 gux2* mutant than wild type (Busse-Wicher, personal communication). This suggests interactions between cell wall components are weaker or altered. To investigate this idea further, changes in cell wall molecular architecture will be explored in chapter 4.

Chapter 4: Lignin:xylan cross-linking may cause the recalcitrance phenotype of *gux1 gux2* mutants

4.1 Introduction

Xylan glucuronic acid substitutions have been shown to be very important for the recalcitrance of secondary cell walls (Lyczakowski et al., 2017). Understanding and modulating the recalcitrance of plant cell walls could be important for improving the resistance of plants to pathogens and improving the economic viability of second generation biofuels (Mahon and Mansfield, 2019; Barros et al., 2019). In chapter 3 of this thesis, it was shown that there was a dramatic increase in the accessibility of xylan to enzymes in the *gux1 gux2* mutant, and that this is likely in part enabled by an increase in cell wall porosity. This alteration in xylan accessibility and cell wall recalcitrance is likely to be caused by an alteration in cell wall molecular architecture, the interactions between, and arrangement of, cell wall constituents. The basis of this change in molecular architecture is unknown and was investigated in the work described in this chapter.

Cell wall molecular architecture and recalcitrance may be affected by the chelation of calcium by acidic sugars. In primary cell walls, the carboxylic acid groups of galacturonosyl backbone residues in homogalacturonan chelate calcium (Thakur et al., 1997). The calcium cross-linking enables the formation of egg-box structures between homogalacturonan molecules (Levesque-Tremblay et al., 2015). Such cross-linking is important for cell expansion and plant development (Levesque-Tremblay et al., 2015), but may also be important for cell wall recalcitrance and pathogen susceptibility (Bellincampi et al., 2014). Reduction of Ca^{2+} cross-linking of pectin seems to result in decreased cell wall recalcitrance (Biswal et al., 2018). In secondary cell walls, pectin makes up a much smaller proportion of the cell wall by weight, and is mostly replaced by xylan (Scheller and Ulvskov, 2010). The lignocellulose of wood contains a significant amount of calcium (Etiégni and Campbell, 1991), and it is the most common ion by weight in wood. The two-fold xylan bound to cellulose microfibrils may project glucuronic acid substitutions away from the cellulose surface (Busse-Wicher et al., 2016; Simmons et al., 2016; Kang et al., 2019). These substitutions may be able to chelate calcium

(Lamport and Várnai, 2013; Kardošová et al., 1990) and thus cross-link separate cellulose microfibrils with xylan bound on the surface, see Figure 4.1a.
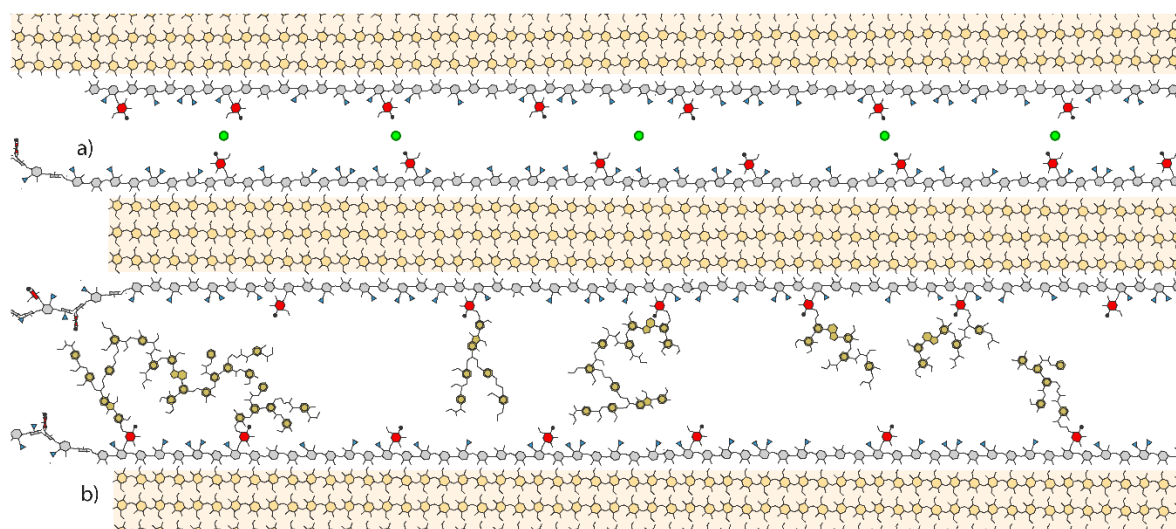
Figure 4.1: Models of the molecular architecture function of glucuronic acid substitutions of xylan. a) Xylan is shown bound to cellulose hydrophilic faces as a two-fold screw. The glucuronic acid substitutions project into the inter-microfibril space and chelate calcium ions. Ca^{2+} chelation between xylans on different microfibrils cross-links the different microfibrils. b) Some glucuronic acid substitutions are shown as lignin-glucuronate esters, cross-linking lignin and xylan together.

The chemistry of lignification may enable glucuronic acid-mediated cross-linking between xylan and lignin, and this could be important for cell wall recalcitrance. When monolignol radicals couple at the β position, a quinone methide intermediate is formed, which is re-aromatised by nucleophilic attack at the α position in the β-O-4 linkage (Terrett and Dupree, 2019), see Section 1.3.2. This nucleophile can be water or a polysaccharide (Nishimura et al., 2018). There is some indirect evidence suggesting that glucuronic acid can act as this nucleophile, forming a lignin-glucuronate ester (Das et al., 1984; Arnling Bååth et al., 2016; Martínez-Abad et al., 2018). Lignin and polysaccharides that are cross-linked are frequently referred to as lignin carbohydrate complexes (LCCs) in the literature (Giummarella et al., 2019). Convincing evidence of a lignin-glucuronate ester has yet to be demonstrated (Terrett and Dupree, 2019; Giummarella et al., 2019). Convincing evidence requires extensive extraction procedures and complex NMR analysis (Nishimura et al., 2018). This covalent cross-linking could be important for cell wall recalcitrance (Terrett and Dupree, 2019). A model of the cell wall with lignin-glucuronate esters cross-linking the cell wall is shown in Figure 4.1b.

Various lignin mutants may provide complementary genetic evidence to the *gux1 gux2* mutant when analysing the function of lignin-glucuronate esters in cell wall recalcitrance to enzymatic digestion. The *cad2 cad6* mutant has lignin which is composed of 13% normal monolignols (Vanholme, personal communication) (see introduction Section 1.2.3.1 for an explanation of cinnamyl alcohol dehydrogenase (CAD) function in lignin biosynthesis). Most *cad2 cad6* lignin monomers have aldehyde functional groups at the γ position, instead of hydroxyls (Sibout et al., 2005). This results in a change in the chemistry of lignification; instead of re-aromatising the quinone methide intermediate by nucleophilic attack, the 7 proton is eliminated (Terrett and Dupree, 2019), thus *cad2 cad6* is unlikely to be able to form the proposed xylan:lignin ester cross-links through glucuronic acid. See introduction Figures 1.7-8 for an illustration of these processes. If *cad2 cad6* shares cell wall phenotypes with *gux1 gux2*, this would support the hypothesis that glucuronic-acid mediated xylan:lignin cross-linking is a causative factor in the *gux1 gux2* recalcitrance phenotype.

The *cad2 cad6* mutant has pleiotropic lignin phenotypes, so multiple lignin mutant controls must be included in a comparison of *cad2 cad6* to *gux1 gux2*. The *cad2 cad6* has a 20% reduction in total lignin content, so mutants such as *Arabidopsis thaliana cytochrome p450 reductase 2* (*atr2*) (Sundin et al., 2014), *laccase 4* (*lac4*) (Berthet et al., 2011) and *4-coumarate:CoA ligase 1* (*4cl1*) (Van Acker et al., 2013) which have a similar reduction in lignin content, but a more wild type lignin composition, are important controls for the specific effect of *cad2 cad6* mutation on recalcitrance. It has also been reported that *CAD* down-regulation results in lower molecular weight lignin in conifers (Tolbert et al., 2014), though the differences are comparable to interspecific variation, and lignin molecular weight has not been investigated in Arabidopsis *cad* mutants. To act as a control for the potential size effect of *CAD* mutation on lignin molecular weight, the *ref8 med5a/b* mutant, which has also been reported to have a reduction in lignin molecular weight (Bonawitz et al., 2014), was included in the analysis. *ref8 med5a/b* likely also has a reduced capacity for this type of xylan:lignin cross-linking due to an increase in dibenzodioxocin linkages (Shi et al., 2016). See introduction Section 1.2.3.1 for an explanation of the roles of these enzymes in lignin biosynthesis.

The overall aim of the work described in this chapter is to investigate if the reduction in recalcitrance to enzymatic digestion of the *gux1 gux2* mutant is related to changes in the secondary cell wall molecular architecture. The specific aims are:

- To analyse the effect of changing Ca^{2+} concentration during growth on the recalcitrance of wild type and *GUX* mutant plants;

- To investigate if specific mutations in lignin synthesis lead to similar reductions in recalcitrance to the *gux1 gux2* mutant;

- To use solid-state NMR to analyse if *in muro* lignin-xylan proximity is affected in the *gux1 gux2* mutant;

- To develop techniques to analyse lignin:glucuronate esters.

4.2   Results

4.2.1   Calcium concentrations do not affect the recalcitrance of secondary cell walls

To analyse the effect of Ca^{2+} concentrations in secondary cell walls on recalcitrance, wild type, *gux1*, *gux2* and *gux1 gux2* were grown hydroponically in rockwool at different Ca^{2+} concentrations.   The Ca^{2+} concentrations selected were previously found to have significant developmental effects in arabinogalactan protein (AGP) glucuronidation mutants (Hernández, 2018). This suggests these growth conditions significantly reduce the amount of Ca^{2+} in the apoplast and affect interactions between Ca^{2+} and glucuronic acid on some polysaccharides. Therefore these Ca^{2+} concentrations may be physiologically relevant to any interactions between Ca^{2+} and the glucuronic acid substitutions of xylan.

To ensure that the growing conditions were effective in producing Ca^{2+}-depleted plants, the height of mature eight week old plants was measured, as this has previously been reported to be reduced in Ca^{2+}-depleted conditions (Hernández, 2018). For each genotype, except for *gux1*, the reduction in Ca^{2+} concentration from 1 mM to 0.2 mM induced statistically significant reductions in plant height, Figure 4.2. The plant height reduction between the 1mM condition and the 0.5 mM condition for each genotype was not significant. At any given calcium concentration, there were no statistically significant differences in the height of plants of different genotype. These results suggest that Ca^{2+}concentration has no genotype-dependent effects on plant height. The dwarfing of each genotype, except *gux1*, at 0.2 mM suggests that the growing conditions were effective in altering *in planta* Ca^{2+} concentrations.

If a Ca^{2+}-glucuronic acid interaction was important for cell wall recalcitrance, one might
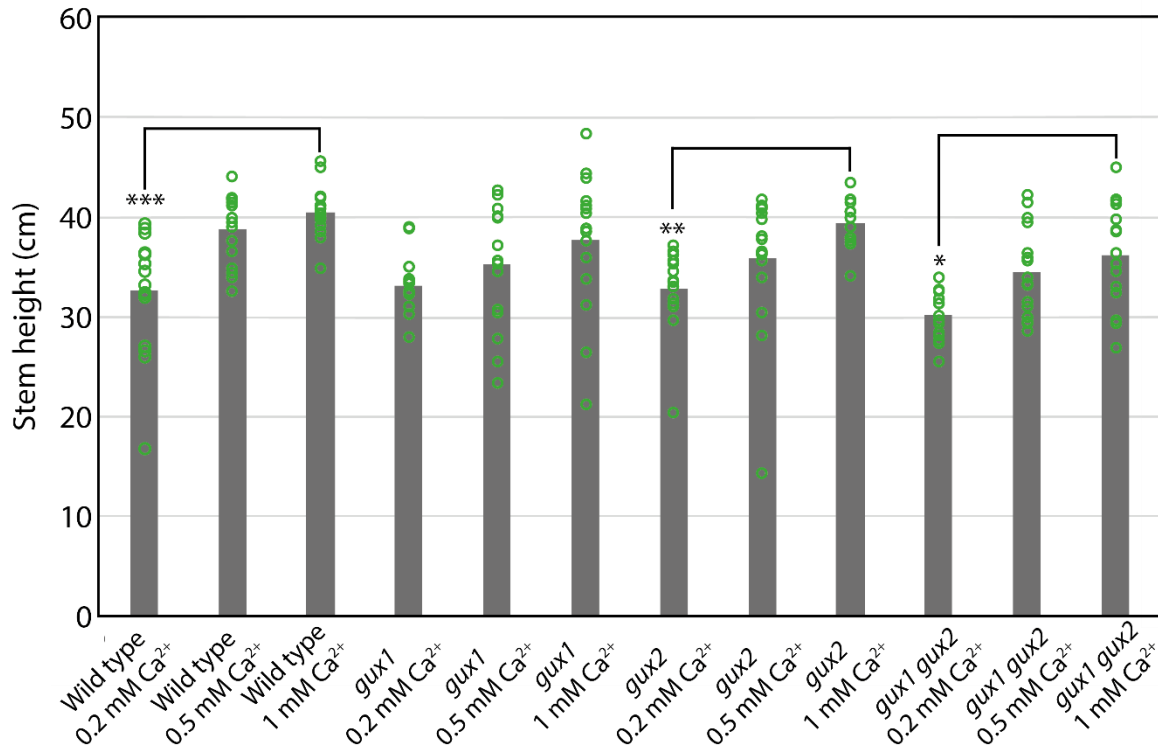
Figure 4.2: Ca^{2+} depleted growth conditions reduce plant growth. Seeds of the four genotypes were sterilised and sown on MS plates. Two week old plantlets were transferred to rockwool in pots that had been soaked in distilled water to remove micronutrients. The plants were then grown for six weeks in hydroponics medium with a CaCl$_2$ concentration of 0.2 mM, 0.5 mM and 1 mM. The height of 8 week old plants was measured from the rosette to the tip of the tallest stem. Statistical significance was assessed by one way ANOVA with pairwise comparison of means, *p<0.05, **p<0.01, ***p<0.001. Each data point is the height of a single plant, 16 plants of each genotype were grown in each CaCl$_2$ condition.

expect that in these Ca^{2+}-depleted growth conditions, the wild type and single *gux* mutant plants have reduced recalcitrance, due to reduced Ca^{2+} binding by glucuronic acid, while the *gux1 gux2* mutant plant would be unaffected. To test if Ca^{2+} depletion affects the recalcitrance of the cell wall, ball-milled dried biomass from each genotype at each Ca^{2+} concentration was saccharified with CTec2, as previously described. The Xyl released was measured using the Megazyme kit, Figure 4.3. In the 1 mM Ca^{2+} growth condition, the Xyl released from wild type, *gux1, gux2* and *gux1 gux2* was 21.7, 40.0, 28.6 and 110.9 µg Xyl mg^{-1} dried biomass respectively. In all genotypes, as the Ca^{2+} decreased the Xyl released also decreased, and these differences were statistically significant. The reduction in released Xyl suggests that the Ca^{2+}-depletion-induced dwarfing may have resulted in a reduction in secondary cell wall content by weight, as has been observed in many *irx* mutants (Lefebvre et al., 2011; Brown et al., 2007). Despite this dwarfing effect, there doesn't seem to be an effect of Ca^{2+}-depletion on recalcitrance. This analysis might benefit from measuring the Glc released as well, but previously it was observed that the larger differences were observed in Xyl release in saccharification of *gux* mutants. As these initial experiments did not suggest a relationship between Ca^{2+} and secondary cell wall recalcitrance, other hypotheses were focussed on instead.
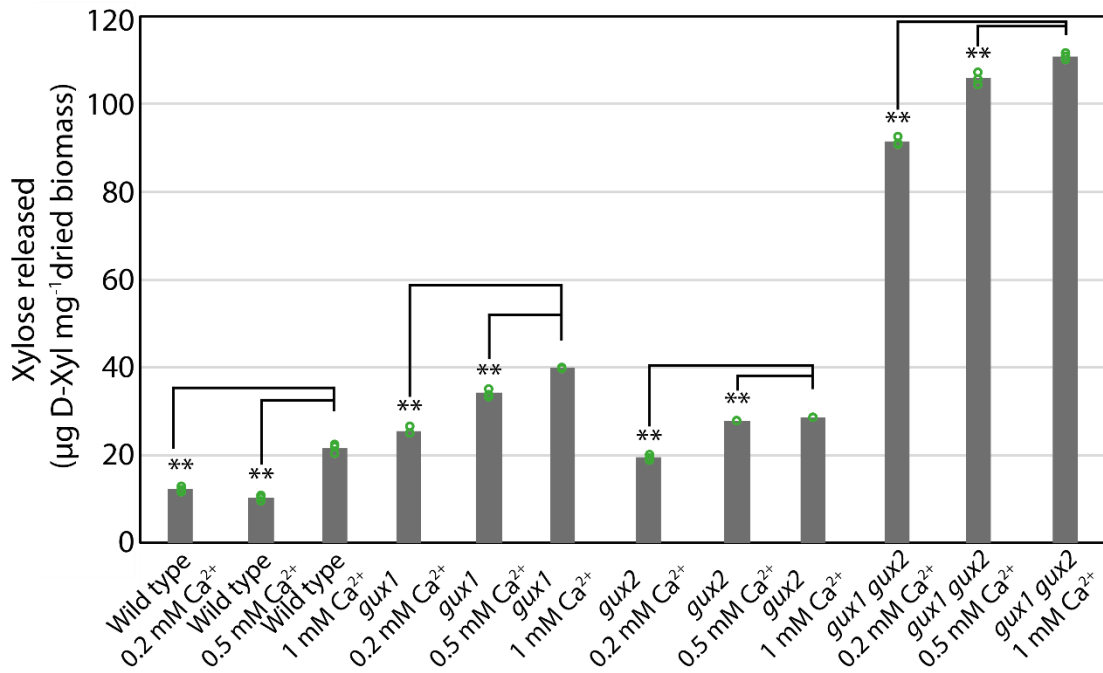
Figure 4.3: Ca²⁺ depleted growth conditions reduce saccharification yields from all genotypes. Ball-milled dried biomass, from the mature hydroponically grown plants, at 1mg ml⁻¹ was saccharified with CTec2 for 24 hours at 45°C. The Xyl released was measured. Statistical significance was assessed by one way ANOVA with pairwise comparison of means, *p<0.05, **p<0.01, ***p<0.001. Only statistical significance between the same genotype in different conditions is marked. Each data point is a technical replicate of saccharification of material collected from all 16 plants in each Ca²⁺ condition.

4.2.2 The *cad2 cad6* mutant shares recalcitrance and accessibility phenotypes of the *gux1 gux2* mutant

Previously, it has been proposed that an ester bond exists between glucuronic acid substitutions of xylan and lignin (Watanabe and Koshijima, 1988). The ester has been proposed to be formed during lignification (Mottiar et al., 2016), by the glucuronic acid acting as a nucleophile to re-aromatise the quinone methide intermediate of two monolignols coupling at the β position (Terrett and Dupree, 2019). The altered lignin composition of the *cad2 cad6* (Sibout et al., 2005) mutant has been proposed to prevent such cross-linking (Kim et al., 2003; Terrett and Dupree, 2019). The data presented here was taken from the same experiments as in Section 3.2.1, so data from wild type and *gux1 gux2* is the same as in that Section.

To investigate if the *cad2 cad6* mutant has reduced recalcitrance beyond that of control lignin mutants, dried biomass from wild type, *gux1 gux2*, *cad2 cad6*, *atr2*, *lac4*, *4cl1*, and *ref8 med5a/b* was saccharified with CTec2, as previously described. The Xyl and Glc released after 24 hours was measured with the relevant Megazyme kit. The Xyl released, Figure 4.4a, from wild type and *gux1 gux2* was 20.5 µg Xyl mg^{-1} dried biomass and 85.9 µg Xyl mg^{-1} dried biomass respectively, as seen in chapter 3. The *cad2 cad6* mutant released 59.3 µg Xyl mg^{-1} dried biomass, which was significantly more than wild type, but significantly less than *gux1 gux2*. The control lignin mutants did not release a significantly higher amount of Xyl than wild type. This data suggests that of the lignin mutants, *cad2 cad6* alone has significantly decreased xylan recalcitrance, similarly to *gux1 gux2*.

It was seen in chapter 3 that there were differences in the amount of Xyl in the cell walls of *gux1 gux2*. To test if there were differences in the percentage of Xyl released from different genotypes, the data in Figure 4.4a was expressed as a percentage of the Xyl released from Saemann hydrolysis of each genotype, Figure 4.4b. The percentage of Xyl released from wild type was 18.7%. The percentage of Xyl released from *gux1 gux2* and *cad2 cad6* was 66.6% and 61.5% respectively, a statistically significant increase over wild type. The difference between *cad2 cad6* and *gux1 gux2* was not statistically significant. The control lignin mutants, *atr2*, *lac4* and *4cl1*, which have a similar reduction in lignin content as *cad2 cad6* (Sundin et al., 2014; Berthet et al., 2011; Van Acker et al., 2013), had a percentage of Xyl released that was not statistically different from wild type.
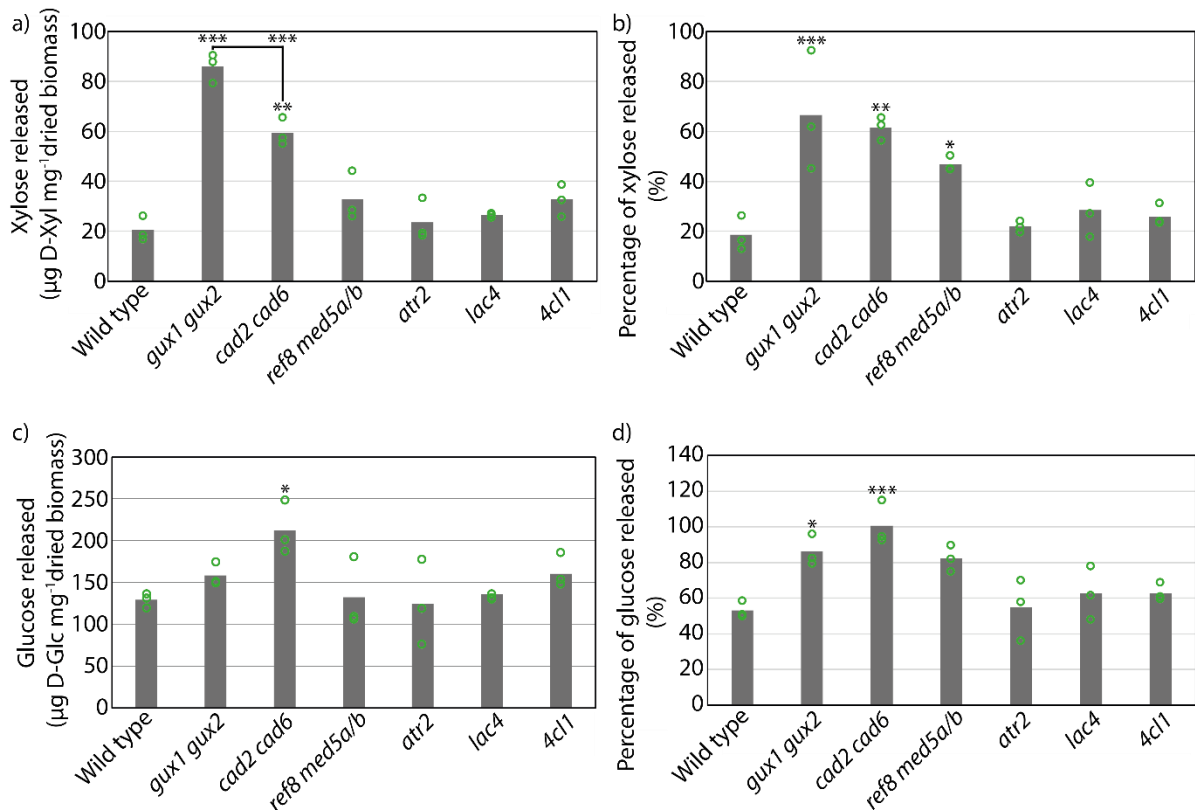
Figure 4.4: *cad2 cad6* **has a similar saccharification phenotype to the** *gux1 gux2* **mutant.** Ball-milled dried biomass at 1mg ml^{-1} was saccharified with CTec2 for 24 hours at 45˚C. a) The Xyl and c) Glc released was measured. b) The Xyl and d) Glc released were expressed as a percentage of the total Xyl and Glc and released from Saemann hydrolysis, measured by HPAEC-PAD. Statistical significance was assessed by one way ANOVA with pairwise comparison of means, *$p<0.05$, **$p<0.01$, ***$p<0.001$. Statistical significance between wild type and a mutant is shown, unless otherwise indicated by black lines. Each data point is the mean of three technical replicates of saccharification of material from one biological replicate. Each biological replicate used material from at least 24 plants.

However, the *ref8 med5a/b* mutant, which has a reduced lignin content, molecular weight and altered composition (Bonawitz et al., 2014), releases 46.9% of the total Xyl. The dwarfing phenotype of this plant likely reduces the secondary cell wall content, concealing the reduced recalcitrance phenotype of this mutant when the Xyl released is expressed per mg of biomass, rather than as a percentage of the total Xyl in the biomass. This data suggests that *gux1 gux2*, *cad2 cad6* and *ref8 med5a/b* have reduced cell wall recalcitrance, though this effect is greatest in *gux1 gux2* and *cad2 cad6*.

To test if the reduced recalcitrance of *cad2 cad6* also extends to cellulose, as in *gux1 gux2*, the Glc released after CTec2 saccharification was measured, Figure 4.4c. The Glc released from wild type was 129.1 µg Glc mg^{-1} dried biomass and from *gux1 gux2* was 158.5 µg Glc mg^{-1} dried biomass, though this increase is statistically insignificant, as seen in Section 3.2.1. The *cad2 cad6* mutant released 212.2 µg Glc mg^{-1} dried biomass, a statistically significant increase over wild type. The four control lignin mutants released a similar amount of Glc as wild type, except for *lac4* which released 162.6 µg Glc mg^{-1} dried biomass, though this increase was not statistically significant.

In Section 3.2.1, the increased amount of Glc released from *gux1 gux2* was only statistically significant when expressed as percentage of the total Glc in the biomass. To compare if the percentage of Glc released was also increased in the *cad2 cad6* mutant, the data in Figure 4.4c was expressed as a percentage of the total Glc released from each genotype during Saemann hydrolysis, Figure 4.4d. As in Section 3.2.1, the percentage of Glc released from wild type and *gux1 gux2* was 53.2% and 86.0% respectively. The percentage of Glc released from *cad2 cad6* and *ref8 med5a/b* was 100.7% and 82.3%, both statistically significant increases over wild type. Although the reduced lignin content control mutants released a higher percentage of Glc than wild type, the differences were not statistically significant. These results suggest that the *cad2 cad6* mutant has a similar reduction in xylan recalcitrance as *gux1 gux2*, but a greater decrease in cellulose recalcitrance. The three control lignin mutants, which only have reductions in lignin content, showed wild type recalcitrance. However, the *ref8 med5a/b* mutant, which has pleiotropic lignin phenotypes, also has greatly reduced xylan and cellulose recalcitrance. This could suggest that the recalcitrance phenotypes of *cad2 cad6* may not just be related to any changes in xylan:lignin cross-linking, but also to changes in lignin molecular weight or lignin composition.

To determine if the *cad2 cad6* mutant has an increase in the accessibility of xylan, like the *gux1 gux2* mutant (see Section 3.2.2), dried biomass from the nine genotypes described was incubated with *Aa*GH10, GH115 and an acetyl esterase for 24 hours at 45°C. The supernatant of this digestion was then saccharified with CTec2 and the Xyl released was measured using the Megazyme kit, Figure 4.5a. The wild type released 3.3 µg Xyl mg^{-1} dried biomass, less than 3% of the total xylan (Figure 4.5b), while *gux1 gux2* released 33.5 µg Xyl mg^{-1} dried biomass, equal to 24.4% of the total xylan in the sample. Interestingly, the *cad2 cad6* mutant released 25.1 µg Xyl mg^{-1} dried biomass, around 26.1% of the total xylan in this mutant. This increase was statistically significant compared to wild type, but not to *gux1 gux2*. The control lignin mutants had statistically insignificant increases in xylan accessibility compared to wild type. However when expressed as a percentage of the total Xyl in the biomass, the *ref8 med5a/b* mutant has a statistically significant increase in xylan accessibility: 19.0% of the total Xyl was released from this mutant.

In Section 3.2.2, it was shown that the increase in enzyme accessibility in the *gux1 gux2* mutant was specific to xylan only. When a similar experiment with cellulose-specific hydrolases was attempted, the release of Glc from different mutants was not distinguishable from wild type. To test if the enzyme accessibility of cellulose in the *cad2 cad6* mutant was the same as wild type, the cellulase mixture of GH45, GH7 and GH1 was applied to the ball-milled dried biomass. The cellulose accessibility experiment was performed by Dr Clelton A. Santos. The supernatant was saccharified with CTec2, and the Glc released was measured using the relevant Megazyme kit. When expressed per mg of dried biomass there were no significant differences in the Glc released from different genotypes, Figure 4.5c. However, when expressed as a percentage of the total Glc in the sample, Figure 4.5d, the wild type released 6.4 % of the total Glc and the *ref8 med5a/b* mutant released 13.8% of the total Glc, one-way ANOVA analysis showed this increase was statistically significant.

Collectively, these results suggest that *cad2 cad6* and *gux1 gux2* have very similar reductions in recalcitrance to enzymatic digestion and an increase specifically in the enzyme accessibility of their xylan. This was predicted on the basis that both mutants should have reduced xylan:lignin cross-linking through glucuronic acid (Terrett and Dupree, 2019). The *ref8 med5a/b* mutant shares some of these phenotypes, suggesting we should be cautious about correlating the phenotypes of *cad2 cad6* and *gux1 gux2.*
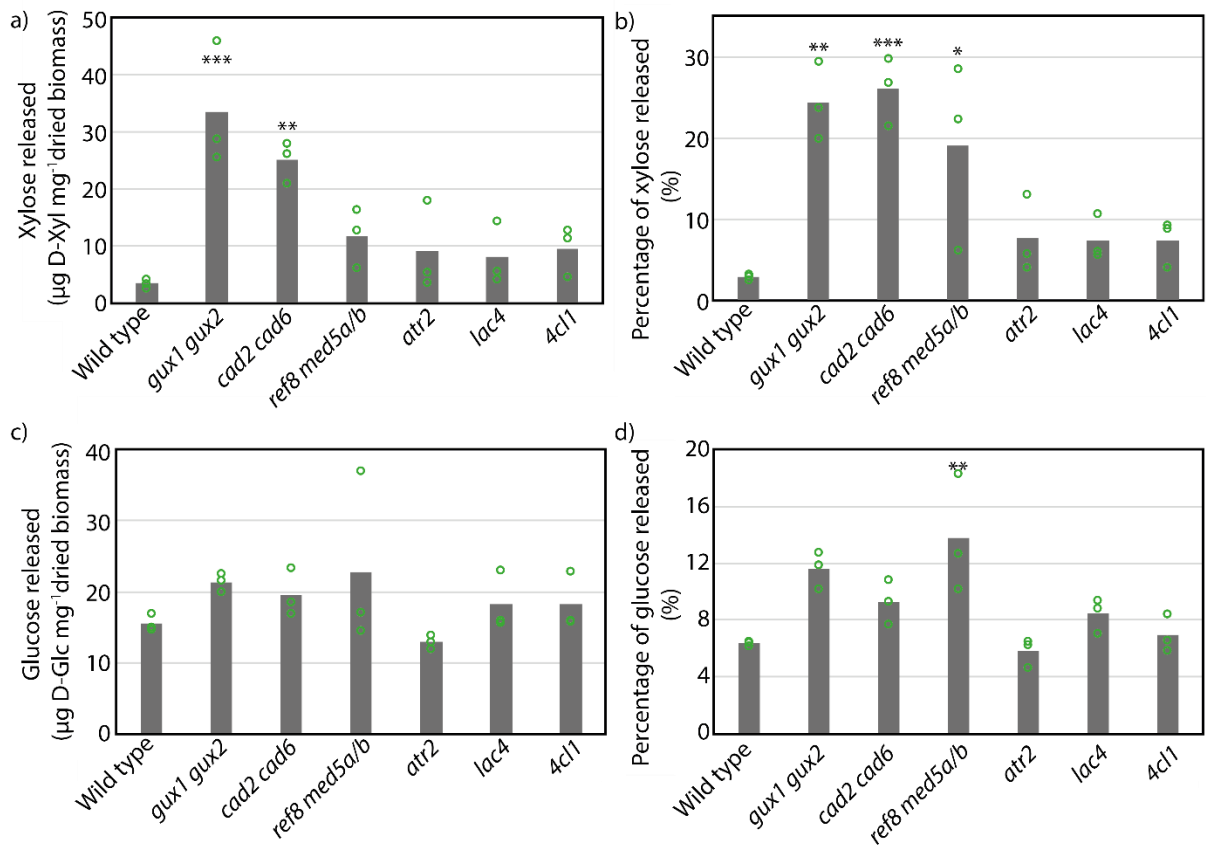
Figure 4.5: *cad2 cad6* has a similar xylan and cellulose accessibility phenotype to the *gux1 gux2* mutant. a) Ball-milled dried biomass at 1mg ml^{-1} digested with AaGH10, acetyl esterase and GH115 α-glucuronidase for 24 hours at 45˚C. The oligosaccharides in the supernatant were removed and saccharified with CTec2. The Xyl released was measured using the Megazyme enzymatic assay. b) Ball-milled dried biomass at 1mg ml^{-1} digested with GH45 endoglucanase, GH7 cellobiohydrolase and GH1 β-glucosidase for 24 hours at 45˚C. The oligosaccharides in the supernatant were removed and saccharified with CTec2. The Glc released was measured using the Megazyme enzymatic assay. c and d) The data in a) and b) is expressed as a percentage of the total xylose and glucose released from Saemann hydrolysis of the dried biomass aliquots, measured by HPAEC-PAD. Statistical significance was assessed by one way ANOVA with pairwise comparison of means, *p<0.05, **p<0.01, ***p<0.001. Statistical significance between wild type and a mutant is shown, unless otherwise indicated, by black lines. Each data point is the mean of three technical replicates of saccharification of material from one biological replicate. Each biological replicate used material from at least 24 plants.

However *ref8 med5a/b* also has a significant increase in cellulose accessibility, suggesting its phenotype is slightly different to *cad2 cad6* and *gux1 gux2*. In addition, the altered lignin linkages in *ref8 med5a/b* may reduce the xylan:lignin cross-linking in this mutant.

4.2.3 Solid-state NMR shows that *cad2 cad6* and *gux1 gux2* have alterations in cellulose structure and a reduction in xylan-lignin proximity

Previously solid-state NMR has been a powerful tool in evaluating the *in muro* interactions of cell wall components and their conformation (Simmons et al., 2016; Grantham et al., 2017; Kang et al., 2019). It was previously shown in grasses, and wild type Arabidopsis, that lignin and xylan are in very close proximity to each other *in muro*, suggesting that lignin coats xylan in the cell wall (Kang et al., 2019), and this may be mediated by covalent linkages between xylan and lignin (Terrett and Dupree, 2019). In Section 4.2.2 it was investigated whether there was genetic evidence supporting the existence of covalent linkages between lignin and glucuronic acid substitutions on xylan by comparing the recalcitrance and enzyme accessibility phenotypes of the *gux1 gux2* and *cad2 cad6* mutants. The similarity of the recalcitrance and xylan accessibility phenotypes of these two mutants suggested that xylan:lignin cross-linking is the cause of the reduction in recalcitrance in the *gux1 gux2* mutant. In this Section, the aim was to investigate if xylan and lignin are in close proximity in Arabidopsis, if this proximity is affected in the *gux1 gux2* and *cad2 cad6* mutants, and if there are other changes in the interactions of cell wall components that might contribute to the reduced recalcitrance to enzymatic digestion in these mutants. The solid-state NMR experiments were performed by Professor Ray Dupree with samples prepared by me, and the data was analysed by me. Experiments were performed on basal stem material from at least 8 plants per genotype. They will need to be replicated on a second biological replicate, but this has not yet been performed due to experimental time restrictions at the National 850 MHz Facility.

Two different solid-state NMR experiments were used in this Section, and they are explained in more detail in Section 6.1. In brief, the CP-INADEQUATE experiment transfers magnetisation between carbons that share a covalent bond. Each carbon in a sugar will have a unique single quantum shift on the x axis of the spectrum, and a double quantum shift on the y axis. The double quantum is the sum of that carbon's chemical shift and the chemical shift of the carbon it is bonded to. The experiment can be used to assign

the chemical shifts of ^{13}C nuclei in different sugars and assess the relative conformations of xylan and cellulose (Simmons et al., 2016; Grantham et al., 2017; Wang et al., 2016b). The CP-PDSD experiment transfers magnetisation between carbons through space. At shorter mixing times the magnetisation travels short distances, and cross-peaks will be limited to carbons within the same molecule. As the mixing time increases, cross-peaks between carbons that are more distant, up to 1 nm, will appear. This experiment can be used to evaluate the sub-nanometre proximity of different cell wall components and therefore draw conclusions about their interactions (Simmons et al., 2016; Kang et al., 2019).

To investigate the conformation of xylan in *gux1 gux2* and *cad2 cad6,* CP-INADEQUATE experiments were performed on wild type, *gux1 gux2* and *cad2 cad6*. Figure 4.6a shows the wild type spectrum, where the carbons 4 and 5 of xylan in both the two-fold ($Xn4^{2f}$ and $Xn5^{2f}$) and three-fold ($Xn4^{3f}$ and $Xn5^{3f}$) conformation have been assigned based on published work (Simmons et al., 2016). Both the *gux1 gux2* (Figure 4.6b) and *cad2 cad6* (Figure 4.6c) mutant spectra have two-fold and three-fold xylan peaks, similarly to wild type. This suggests the xylan in these mutants adopts similar conformations to wild type. The $Xn4^{2f}$ peak for all three genotypes has a range of chemical shifts from 80.5-82.5 ppm, with the greatest intensity at 82.4 ppm. As the chemical shift of Xn4 varies with the glycosidic bond angle (Simmons et al., 2016), it is likely the xylan adopts a range of conformations that are intermediate between a two-fold and three-fold screw.

In the quantification of the amount of Xyl released from Saemann hydrolysis, the *gux1 gux2* mutant released much more Xyl, suggesting an increased amount of xylan in this mutant, Section 3.2.1. To compare the amount of two-fold xylan in the different mutants, a sum projection of the CP-INADEQUATE region between DQ=145.1-148.6 ppm (the $Xn4/5^{2f}$ region) was produced from wild type, *gux1 gux2* and *cad2 cad6*. In addition, the same sum projection was produced from a CP-INADEQUATE spectra of the *eskimo1* mutant, which has been reported to have very little two-fold xylan (Grantham et al., 2017). The projections are shown in Figure 4.7, and the Xn4 and 5^{2f} peaks are labelled. In both wild type and *cad2 cad6* the height of the Xn4 and 5^{2f} peaks are similar, while in the *eskimo1* mutant, as previously reported, the two-fold xylan peaks are virtually absent.

In the *gux1 gux2* mutant, the Xn4 and 5^{2f} peaks are relatively larger than wild type or *cad2 cad6*. This supports the Saemann hydrolysis data and suggests that a larger amount

89
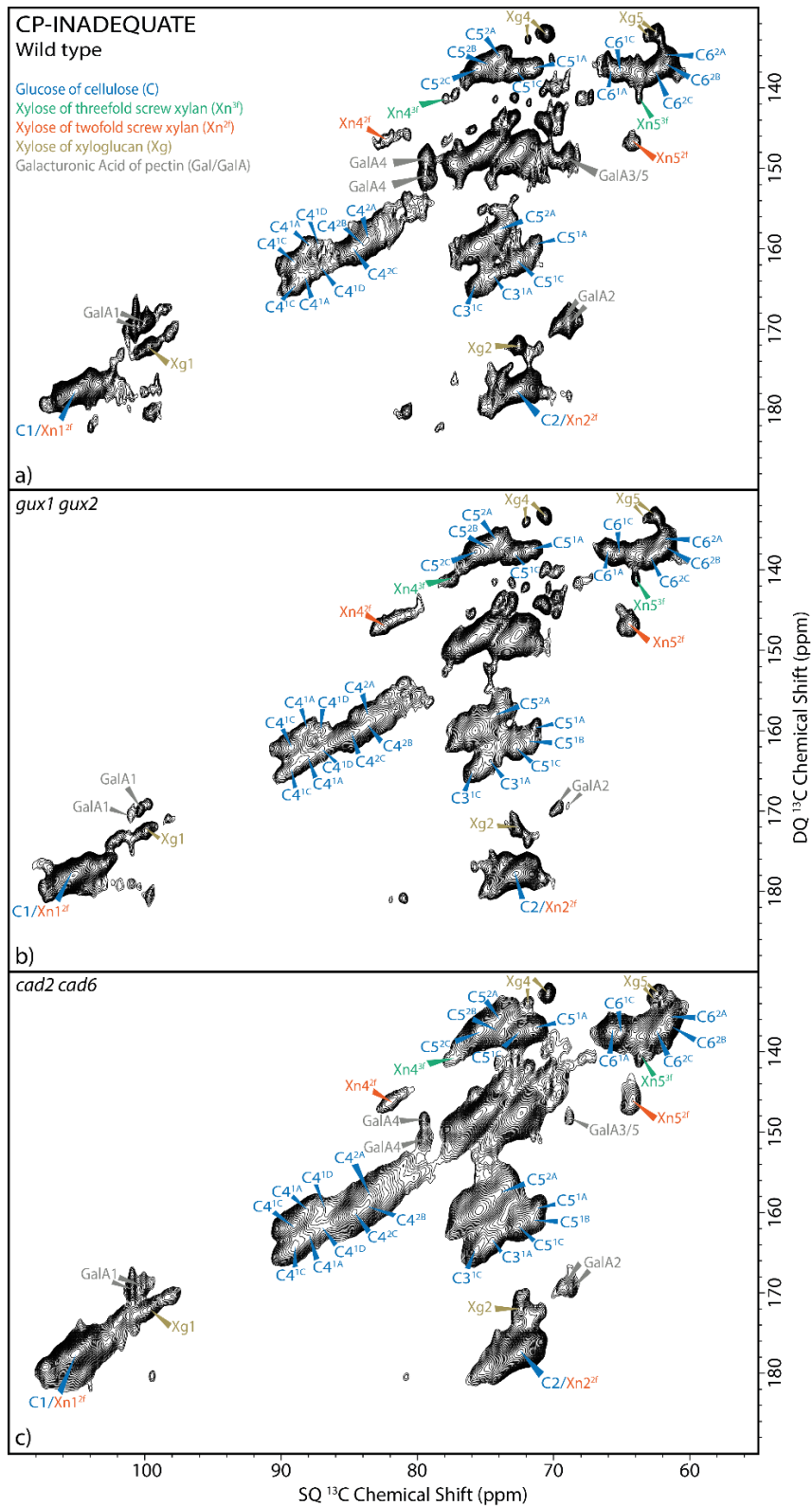
Figure 4.6: *gux1 gux2* and *cad2 cad6* have two-fold xylan and cellulose environments with similar chemical shifts to wild type. [13]C CP-INADEQUATE spectra of a) wild type, b) *gux1 gux2* and c) *cad2 cad6* Arabidopsis never-dried basal stems. Carbons of selected cell wall polysaccharides are labelled, based on comparison with literature values in (Simmons et al., 2016). The spectra are Fourier transformed with exponential line

broadening of 50 Hz in the SQ dimension and 90°-shifted squared sine bell processing in the DQ dimension.
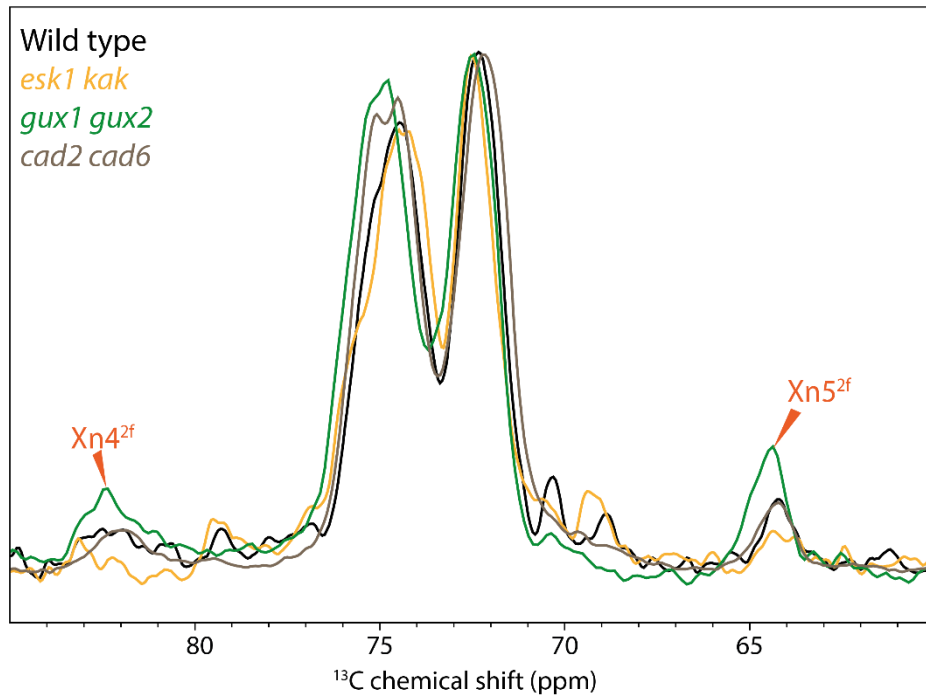
Figure 4.7: The *gux1 gux2* mutant has an increased amount of two-fold xylan. The DQ=145-148 ppm sum projections of ^{13}C CP-INADEQUATE spectra are shown. The slices are scaled so the tallest peak (which arises from carbon 2 in many polysaccharides) is the same height.

of xylan results in a larger amount of two-fold xylan in the *gux1 gux2* mutant. Previously, two-fold xylan formation has been shown to be dependent on the presence of cellulose, and the two-fold xylan is within a nanometre of the cellulose surface (Simmons et al., 2016). The presence of two-fold xylan signals in *gux1 gux2* and *cad2 cad6* suggests that xylan binds to cellulose in these mutants. The proximity of xylan and cellulose will be investigated later in this Section.

To compare the different cellulose environments in each genotype, the cellulose sub-domains in the CP-INADEQUATE spectra were assigned based on previous publications (Dupree et al., 2015; Wang et al., 2016b), Figure 4.6. The environments are split into two groups, domain 1 and 2 (Dupree et al., 2015). The differences in domain 1 and 2 chemical shift are thought to be largely due to differences in carbon 6 hydroxymethyl conformation (Phyo et al., 2018). Each domain is split into sub-domains, by a letter, e.g. C^{1A}, and is labelled based on positional similarity with the data in Section 6.2.2. In wild type Arabidopsis there are three domain 1 cellulose sub-domains, resolvable in the carbon four (C4) region, C4^{1A}, ^{1C} and ^{1D}, Figure 4.6a. There are three domain 2 sub-domains, resolvable at carbon 6 (C6), C6^{2A}, ^{2B} and ^{2C}. In *gux1 gux2* and *cad2 cad6*, there is an additional resolvable domain 1 sub-domain, C^{1B}, in addition the C4^{1D} peaks seem to be more intense, Figure 4.6b and 4.6c. Broadly speaking, the cellulose sub-domains seem to be similar in number and chemical shift between wild type and the two mutants, however the relative intensity of domain 1 sub-domains may be increased in the *gux1 gux2* and *cad2 cad6* mutant. In short mixing time PDSD spectra, the C^{1B} cross-peaks are present, suggesting there is relatively less C^{1B} in wild type. The short mixing time spectra are not shown, due to an incomplete set of experiments for all three genotypes, due to experimental time restrictions. A more detailed investigation of the number of cellulose domains and their chemical shifts would benefit from analysis of 30ms CP-PDSD spectra of each genotype.

Previously, it was observed that the loss of two-fold xylan in *eskimo1* correlates with alterations in the relative proportions of cellulose domains (unpublished data), thus the increased amount of two-fold xylan in *gux1 gux2* may also affect the proportions of cellulose domains. To assess the relative amounts of the cellulose domains between wild type and the two mutants, a sum projection of the cellulose carbon 5 and 6 region of the CP-INADEQUATE spectra between DQ= 135.4-140.3 ppm was produced, Figure 4.8. In addition to wild type, *gux1 gux2* and *cad2 cad6*, the *eskimo1 kaktus* mutant was also
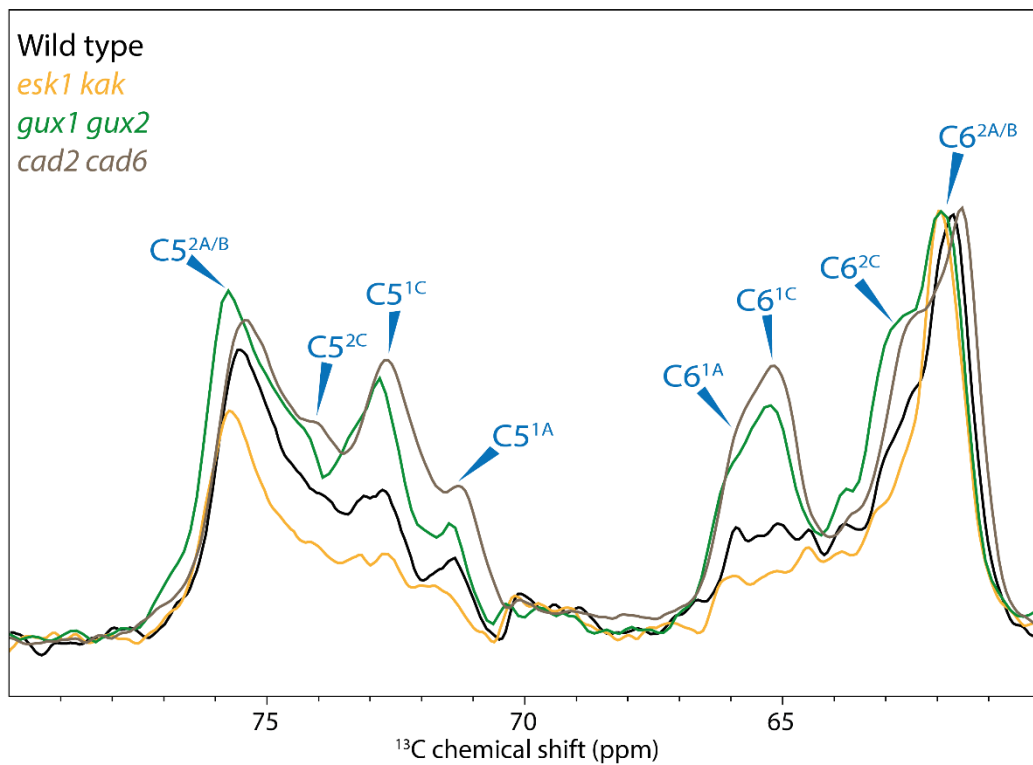
Figure 4.8: Relative proportions of different cellulose environments vary in *gux1 gux2* and *cad2 cad6*. DQ=135-140 ppm sum projection taken from ^{13}C CP-INADEQUATE spectra of wild type and *esk1 kak*, *gux1 gux2* and *cad2 cad6* mutant stems. All projections are scaled to the largest carbon 6 peak.

analysed in the same way. Each sum projection is scaled to the largest peak, C6$^{2A/B}$. The loss of two-fold xylan in *esk1 kak* correlates with relative reductions in the amount of C5 and C6$^{1A/C}$. In the *gux1 gux2* mutant there are increases in the relative amounts of C5 and C6$^{1A/1C}$. Since *gux1 gux2* has a larger amount of two-fold xylan (Figure 4.7), these changes suggest there is a correlation between the amount of two-fold xylan and the conformation of cellulose glucan chains. However, as can be clearly seen in Figure 4.8, the *cad2 cad6* mutant has a relatively increased amount of C5 and C6$^{1A/1C}$, just like the *gux1 gux2* mutant. The *cad2 cad6* mutant has the same amount of two-fold xylan as wild type, Figure 4.7, so if the cellulose conformation correlates solely with the amount of two-fold xylan, the *cad2 cad6* should have the same relative amounts of cellulose sub-domains as wild type. This suggests that the cellulose structural changes may correlate with other changes in the molecular architecture of the *gux1 gux2* and *cad2 cad6* mutants, such as the interactions between cell wall components.

To compare the interactions of xylan with cellulose in wild type and the *gux1 gux2* and *cad2 cad6* mutants, 1500 ms mixing time CP-PDSD spectra were performed on each mutant, Figure 4.9. At this mixing time intermolecular cross-peaks are observable, and these indicate that carbons in different molecules are less than approximately 10 Å apart in the cell wall (Kang et al., 2019). Each spectrum is normalised to the intensity of the C1/Xn1 self-cross-peak at 105 ppm. In all genotypes there is evidence of two-fold xylan interacting with cellulose, as has been previously reported in wild type (Simmons et al., 2016). There are cross-peaks between the acetate methyl (Ac^{Me}) of xylan and C4$^{1C/2A-C}$, C6$^{1A/2A-C}$. In addition, there are cross-peaks between Xn4^{2f} and C4^{1C}, C6$^{1C/2A/B}$. These results confirm that these mutations do not affect the binding of xylan to cellulose in a two-fold screw conformation.

To compare the interactions of xylan with lignin in wild type and the *gux1 gux2* and *cad2 cad6* mutants, the 1500 ms CP-PDSD spectra were examined for cross-peaks between lignin and xylan, Figure 4.9. In the wild type there are cross-peaks between the lignin methoxyl (lignin $^{CH_3}O^-$) and Xn4^{2f}, C4$^{1C/2A-B}$, as well as Xn1/C1, suggesting some lignin is within 10 Å of xylan in Arabidopsis, as has been reported recently (Kang et al., 2019). Interestingly, in both the *gux1 gux2* and *cad2 cad6* mutants, these cross-peaks between the lignin $^{CH_3}O^-$ and polysaccharide backbone carbons are greatly reduced in intensity, suggesting a reduction in the proximity of lignin and xylan.

a)

1500 ms CP-PDSD
Wild type
gux1 gux2
cad2 cad6

Glucose of cellulose (C)
Xylose of twofold screw xylan (Xn^{2f})
Methoxyl group of lignin ($Lignin^{CH_3O-}$)
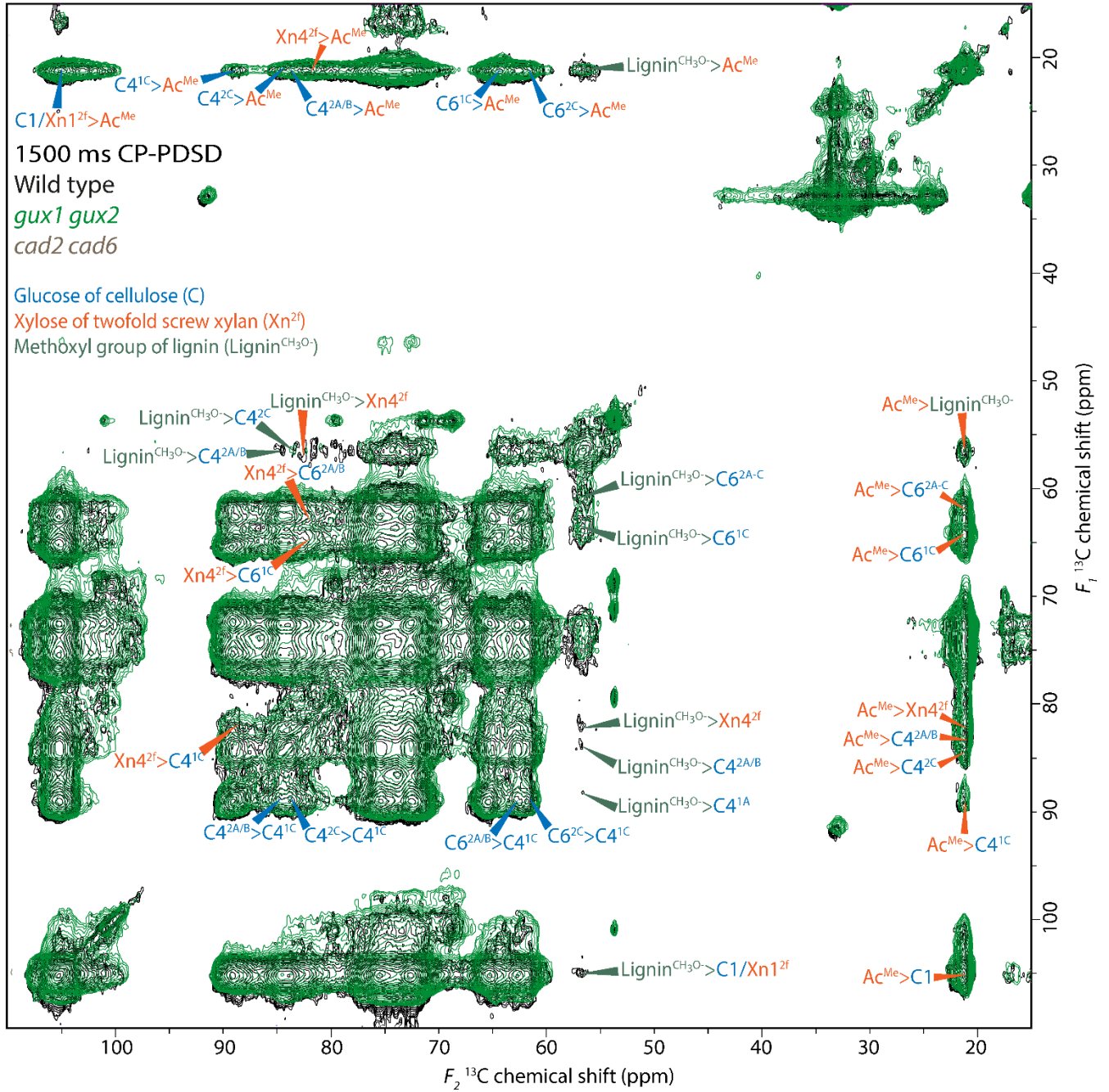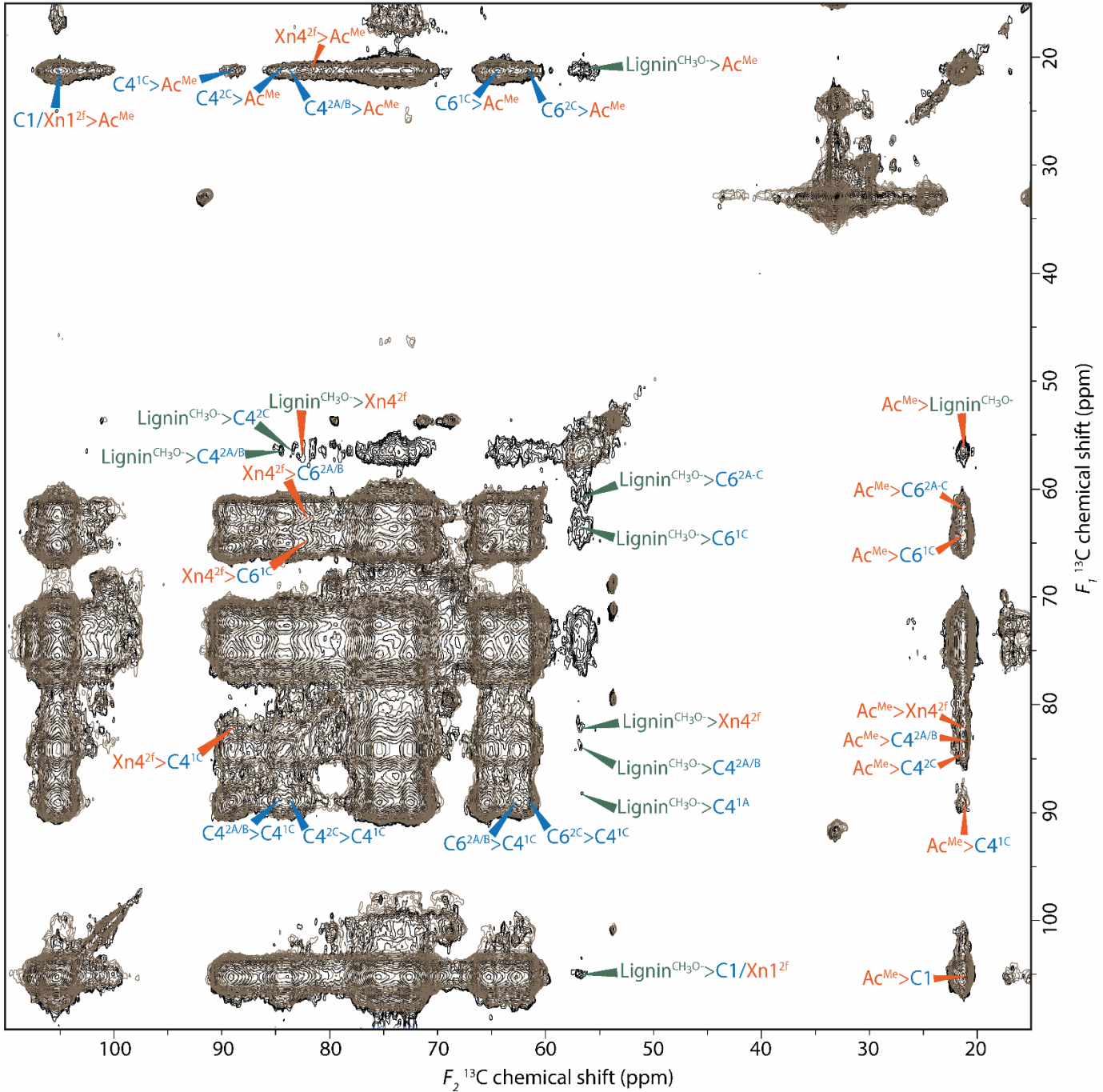
Figure 4.9: The interactions between cell wall constituents vary between wild type, *gux1 gux2* **and** *cad2 cad6*. a) The carbohydrate region of long (1500 ms) mixing time CP-PDSD spectra of wild type and *gux1 gux2* never-dried basal stems are overlaid. b) The carbohydrate region of 1500 ms mixing time CP-PDSD spectra of wild type and *cad2 cad6* are shown. Cross-peaks that indicate intermolecular contacts between xylan, cellulose and lignin are labelled. The spectra are scaled so the C1>C1 cross-peak at 105 ppm is the same height. The spectra are Fourier transformed with exponential line broadening of 50 Hz in the F_2 dimension and 90°-shifted squared sine bell processing in the F_1 dimension.

97

In all three genotypes there are also cross-peaks between the Ac^{Me} and lignin $^{CH_3}O^-$. To determine if these cross-peaks are relatively weaker in the *gux1 gux2* and *cad2 cad6* mutants, a sum projection of the Ac^{Me} region F_1=18-22 ppm was made, Figure 4.10a. Each projection is scaled to the Ac^{Me} self-peak at 21.4 ppm. In the *gux1 gux2* and *cad2 cad6* mutants, the cross-peak from lignin $^{CH_3}O^-$ to Ac^{Me} is weaker than in wild type, suggesting that there is a reduced interaction between xylan and lignin. The lignin $^{CH_3}O^-$ to Ac^{Me} cross-peak is weaker in *cad2 cad6* than *gux1 gux2*. This is likely due to a reduced lignin content in the *cad2 cad6* mutant (Sibout et al., 2005). This can be seen in the quantitative one-dimensional DP spectra shown in Figure 4.10b, where the lignin $^{CH_3}O^-$ peak is weaker in *cad2 cad6*. A sum projection of the region F1= 55-57 ppm from each genotype was also produced, Figure 4.10c. Each projection is scaled to the lignin $^{CH_3}O^-$ self-peak at 56.5 ppm. The cross-peak from Ac^{Me} to lignin $^{CH_3}O^-$ is also weaker in *gux1 gux2* and *cad2 cad6*, however this is less significant than the decreased intensity of the lignin $^{CH_3}O^-$ to Ac^{Me} cross-peak in Figure 4.10a. These results suggest that lignin and xylan have significantly reduced proximity in *gux1 gux2* and *cad2 cad6*, and thus both these mutants are affected in the interaction of lignin and xylan. It would be useful to investigate the *atr2*, *lac4* or *4cl1* reduced lignin content mutants with this method to examine if there are normal levels of lignin:xylan proximity, this was not done due to a lack of experimental time at the National 850 MHz NMR Facility.

4.2.4   Potential lignin-carbohydrate complexes are extractible from birch and wild type Arabidopsis

In this chapter it has been found that the *gux1 gux2* and *cad2 cad6* mutants have similar phenotypes of a reduction in cell wall recalcitrance, an increase in the enzyme accessibility of xylan and a reduction in the *in muro* interaction of xylan and lignin. Collectively, these results provide genetic evidence that lignin-glucoronate ester mediated xylan:lignin cross-linking is important for cell wall recalcitrance. To conclude that such cross-links are important for cell wall recalcitrance it is necessary to identify the proposed structure of glucuronic acid substitution on xylan covalently linked to lignin. Although there is some indirect evidence for the lignin-glucuronate ester, see Section 1.3.2, there is not yet convincing solution-state NMR data demonstrating this structure exists. In the literature, such cross-linked xylan and lignin have been referred to as lignin carbohydrate complexes or LCCs (Giummarella et al., 2019), and for simplicity will be
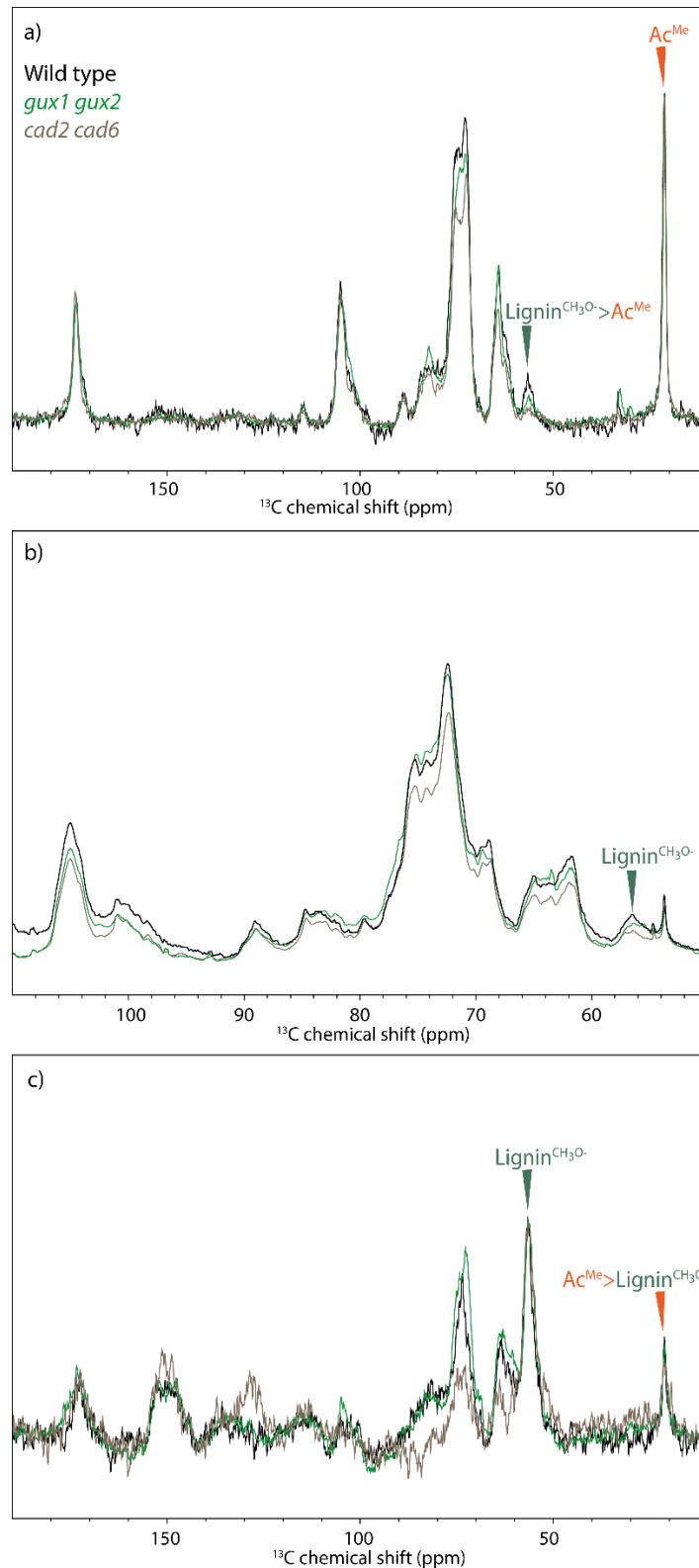
Figure 4.10: Lignin and xylan are less tightly associated in the *cad2 cad6* and *gux1 gux2* mutants. a) F_1=20-22 ppm sum projection of 1500 ms ^{13}C CP-PDSD spectra of wild type and the *cad2 cad6* and *gux1 gux2* mutant. b) 20 s DP one-dimensional spectra of each sample, scaled to number of scans of each spectra. c) F_1=55-57 ppm sum projection of 1500 ms ^{13}C CP-PDSD spectra of the above mutants. In a) and c) each spectra is scaled so that the self-peaks, in a) 21.4 ppm and in b) 56.6 ppm, are the same height.

referred to as LCCs in this Section.

DMSO can be used to extract acetylated xylan, and it is possible that lignin-glucuronate esters would also be preserved in such an extraction (Busse-Wicher et al., 2014). Previous unpublished results suggest that DMSO extractions of holocellulose, with the intention of extracting pure xylan, solubilise a combination of xylan, lignin and glucomannan, with higher temperatures releasing a lower purity xylan (Busse-Wicher., personal communication). It was hypothesised that high temperature DMSO extractions solubilise LCCs. Due to the limited amount of Arabidopsis material available, material from birch was used as a starting point to extract these potential LCCs. AIR was produced from ball-milled birch chips from Sigma Aldrich. To extract LCCs the AIR was depectinated with ammonium oxalate and partially delignified with peracetic acid to produce holocellulose. A high temperature ($80°C$) DMSO extraction was then performed on the holocellulose. The LCC extract was then buffer exchanged by gel filtration into pH 5.5 ammonium acetate to enable enzymatic digestion.

The LCC extract was then digested with *Aa*GH10 (the β-1,4 xylanase used in enzyme accessibility experiments, Section 3.2.1) and a CE4 acetyl esterase. *Aa*GH10 can cut the xylan backbone when the -1 subsite is unsubstituted, but can accommodate substitutions in the +1 subsite (Rantanen et al., 2007). Acetyl esterases from carbohydrate esterase family 4 can remove acetyl substitutions from singly substituted xylosyl residues (Neumüller et al., 2015). To separate any LCC which would be hydrophobic because of the lignin moiety, the oligosaccharides produced were then separated by solid-phase extraction with a C-18 column, which has previously been shown to separate grass hydrophobic ferulated arabinose structures from acetylated xylan oligosaccharides (Feijão, 2016). The column was washed with water and 10%, 20% and 99.8% ethanol. The oligosaccharides in each wash were then labelled with ANTS and separated and visualised by polysaccharide analysis by carbohydrate electrophoresis (PACE), a gel-based technique for structural characterisation of oligosaccharides. Figure 4.11a shows the PACE gel of the ANTS-labelled oligosaccharides in each fraction. In the H_2O wash, based on co-migration with a xylo-oligosaccharide ladder, the main products are xylose, xylobiose and xylotriose. In the 10% ethanol fraction there are higher degree of polymerisation (DP) oligosaccharides, such as xylotriose, xylotetraose and xylotetraose with a glucuronic acid substitution because they do not comigrate with the standards. In the 20% ethanol fraction there is a smear of ANTS labelled oligosaccharides that do not

Figure 4.11: Potential LCCs can be purified from digests of DMSO extracts of birch. High temperature DMSO extracts were digested as described. The products of digestion were separated by solid-phase extraction with a C-18 column, using water and different concentrations of ethanol. a) The oligosaccharides from each C-18 fraction were labelled with ANTS and visualised by PACE. b) The oligosaccharides in the 20% ethanol fraction were de-esterified and then digested further with GH11 xylanase and/or GH115 α-glucuronidase to characterise the oligosaccharides.

resolve into bands. This effect has previously been reported to occur when xylo-oligosaccharides are acetylated (Busse-Wicher et al., 2014). The 20% ethanol fraction also elutes the hydrophobic ferulated arabinose (Feijão, 2016). Therefore the presence of the xylo-oligosaccharides in the 20% ethanol fraction suggests these oligosaccharides are relatively hydrophobic.

Further glycosyl hydrolase digestions were used to characterise the oligosaccharides in the 20% ethanol fraction, Figure 4.11b. The 20% ethanol fraction was alkali treated to remove any acetyl substitutions or other esterified moieties. There are three main oligosaccharides released by alkali treatment. The three oligosaccharides migrated ahead of the xylo-oligosaccharide ladder, suggesting they are glucuronosylated xylo-oligosaccharides. Digestion with GH11 reduced the intensity of the two higher DP oligos, suggesting they are xylo-oligosacchrides. Digestion with the GH115 α-glucuronidase shifted the oligosaccharide migration so that the three oligosaccharides co-migrate with the xylo-oligosaccharide ladder, demonstrating that the oligosaccharides are glucuronated. Co-digestion of the oligosaccharides with GH11 and GH115 digests the three oligosaccharides. Thus the hydrophobic 20% ethanol fraction contains glucuronated oligosaccharides that are released by alkali-treatment. In Figure 4.11b the oligosaccharides are labelled after the conventions in (Fauré et al., 2009), where X means an unsubstituted xylosyl residue and U means a glucuronated xylosyl residue. The position of the U in the oligosaccharides is unknown but is labelled as $XU(X)_n$ for simplicity. These oligosaccharides could be esterified to lignin in an LCC via glucuronic acid.

Solution NMR could provide unambiguous identification of the structures, and has previously enabled unambiguous investigation of a glucomannan LCC structure (Nishimura et al., 2018). The NMR experiments described here were performed by Dr Katherine Stott. Experiments were initially performed on the 20% ethanol fraction from birch wood, but the lack of ^{13}C enrichment made multidimensional experiments unworkable. To enable solution state NMR structural characterisation of the 20% ethanol fraction the same enzyme digestion and solid-phase extraction procedure was performed on the high temperature DMSO-extraction of 99% ^{13}C-enriched Canadian poplar, purchased from IsoLife. The 20% ethanol fraction from this sample was then dried and resuspended in fully deuterated DMSO and characterised using solution-state NMR. Figure 4.12 shows the heteronuclear single quantum coherence (HSQC) spectrum of the
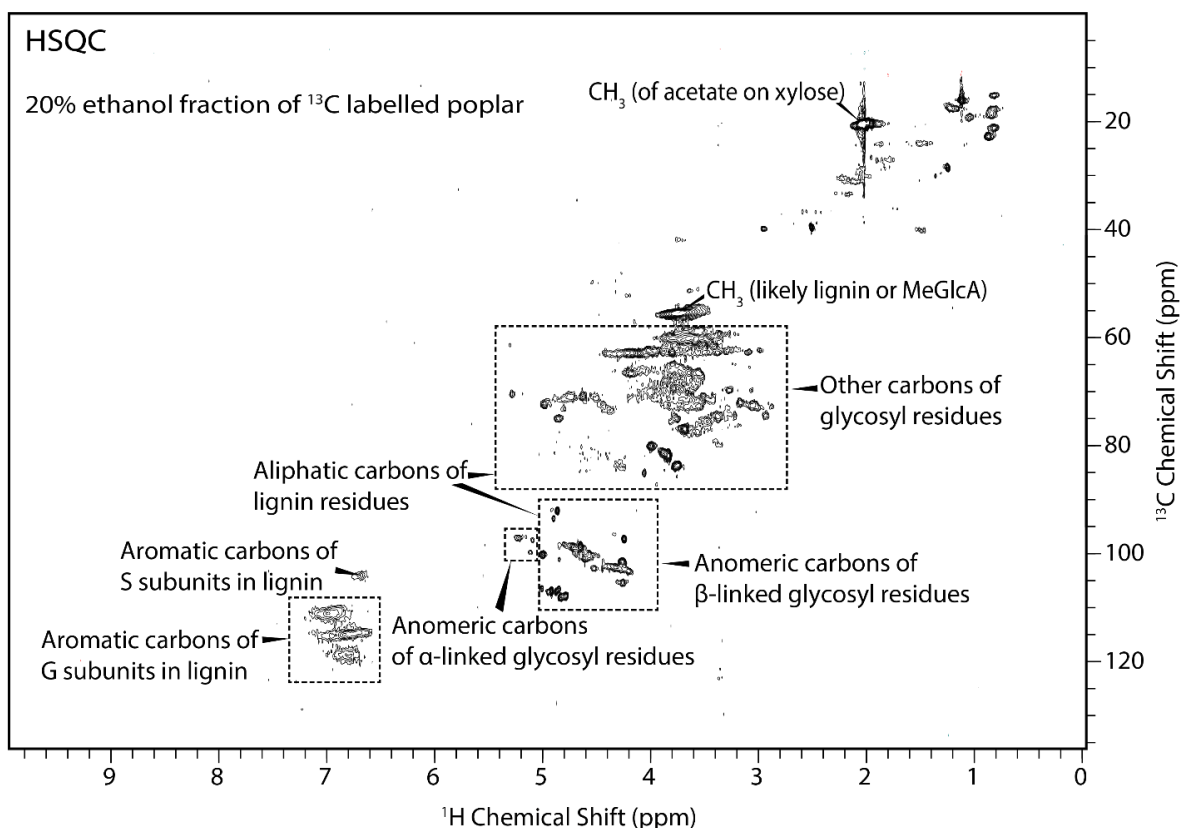
Figure 4.12: The 20% ethanol fraction LCC contains lignin and carbohydrates. The HSQC spectra of the 99% [13]C enriched Canadian poplar sample is shown. The sample was prepared using the same enzyme digestions and solid-phase extraction procedure as for birch. Areas of the spectra are labelled and assigned based on the following publications (Busse-Wicher et al., 2014; Kim and Ralph, 2014; 2010).
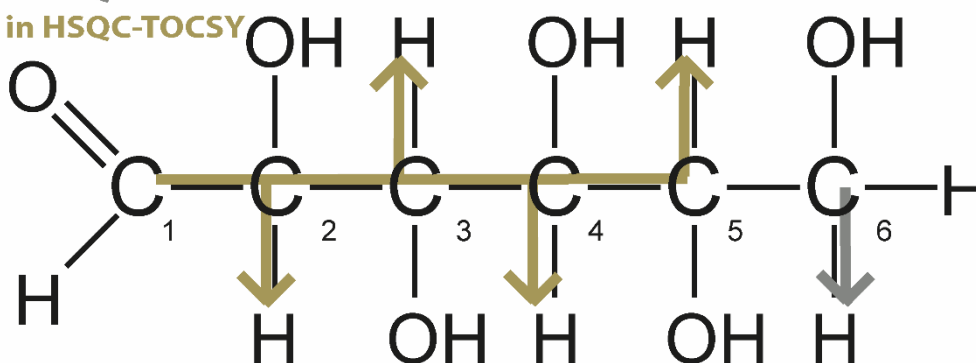
Figure 4.13: HSQC versus HSQC-TOCSY experiments. A non-cyclic glucose is shown, with carbons and protons labelled 1 to 6. Illustrative examples of magnetisation transfer between nuclei in specific NMR experiments is shown by coloured arrows. The HSQC only shows cross-peaks between ^{13}C nuclei and directly bonded ^{1}H, while the HSQC-TOCSY shows cross-peaks between ^{13}C nuclei and up to all other ^{1}H in the glucose (Cavanagh et al., 2007).

103

20% ethanol fraction. In the HSQC experiment, magnetisation is transferred through a single covalent bond, and thus the HSQC spectrum shows cross-peaks between carbons and their covalently linked protons only, see Figure 4.13. The aromatic carbon-proton cross-peaks at 105-125ppm to 6-8 ppm are consistent with the presence of lignin in this sample, Figure 4.12 (Dou et al., 2018). The carbon-proton correlations in the region 55-110 ppm > 3-5 ppm are consistent with the presence of xylo-oligosaccharides, and the aliphatic carbons of lignin (Busse-Wicher et al., 2014; Dou et al., 2018). The ^{13}C chemical shifts of the carbon 1 (C1) of glycosyl residues are shifted downfield (higher numerical value) than other glycosyl carbons, due to the magnetic field effect of the two covalently bonded oxygen atoms. In addition, the ^{1}H shift of proton 1 (H1) is indicative of the α or β conformation of glycosyl residues. In Figure 4.12 three clear C1H1 signals with a ^{1}H shift over 5 ppm can be seen, these are likely to be indicative of α-linked glucuronic acid or the α-linked reducing end of xylo-oligosaccharides. The majority of C1H1 cross-peaks have a ^{1}H shift of less than 5 ppm, indicative of a β-linkage if they form part of a glycosyl residue. These signals suggest the 20% ethanol fraction contains both lignin and polysaccharides, and thus perhaps an LCC.

To group the pairs of protons and carbon in the polysaccharide region into different residues, a heteronuclear single quantum coherence total correlation spectroscopy (HSQC-TOCSY) experiment was performed. In this experiment, magnetisation is transferred by J-coupling, but through any number of covalent bonds in a spin system, thus this experiment shows cross-peaks from each carbon to, potentially, every proton in the same glycosyl residue, see Figure 4.13. Thus the HSQC-TOCSY can be used to group the pairs of protons and carbons identified in the HSQC into a "spin system" of multiple pairs of directly linked protons and carbons. Figure 4.14 shows a restricted region of the HSQC spectrum overlaid with the HSQC-TOCSY spectrum. Most peaks are labelled xCnHn where x is the name of the residue (labelled with a single letter from a-k) and n is the number of the carbon and proton, according to Figure 4.13. Due to the lack of a H2BC experiment, the n numbers except for 1 and 5 do not indicate confirmed relative positions of the different carbons within the glycosyl residues, but enable differentiation of different ^{1}H-^{13}C pairs and probable relative positions (Busse-Wicher et al., 2014). The assignments are summarised in table 4.1. Full assignment of glycosyl residues was complicated by overlapping signals. To more confidently assign ^{1}H-^{13}C pairs into a spin system, a 3D HSQC-TOCSY was performed, and selected strips from this experiment are shown in Figure 4.14, with intramolecular cross-peaks labelled.

a)

HSQC
HSQC-TOCSY

20% ethanol fraction of ^{13}C labelled poplar

Residue a/b/c
Residue d/e/f
Residue g/h/
Residue j/k/

^{13}C Chemical Shift (ppm)

^{1}H Chemical Shift (ppm)

b)

3D-HSQC TOCSY

^{1}H Chemical Shift (ppm)

^{1}H Chemical Shift (ppm)

105
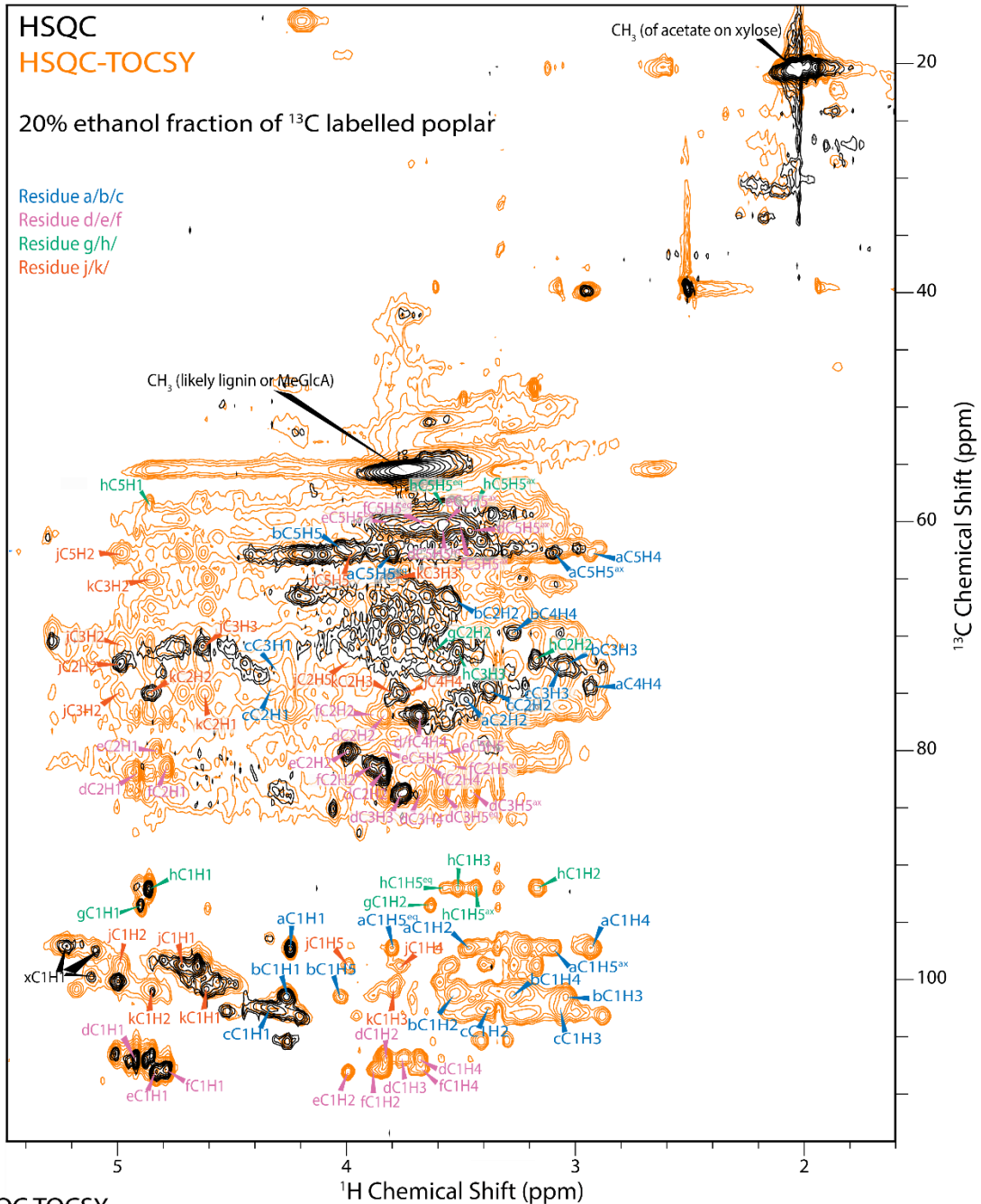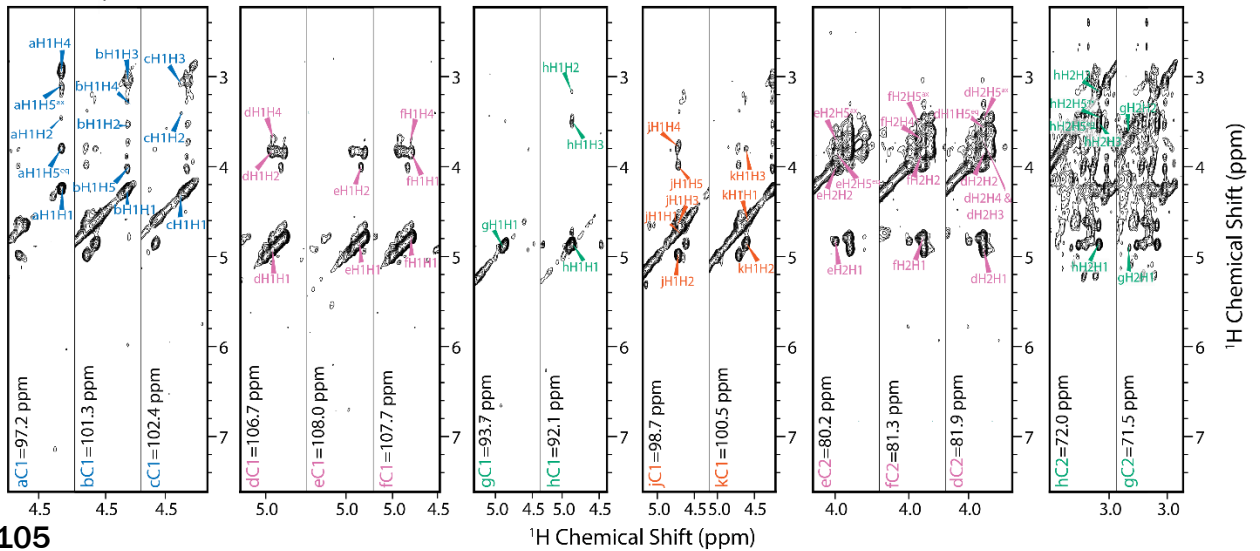
Figure 4.14: Characterisation of 20% ethanol fraction of ^{13}C Canadian poplar LCC. a) The carbohydrate region of the HSQC and 2D HSQC-TOCSY spectra of the 20% ethanol fraction are overlaid. b) Selected areas of the 3D HSQC-TOCSY spectra are shown. The ^{13}C and ^{1}H shifts are labelled based on their connectivities shown in the spectra.

Table 4.1: Chemical shift assignments of residues in 20% ethanol fraction of ^{13}C Canadian poplar LCC.

Residue	^{13}C or ^{1}H Chemical Shift (ppm)											
	C1	H1	C2	H2	C3	H3	C4	H4	C5^{eq}	H5^{eq}	C5^{ax}	H5^{ax}
a	97.31	4.24	75.53	3.47	-	-	74.48	2.93	62.70	3.80	62.70	3.09
b	101.45	4.26	66.90	3.53	72.76	3.04	69.85	3.27	62.36	4.03	-	-
c	102.53	4.33	74.43	3.37	72.70	3.07	-	-	-	-	-	-
d	106.96	4.92	81.84	3.84	83.63	3.75	76.89	3.68	60.65	3.58	60.78	3.46
e	108.03	4.83	80.10	3.99	-	-	76.89	3.68	60.07	3.79	60.07	3.56
f	107.89	4.79	81.41	3.88	-	-	-	-	60.14	3.65	60.14	3.52
g	93.65	4.90	71.64	3.63	-	-	-	-	-	-	-	-
h	92.06	4.86	72.09	3.17	71.44	3.51	-	-	-	-	-	-
j	98.87	4.71	72.43	4.99	70.82	4.63	75.04	3.76	62.74	3.99	-	-
k	100.76	4.61	74.95	4.86	65.02	3.80	-	-	-	-	-	-

Four of the assigned spin systems are likely to be β-linked pentoses, residues a, d, e and f. Each of these residues has five assigned carbons (or less) and one carbon in the 60-65 ppm range with two distinct ^{1}H shifts, confirming that the residues contain a CH_2 carbon. These results suggest they are pentosyl residues. In addition, all four residues have a H1 chemical shift of less than 5 ppm, which is suggestive of β-D-linked glycosyl residues (Peña et al., 2016; Busse-Wicher et al., 2014). The PACE analysis of the 20% ethanol fraction from birchwood would suggest that these are xylosyl residues, though PACE analysis of this Canadian poplar sample would be helpful to confirm that this sample has a similar composition. Unusually, the C1 shifts of residues d, e and f are high, at between 107-108 ppm, whereas xylosyl residues in DMSO normally have chemical shifts of between 96-101 ppm (Dou et al., 2018). Such a high C1 shift is normally indicative of arabinofuranose residues, however, arabinofuranose H1 have a shift above 5 ppm (Peña et al., 2016), so the identity of these residues is unclear. Based on the C1 assignments of residue b and c, b and c could also be β-linked glycosyl residues. However, only three ^{1}H-^{13}C pairs were assigned to residue b, thus it is unclear whether it is a pentose or hexose. For residue b, five ^{1}H-^{13}C pairs were assigned, this could indicate it is a hexose, as there are no carbons which clearly have two protons, however another ^{1}H-^{13}C cross-peak could be significantly overlapped.

Residues g and h have only 3 carbons assigned. Their C1 shifts could be consistent with an α carbon in a lignin β-O-4 linkage, though the H1 shifts are a little higher than previously published data (Dou et al., 2018). Residues j and k could also be β-linked glycosyl residues, however their assignments are incomplete, so it is unclear if they correspond to hexose or pentose residues. In addition, their H2 (and for residue j H3) shifts are unusually high for glycosyl residues, and are more consistent with aliphatic carbons in lignin (Dou et al., 2018). The xC1H1 shifts are very similar to previously reported assignments for glucuronic acid (Busse-Wicher et al., 2014; Kim and Ralph, 2014), however, due to overlapping signals in the HSQC-TOCSY (2D and 3D), it was not possible to group other ^{1}H-^{13}C cross-peaks into the same spin system.

To investigate the relative proximity of each residue to each other and to lignin carbons within the sample, nuclear overhauser effect spectroscopy experiments were performed (NOESY) experiments were performed. In a NOESY experiment, magnetisation is transferred through space, and cross-peaks between pairs of protons, not necessarily in the same spin system, are visible in the spectra. However, very few cross-peaks were

visible (data not shown), even the intramolecular cross-peaks observed in the TOCSY experiments were mostly not observed. This suggests the experiment needs optimising to improve the NOE transfer, possibly by altering the solvent mix from pure DMSO or changing the temperature of the experiment. The assignment of these residues provides a basis to perform these experiments in the future and assign the relationship between glycosyl residues and lignin in this sample.

4.3   Discussion

In chapter 3 it was proposed that the large reduction in recalcitrance to enzymatic digestion of the *gux1 gux2* is due to changes in the molecular architecture of the cell wall. In this chapter, two hypotheses of the role of xylan glucuronic acid substitutions in cell wall molecular architecture were investigated. Firstly, it was proposed that Ca^{2+} chelation by xylan glucuronic acid substitutions is important for cell wall molecular architecture. The effect of Ca^{2+}-depleted growth conditions on recalcitrance to enzymatic digestion was investigated, but there seemed to be no effect. Secondly, the role of lignin:xylan cross-linking was investigated, by analysing various mutants of lignin synthesis. A lignin mutant, *cad2 cad6*, uniquely shared similar recalcitrance and xylan accessibility phenotypes of *gux1 gux2*. The *in muro* conformation of xylan, cellulose and the intermolecular interactions of xylan, cellulose and lignin were investigated. In *gux1 gux2* and *cad2 cad6*, interactions between xylan and lignin were significantly reduced. A method of extracting and identifying potential LCCs using PACE and NMR was developed, and may be applied to Arabidopsis in the future.

4.3.1   Calcium binding by glucuronic acid substitutions does not play a significant role in recalcitrance

Previously, it has been suggested that calcium-pectin interactions may have important effects on the digestibility of primary cell walls. Regions of de-methylesterified HG can bind Ca^{2+} ions and form egg box structures (Jarvis and Apperley, 1995). Ca^{2+} binding by HG may be an important mechanism of pectins binding to each other in the plant, and could restrict the extensibility and porosity of the cell wall (Levesque-Tremblay et al., 2015). In banana, infection by highly pathogenic *Fusarium* race 4 is associated with decreases in Ca^{2+} cross-linked homogalacturonan, while the less infectious *Fusarium* race 1 had little effect on this pectin epitope (Fan et al., 2017). This suggests removal of

Ca^{2+} cross-linked pectin by the fungus facilitates infection. In Arabidopsis, endogenous pectin methylesterases are induced during infection by *Pseudomonas syringae*, and this is important for resistance to infection (Bethke et al., 2014). These Arabidopsis pectin methylesterases produce de-esterified regions of HG, which may cross-link through Ca^{2+}, potentially preventing *P. syringae* infection.

Analogously, it was hypothesised that glucuronic acid substitutions of xylan may bind Ca^{2+} in the cell wall. This function could explain some of the phenotypes of the *gux1 gux2* mutant. Ca^{2+} binding by xylan on the cellulose surface could restrict the size of pores between xylan-coated cellulose microfibrils, preventing digestion by enzymes. This model could explain phenotypes of the *gux1 gux2* mutant: the dramatic reduction in recalcitrance, increase in accessibility of xylan to xylanases, and increase in cell wall porosity (chapter 3). To test this idea, wild type, *gux1*, *gux2* and *gux1 gux2* plants were grown hydroponically, with different concentrations of calcium in the growth medium. While the Ca^{2+}-depletion was sufficient to cause dwarfing, suggesting *in vivo* calcium concentrations were altered, there was no effect on the recalcitrance of the cell wall in any mutant tested. This suggests that changes in Ca"+ binding by xylan is not an important mechanism of secondary cell wall recalcitrance.

4.3.2 The *cad2 cad6* and *gux1 gux2* mutants share important phenotypes that suggest a biological function for lignin:xylan cross-linking

In the 1980s it was reported that some glucuronic acid substitutions of xylan in wood are esterified (Das et al., 1984), however the extreme extraction procedures in these methods has been suggested to generate artefacts (Mottiar et al., 2016; Giummarella et al., 2019) and the biological function of such lignin-glucuronate esters is unknown. As explained in Section 1.3.2, the proposed mechanism for such xylan-lignin cross-linking is the re-aromatisation of quinone methide lignification intermediates by polysaccharide nucleophiles, such as the carboxylate group of glucuronic acid (Terrett and Dupree, 2019). There is convincing evidence that polysaccharide nucleophiles do re-aromatise quinone methides. Recently, work on pine demonstrated with multidimensional NMR the presence of a mannan etherified to a β-O-4 linkage in lignin, which would likely be formed by the same mechanism (Nishimura et al., 2018). The identification of CE15 esterases, that can de-esterify synthetic lignin-glucuronate esters provides indirect evidence for the presence of such cross-linking (d'Errico et al., 2015). There is not yet

convincing solution NMR verification of the existence of these structures (Nishimura et al., 2018). These studies suggest that glucuronic acid may be esterified to lignin *in vivo*, but the specific identification of this linkage and its biological function have not been investigated (Giummarella et al., 2019; Terrett and Dupree, 2019).

Here some further progress was made towards identifying LCCs. A xylan-rich high temperature DMSO extract of birch wood was produced. A putative LCC was prepared by enzymatic digestion and solid-phase extraction. This approach was similar in principal to that in Nishimura et al., (2018), though the extraction method differed significantly. Characterisation of the oligosaccharides with PACE showed that the three released xylo-oligosaccharides all contained at least one glucuronic acid substitution, as evidenced by α-glucuronidase sensitivity. The oligosaccharides could form part of an LCC. Esterification to lignin would mean the normally acidic xylo-oligosaccharides are neutral and hydrophobic, enabling binding to the C-18 column. Thus the hydrophobicity of the lignin-glucuronate etser may explain how these otherwise acidic oligosaccharides elute in the higher percentage ethanol washes, which previously have been shown to elute hydrophobic compounds (Feijão, 2016). It would be useful to investigate whether other LCC extraction techniques, such as hot water or dioxane, can produce similar or improved results (Giummarella et al., 2019). It may also be useful to utilise other enzymes, such as GH3 β-1,4 xylosidases and more promiscuous acetyl esterases, to simplify the oligosaccharides.

To investigate the structure of this potential LCC, the same extraction was performed on ^{13}C enriched Canadian poplar. Multidimensional NMR experiments were performed on the LCC. Encouragingly, clear signals from lignin could be identified, enabled by ^{13}C enrichment. In addition some progress was made assigning carbons into glycosyl residues. Unfortunately, through space experiments, that could aid further assignment and provide evidence for cross-linking between the lignin and glucuronic acid, have not yet been successful. In the future, to improve the assignments of carbons into residues, H2BC experiments could be used to provide complementary data to the HSQC-TOCSY. To improve the NOESY data, higher temperatures or an altered solvent mix could be used.

To investigate the idea of glucuronic acid mediated lignin:xylan cross-linking, the *cad2 cad6* mutant was studied. The *cad2 cad6* mutant has lignin with γ aldehyde groups instead of γ hydroxyls (Sibout et al., 2005). The aldehyde group alters the chemistry of the lignification. Instead of re-aromatisation by nucleophilic attack the now acidic proton

on carbon 7 is eliminated (Kim et al., 2003). Thus, the lignin of *cad2 cad6* should be unable to form lignin-glucuronate esters and would be predicted to have the same recalcitrance and xylan accessibility phenotypes of the *gux1 gux2* mutant (Terrett and Dupree, 2019). Previously, the effect of CAD mutation has been investigated in multiple species. In Wassilewskija Arabidopsis, the *cad2 cad6* mutant has cellulose that is almost twice as digestible as wild type, though the recalcitrance of xylan was not measured (Anderson et al., 2015). CAD mutation or downregulation in grass species such as maize (Fornalé et al., 2012) and *Brachypodium* (Bouvier d'Yvoire et al., 2013) also causes increases in cellulose digestibility. A CAD RNAi-downregulated poplar shows small decreases in the recalcitrance of xylan (Van Acker et al., 2017).

In this chapter it was found that the *cad2 cad6* mutant has very similar phenotypes to the *gux1 gux2* mutant. Specifically, when saccharified with CTec2, around 60% of the Xyl in both mutants is released, along with 80% of the Glc in *gux1 gux2* and likely all of the Glc in *cad2 cad6*. In addition, a GH10 xylanase (and acetyl esterase and α-glucuronidase) could digest approximately 25% of the xylan in both mutants in the absence of other enzymes. This is a dramatic 800% increase in the enzyme accessibility of xylan compared to wild type. This shows extensive digestion can occur without any pretreatment beyond ball-milling. The cellulose accessibility was unchanged in both *cad2 cad6* and *gux1 gux2*. The similarity of these two mutants supports the idea that re-aromatisation of the quinone methide intermediates of lignification is an aspect of cell wall molecular architecture that is important for recalcitrance to enzymatic digestion.

cad2 cad6 has pleiotropic lignin phenotypes. As well as the increase in hydroxycinnamladehydes, *cad2 cad6* has a 24% reduction in Klason lignin content compared to wild type, and lignin composed of 13% H units compared to 2% in wild type (Vanholme, personal communication). Reductions in lignin content have been reported to result in reduced recalcitrance (Van Acker et al., 2013). An increased proportion of H units has also been shown to reduce recalcitrance and result in more extractible hemicelluloses (Shi et al., 2016). In addition *cad2 cad6* has approximately double the proportion of S units as wild type (Vanholme, personal communication). An increased proportion of S units may reduce the length of lignin polymers (Stewart et al., 2009), but doesn't seem to have an effect on recalcitrance (Shi et al., 2016; Mansfield et al., 2012).

To control for the effects of these lignin phenotypes, four other lignin mutants were analysed in addition to *cad2 cad6*. *lac4* has a 13% decrease in Klason lignin content,

111

and a wild type lignin composition (Berthet et al., 2011). The *4cl1* mutant has a 26% decrease in acetyl bromide lignin content and 30% decrease in thioglycolic acid lignin content and an increased proportion of S units (Li et al., 2015; Van Acker et al., 2013). The *atr2* mutant has a 6% decrease in acetyl bromide lignin content, and 10% of *atr2* lignin is composed of H units (Sundin et al., 2014). The *ref8 med5a/b* has a 12% reduction in thioglycolic acid lignin content, 27% reduction in acetyl bromide lignin content and 11% increase in Klason lignin content, and its lignin is 95% composed of H units (Bonawitz et al., 2014). These methods for assessing lignin content do not correlate perfectly, and clearly have genotype-dependent differences due to different lignin compositions or proportions of interunit linkages (Fukushima et al., 2015; Moreira-Vilar et al., 2014). Thus it is unknown if these different mutants have the same amount of lignin, as this is not yet a fully measurable property.

Three of these controls, *lac4*, *4cl1* and *atr2*, were indistinguishable from wild type in terms of their cell wall recalcitrance and the enzyme accessibility of xylan and cellulose. This suggests that the reduction in lignin content, altered S:G ratio and small increase in H units of *cad2 cad6* are not the cause of the *cad2 cad6* phenotypes, and that it is the increase in hydroxycinnamaldehydes that causes the *cad2 cad6* phenotype. This supports the idea that cross-linking between glucuronic acid and lignin is important for cell wall recalcitrance. However, the *ref8 med5a/b* mutant also had somewhat similar phenotypes to *gux1 gux2*. 47% of *ref8 med5a/b* Xyl and 82% of *ref8 med5a/b* Glc was released during CTec2 saccharification. Although this is approximately 25% less Xyl than was released from *gux1 gux2* or *cad2 cad6*, the difference was not statistically significant. The *ref8 med5a/b* also had statistically significant increases in the percentage of Xyl and Glc released during xylan and cellulose accessibility experiments. Thus *ref8 med5a/b* has similar recalcitrance and xylan enzyme accessibility phenotypes to *gux1 gux2* and *cad2 cad6;* in addition it has an increase in cellulose enzyme accessibility.

How does this affect our interpretation of the shared phenotypes of *gux1 gux2* and *cad2 cad6*? The similarity of *cad2 cad6* and *ref8 med5a/b* could suggest that the increase in H units in *cad2 cad6* has an effect on recalcitrance. However, the *atr2* has a similar proportion of H units as *cad2 cad6*, but no effect on the phenotypes measured here (Sundin et al., 2014; Vanholme, personal communication). The *ref8 med5a/b* also has reduced average molecular weight of lignin (Bonawitz et al., 2014), and some CAD

pine mutants also have this phenotype (Tolbert et al., 2014; Dimmel et al., 2001). However, this is likely to be because coniferaldehyde cannot couple with coniferyl alcohol, i.e. coniferaldehydes cannot couple with normal conifer lignin (Kim et al., 2003). However hydroxycinnamaldehydes can couple with each other and *p*-coumaroyl alcohol, suggesting the Arabidopsis *cad2 cad6* mutant may not have an effect on lignin polymer size (Ralph et al., 2019). In addition, in a previous study the molecular weight of *ref8 med5a/b* lignin was greater than the *fah1 pC4H::F5H* transgenic Arabidopsis which has shorter lignin molecules due to being composed almost entirely of S units (Shi et al., 2016). Despite having larger lignin molecules than *fah1 pC4H::F5H*, *ref8 med5a/b* had significantly reduced recalcitrance in this study, while *fah1 pC4H::F5H* did not. Thus it is unlikely that similarities between *cad2 cad6* and *ref8 med5a/b* are due to reductions in lignin molecular weight. It would be useful to measure the lignin molecular weight of the plants used in this study.

The *ref8 med5a/b* plants have been shown to have a reduced proportion of β-O-4 linkages (Bonawitz et al., 2014) and a Medicago *c3h* mutant increased proportions of dibenzodioxocin, phenylcoumaran and resinol structures (Ralph et al., 2006). Thus very high H lignin plants have a reduced amount of β-O-4 linkages which can be re-aromatised by polysaccharide nucleophiles, and instead they re-aromatise quinone methides more frequently with internal nucleophiles (demonstrated by the presence of more phenylcoumaran, resinol and dibenzodioxocin linkages). Therefore, the *ref8 med5a/b* mutant may actually share a reduction in the amount of xylan:lignin cross-linking with the *gux1 gux2* and *cad2 cad6* mutants. These results are consistent with the hypothesis that xylan:lignin cross-linking is important for cell wall recalcitrance.

4.3.3 Xylan:lignin cross-linking may affect cell wall porosity and "shield" xylan from saccharification enzymes

In chapter 3 it was found that the *gux1 gux2* has an increased enzyme accessible pore volume in the cell wall. This likely contributes to the reduced recalcitrance and increased enzyme accessibility of xylan in this mutant, as larger pores give saccharification enzymes a greater surface area of accessible substrate (Grethlein, 1985). Reductions in lignin:xylan cross-linking may contribute to this phenotype: xylan bound to lignin and cellulose may cause the spaces between microfibrils to be filled with xylan and lignin, reducing the porosity of the cell wall.

However, the increases in pore size are comparable to those observed in other studies, while the decrease in recalcitrance of *gux1 gux2* is noticeably greater (Assis et al., 2018). This suggests that the molecular architecture changes go beyond the small increase in cell wall porosity. Recently, solid-state NMR has been used to investigate the *in muro* interactions between xylan and lignin (Kang et al., 2019). It was found that there are extensive interactions between lignin and xylan, suggesting that xylan may be coated by lignin. It is unknown what role covalent linkages may play in these non-covalent interactions, though Kang et al., (2019) speculated that covalent ferulate linkages are unimportant for the coating of grass xylan by lignin.

Here, we investigated the xylan:lignin interaction of Arabidopsis with solid-state NMR to see if it is altered in the *gux1 gux2* and *cad2 cad6* mutants. In wild type, it was found that there were cross-peaks between the lignin CH_3O- and acetyl substitutions on xylan, as well as two-fold xylan backbone carbons. This shows that some xylan and lignin are separated by less than 1 nm in the cell wall, as was recently also shown in Kang et al., (2019). The cross-peaks between xylan and lignin were also found in *gux1 gux2* and *cad2 cad6*, showing that some xylan is still closely interacting with lignin in these mutants. However a comparison of the strength of cross-peaks between xylan and lignin shows that these cross-peaks are significantly weaker in *cad2 cad6* and *gux1 gux2* than in wild type. This suggests that lignin is either relatively more distant from xylan, or that the sites of lignin:xylan contact are less frequent, or some combination of the two. This is the first report suggesting that xylan:lignin cross-linking affects non-covalent interactions between these two polymers in the cell wall. This may suggest that covalent cross-linking plays an important role in cell wall assembly, localising the lignin to specific sites in the cell wall during the polymerisation process. This change in cell wall molecular architecture may explain the dramatic recalcitrance phenotype of the *gux1 gux2* and *cad2 cad6* mutants. In wild type cell walls, the proximity between xylan and lignin probably indicates that xylan is partially coated by lignin (Kang et al., 2019), and thereby shielded from enzymatic hydrolysis. In *gux1 gux2* and *cad2 cad6*, where these interactions are weaker or less frequent, more xylan may be "naked" and accessible to enzymes.

The xylan:cellulose interaction is relatively unchanged in *gux1 gux2* and *cad2 cad6* compared to wild type. All three genotypes had two-fold xylan and three-fold xylan, without differences in the proportion of each conformation in the different genotypes. The

gux1 gux2 mutant had a relatively higher amount of two-fold xylan, but this is likely due to a general increase in xylan content (Chong et al., 2015) and chapter 3. In both mutants there are cross-peaks between two-fold xylan and C^1 and C^2 domains of cellulose in 1500 ms CP-PDSD spectra that show xylan interacts with cellulose. These results suggest that the xylan:cellulose interaction is relatively unchanged from wild type in these two-mutants. Previously it was shown that the loss of the xylan:cellulose interaction in the *esk1* mutant does not reduce cell wall recalcitrance (Lyczakowski et al., 2017; Grantham et al., 2017). Thus two-fold xylan formation does not seem to be important per se for the digestibility of secondary cell walls. Instead two-fold xylan formation seems to be important for the mechanical stability of vessels during water transport (Lefebvre et al., 2011), and potentially other properties.

The CP-INADEQUATE spectra suggest there are also some cellulose structural changes in *gux1 gux2* and *cad2 cad6*. Domain C^1 of cellulose has been shown, with direct measurements and DFT simulations, to have a carbon 6 hydroxymethyl in the *tg* conformation (Yang et al., 2018; Phyo et al., 2018), while domain C^2 has a *gg/gt* hydroxymethyl confromation. The *tg* conformation is thought to be induced by interchain hydrogen bonding, Figure 1.2, while the *gg/gt* conformation is thought to occur in the absence of an adjacent polysaccharide chain to induce the *tg* conformation (Jarvis, 2018). It has been proposed that the domain C^1 mainly occurs in the interior of microfibrils, where adjacent glucan chains can fix the hydroxymethyl *tg* conformation, while the *gg/gt* conformation may mainly occur on the surface of microfibrils. Both mutants have a relatively increased amount of domain C^1, as observed by changes in the relative heights of C5 and C6 peaks. This suggests that cellulose microfibrils in these mutants have relatively more glucan chains in the *tg* conformation. It has been proposed that xylan binding to the hydrophilic face may convert surface glucan chains from domain C^2 to domain C^1 through hydrogen bonding (Jarvis, 2018), see chapter 6. Supporting this interpretation, the *esk1* mutant, which has no two-fold xylan, and presumably no xylan hydrogen bonding to cellulose, has a relatively reduced amount of domain C^1. However, as discussed in the previous paragraph, *cad2 cad6* has a wild type amount of two-fold xylan, while *gux1 gux2* has an increased amount of two-fold xylan. Yet both mutants share a similar increase in the proportion of domain C^1. Thus it seems unlikely that changes in the amount of two-fold xylan are affecting the conformation of cellulose.

Why is there an increased number of tg/domain C^1 glucan chains in *cad2 cad6* and *gux1 gux2*? The increased distance between xylan and lignin in these mutants may affect the conformation of glucan chains in cellulose through an unidentified mechanism. *In silico* simulations of mono- and dilignols with cellulose suggest that lignin can stack onto the hydrophobic face of cellulose (Houtman and Atalla, 1995). The xylan:lignin interactions may have a secondary effect of localising lignin near to cellulose, enhancing the binding of lignin to cellulose. The lignin binding to cellulose may affect the conformation of surface chains, causing relative changes in the amount of different cellulose domains. Modelling of lignin on different faces of cellulose and the effect on hydroxymethyl conformations could be investigated with molecular dynamics. Alternatively, the binding of lignin and xylan together may be an important mechanism of separating cellulose microfibrils. When these interactions are reduced, microfibrils may interact with each other more frequently. If such an interaction occurs between hydrophilic faces, hydrogen bonding could alter the hydroxymethyl conformation of surface chains. Microfibril bundling could be monitored with AFM or cryoEM, though preliminary experiments on the size of cellulose macrofibrils in *gux1 gux2* suggest they are the same size as wild type (data not shown).

Chapter 5:   Engineering ferulate substitutions on Arabidopsis xylan to complement the recalcitrance phenotypes of the *gux1 gux2* and *cad2 cad6* mutants

5.1   Introduction

In previous chapters of this thesis, the recalcitrance phenotype of the *gux1 gux2* mutant, which lacks glucuronic acid on its xylan, was investigated. It was found that the loss of recalcitrance is partially due to an increase in the porosity of the cell wall, which contributes to an increase in enzyme accessibility of xylan, leading to up to 500% more xylan being digested and 50% more cellulose being digested in saccharification assays, see chapter 3.   The hypothesis of lignin-glucuronate esters was investigated as a molecular level explanation for secondary cell wall recalcitrance. To test this idea, the *cad2 cad6* mutant was investigated. This mutant has γ aldehydes on the majority of its monolignols instead of γ hydroxyls (Sibout et al., 2005).   The aldehyde group prevents nucleophilic attack by polysaccharides or water during β coupling of monolignols (Terrett and Dupree, 2019), Thus, in the *cad2 cad6* mutant, xylan might not cross-link to lignin in this way. Supporting this idea, *cad2 cad6* shares the reduced recalcitrance and increased xylan accessibility of the *gux1 gux2* mutant, while this effect is not observed in some other lignin mutants that were tested. Interestingly, the *in muro* proximity between lignin and xylan seems to be significantly decreased in these mutants when investigated using solid-state NMR. These data suggest *gux* or *cad* mutation has profound effects on the molecular architecture of the cell wall.

In grasses, it is thought that xylan:lignin cross-linking mainly proceeds through the participation of ferulate substitutions in lignification (Buanafina, 2009), Section 1.3.2. The proposed model of ferulate substitution synthesis is extremely complex, involving multiple genes in different cellular compartments, see the following paragraphs. Reconstitution of the ferulation pathway in Arabidopsis could enable us to test current models of ferulation synthesis. In addition, investigation of *gux1 gux2* or *cad2 cad6* mutant plants transformed to synthesise ferulated xylan would significantly strengthen the idea that a lack of xylan:lignin cross-linking is the cause of reduced recalcitrance in these mutants.

Ferulation of xylan is an incompletely understood and complex process, involving both cytosolic and Golgi-localised enzymes; see Figure 5.1 for a model of our current understanding of xylan ferulation. In plants, feruloyl-CoA is cytosolically synthesised from the general phenylpropanoid pathway (Ralph et al., 2004). The ferulic acid is then thought to be transferred to UDP-arabinofuranose by cytosolic enzymes in the BAHD family of acyltransferases (Mitchell et al., 2007; de Souza et al., 2018). The metabolite UDP-feruloyl arabinofuranose has not been successfully extracted from plants and the *in vitro* activity of a ferulation-specific BAHD enzyme has never been demonstrated, though the *in vitro* acyl CoA transferase activity of other BAHD enzymes has been shown (Moglia et al., 2016). It has been proposed that ferulated glucose is the ferulate donor
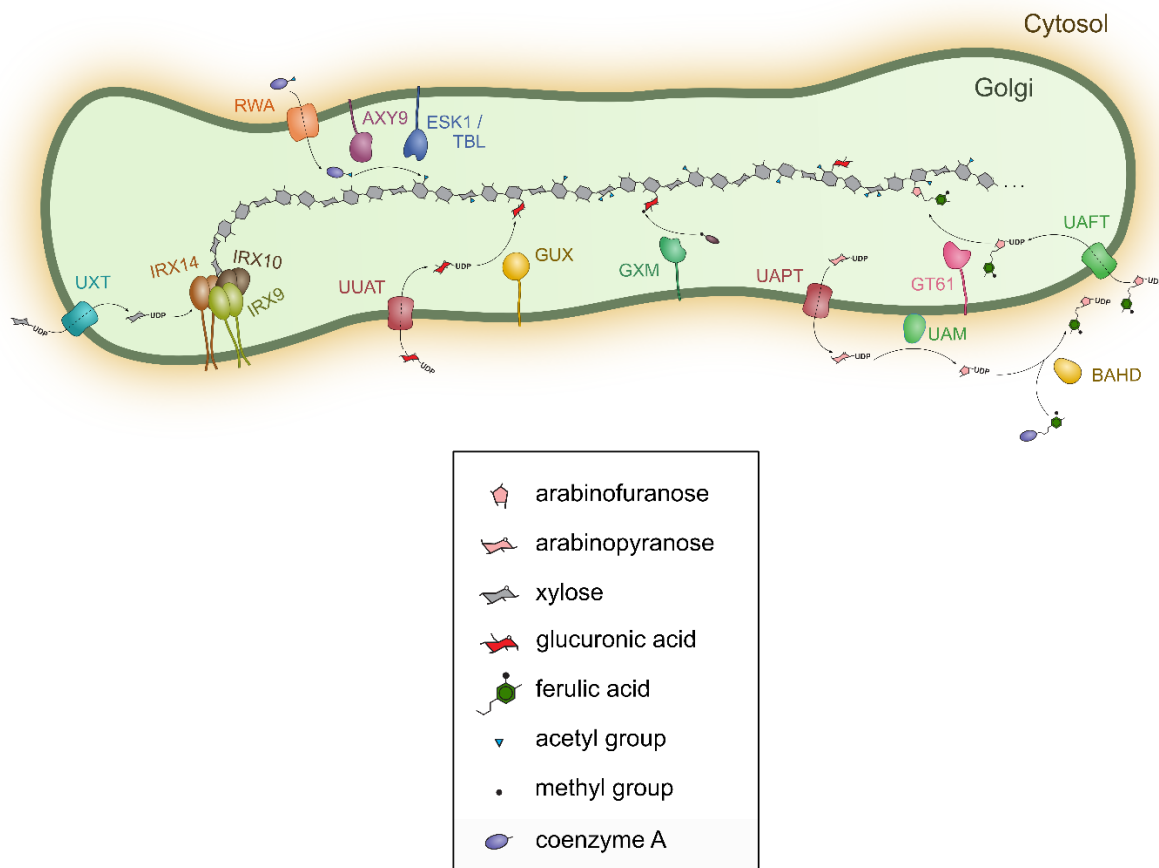
Figure 5.1: A model of ferulated xylan synthesis in the Golgi. Some proteins involved in xylan backbone extension, acetylation, glucuronidation and methylation are shown. UDP-GlcA imported through UUAT proteins is converted into UDP-Ara*p* in a two-step process not shown. The UDP-Ara*p* is then exported by a UAPT activity and UAM converts it to UDP-Ara*f*. The UDP-Ara*f* is thought to be ferulated by a BAHD, possibly while attached to the UAM protein. The ferulated UDP-Ara*f* is then thought to be imported by a UAFT transporter and then is thought to be transferred to the xylan backbone by a GT61. Adapted from an unpublished figure by Louis Wilson.

for the ferulation of UDP-arabinose (Obel et al., 2003), though this is inconsistent with the described thioester substrate specificity of BAHD acyltransferases.

Increasing the complexity of this process is the synthesis of UDP-arabinofuranose, derived from UDP-glucuronic acid (UDP-GlcA), a pyranose, in the Golgi apparatus (Rautengarten et al., 2017). UDP-GlcA is synthesised in the cytosol and it is imported into the Golgi lumen by a UDP-Uronic Acid Transporter (UUAT) (Saez-Aguayo et al., 2017). There, UDP-GlcA is decarboxylated to UDP-xylopyranose, which is then converted by an epimerase to UDP-arabinopyranose (UDP-Ara*p*) (Burget, 2003; Rautengarten et al., 2017). Finally, UDP-ArabinoMutase (UAM) enzymes convert the UDP-Ara*p* to UDP-arabinofuranose (UDP- Ara*f*) (Konishi et al., 2007). However, these enzymes are cytosolic. UAMs have been described to associate with the cytosolic face of the Golgi membrane (Delgado et al., 1998), and were originally thought to be Golgi localised (Dhugga et al., 1997). UAM proteins were found to strongly associate with the xylan synthase complex (XSC) members, IRX14 and 9 in wheat, and it has been suggested this facilitates substrate channelling of UDP-Ara*f* to arabinoxylan synthesis (Zeng et al., 2010). The BAHD genes may act on the UDP-Ara*f* when it is attached to the UAM proteins, since proteins of a similar size to UAMs are radioactively labelled by ^{14}C ferulic acid and ^{3}H arabinose in wheat suspension-cell cultures (Obel et al., 2003). If this is the case, UAM and BAHD enzymes could both interact with the XSC and act to channel the ferulated UDP-Ara*f* toward ferulated xylan synthesis. The mutarotation of arabinose and its ferulation requires two further UDP-sugar transporters (Rautengarten et al., 2017). The UDP-Ara*p* needs to be exported to serve as substrate to UAMs and the UDP-Ara*f* needs to be imported back into the Golgi. It is likely ferulated UDP-Ara*f* would be imported by a nucleotide sugar transporter of the UDP-Arabinose Furanose Transporter (UAFT) family (Konishi et al., 2007; Rautengarten et al., 2017).

Once ferulated UDP-Araf has been synthesised and is present in the Golgi lumen, it must be added to the xylan backbone by a glycosyltransferase (GT). Enzymes from family GT61 clade A and B add α-1,2/3 Ara*f* and β-1,2 Xyl substitutions to the xylan backbone (Anders et al., 2012; Voiniciuc et al., 2015; Zhong et al., 2018; Temple-Sanchez, personal communication). The rice GT61 enzyme Xylosyl Arabinosyl substitution of Xylan 1 (XAX1) has been published to be a xylosyl transferase that adds a xylosyl residue onto an arabinosyl residue on the xylan backbone, forming the $D^{2,3}$ substitution (Chiniquy et al., 2012). This xylosylated arabinose substitution is normally ferulated or coumaroylated.

119

However, the *in vitro* xylosyltransferase activity of this enzyme was very weak, the acceptor was unknown, and the proposed activity is significantly different to other reported activities for this clade, casting doubt on this hypothesis.

In the same paper it was also reported that there was a reduction in wall bound coumaric and ferulic acid in *xax1* rice stems (Chiniquy et al., 2012). The authors interpreted this result as evidence for a model where ferulation and coumaroylation occur in the Golgi, and that some arabinosyl acylation requires the xylosylation by XAX1 as a recognition site for an acyltransferase. In this model, the lack of the xylosylated arabinosyl substitutions reduces the acylation of arabinosyl substitutions, therefore the acylation must occur in the Golgi. However, unpublished work suggests that, in the *xax1* mutant, both the ferulated $D^{2,3}$ substitutions and unxylosylated ferulated arabinosyl susbtitutions are reduced (Feijão, 2016). The author concluded that XAX1 is an arabinosyl ferulate arabinosyl transferase, which acts directly on the xylan backbone. This is more consistent with other published data on the role of cytosolic BAHD enzymes in generating ferulate substitutions and GT61 activity acting directly on the xylan backbone, rather than extending arabinosyl side chain substitutions (de Souza et al., 2018; Feijão, 2016).

As described, in monocots, xylan ferulation is a complex pathway. In order to attempt to reconstitute this pathway in Arabidopsis, several activities need to be introduced in highly engineered plants, including NSTs, UAM, BAHD and XAX1.

The aims of this chapter were:

- To produce a genetic construct that expresses candidate ferulation genes in secondary cell wall synthesising cells of Arabidopsis and evaluate their expression
- To evaluate xylan structural changes in transgenic plants

5.2   Results

5.2.1   Gene and promoter selection and construct design

This project was executed in collaboration with Dr Henry Temple-Sanchez. The genes were selected by him, though a description of the rationale is provided here.

A potential ferulated arabinosyltransferase from the GT61 family, XAX1 (*Os02g22380*), was selected as a strong candidate, due to the general reduction in ferulated arabinosyl substitutions in the *xax1* mutant (Chiniquy et al., 2012; Feijão, 2016).

A putative BAHD feruloyl-CoA transferase was selected. The *Brachypodium distachyon* gene BAHD01 (*BdBAHD01/Bradi02g05480*) is an orthologue of the same clade as the *Setaria viridis* gene that, when silenced, reduces ferulation of xylan (de Souza et al., 2018). In the same paper, silencing this gene in *Brachypodium distachyon* also reduced the ferulate content of stems, but to a lesser extent than in *Setaria viridis*. *BdBAHD01* is also highly expressed in the first and last internode 60 days after germination, in mature leaves and in roots, suggesting it could be a good candidate for secondary cell wall synthesis (Sibout et al., 2017).   In addition, *BdBAHD01* co-expresses with a GT61 (*Bradi1g06560*) that is the closest ortholog of *OsXAX1* (Anders et al., 2012; Chiniquy et al., 2012). Additionally, the RICEFREND co-expression database shows that *OsXAX1* co-expresses with the rice orthologue of *BdBAHD01*, *Os01g09010* (de Souza et al., 2018), with a mutual rank of 14 (Sato et al., 2013), while the ATTED database shows that *OsXAX1* and *OsBAHD01* co-express with a mutual rank of 6 (Obayashi et al., 2018). The mutual rank metric is a measure of co-expression that that is more robust than correlation rank: "correlation rank is asymmetric, namely the rank of gene B from gene A is not the same as the rank of gene A from gene B. And thus, those two ranks are geometrically averaged, which we call Mutual Rank (MR)" (Obayashi et al., 2018). This BAHD has been silenced in rice, along with four other genes, causing reductions in ferulate content (Piston et al., 2010). These *in vivo* and co-expression data strengthen the case that GT61s and BAHDs from these specific clades co-operate together in the synthesis of xylan ferulate substitutions.

Since protein-protein interactions may be important for ferulation, the rest of the candidate genes were all selected from rice. To select rice *UAM* and *UAFT* candidates, *XAX1* was then used as bait in the RICEFREND and ATTED co-expression databases. The only *OsUAM* candidate from RICEFREND was *Os03g40270*, which co-expressed with a large mutual rank of 1740 (suggesting weak co-expression). In ATTED the mutual rank was much higher at 144, and this gene was the most highly co-expressed *OsUAM* candidate in ATTED. This protein has been expressed in insect cells and has *in vitro* UDP-arabinomutase activity (Konishi et al., 2007; Konishi et al., 2010). The *OsUAFT* gene

Os05g41480 co-expresses well with *OsUAM* in both databases, with mutual ranks of 22 in RICEFREND and 39 in ATTED.

These five genes were assembled into a binary vector (*EC62261*) using the GoldenGate cloning technology. Figure 5.2 shows the plasmid map with details of the order of transcriptional units and the promoters, tags and terminators in each transcriptional units. Given that xylan is abundant in dicot secondary cell walls, each gene was placed under the control of promoters from Arabidopsis genes involved in secondary cell wall synthesis. The use of the same promoter to express multiple genes increases the risk of silencing. Thus unique promoters that were thought to be secondary cell wall specific were selected for each gene. The promoters were selected based on compatibility with GoldenGate cloning (lack of *BsaI* and *BpiI* restriction sites), to enable use of the native promoter sequence. IRX3 and 5 are members of the cellulose synthase family (CesA) and form part of the secondary cell wall cellulase synthase complex (Taylor et al., 2003). Previously, both *pIRX3* and *pIRX5* have been effective for inducing expression of glycosyl transferases in secondary cell walls (Lyczakowski et al., 2017; Aznar et al., 2018). KORRIGAN is a plasma membrane endoglucanase that is expressed in both primary and secondary cell walls (Zuo et al., 2000; Bhandari et al., 2006; Vain et al., 2014). The *irx2* mutant is a mutant allele of *KORRIGAN*, suggesting KORRIGAN is expressed in secondary cell walls (Szyjanowicz et al., 2004). COBL4 is a protein expressed in secondary cell walls and is involved in cellulose synthesis (Roudier et al., 2005). IRX14 is a component of the secondary cell wall xylan synthase complex (Brown et al., 2007). 1.5 kb of the 3' end of the promoters of KORRIGAN, COBL4, and IRX14 were used as the promoters to drive expression of three of the ferulation genes. In addition to the five ferulation genes, a seed specific GFP selection marker, and a kanamycin resistance selection marker were included in the construct. Once produced, the construct was sequenced (data not shown) with pairs of primers to verify correct assembly. The construct was transformed into the GV3101 strain of *Agrobacterium tumefaciens*.

Several different Arabidopsis genotypes were transformed with *A. tumefaciens GV3101:: EC62261* using the floral dip method (Clough and Bent, 1998). Wild type was transformed. The *eskimo1* mutant was also used in this work. *esk1* is a mutant in a xylan acetyltransferase gene that is responsible for 60% of xylan acetylation, the major substitution of xylan in dicot plants (Grantham et al., 2017). Therefore, in *esk1* the xylan substitution level falls from 50% to around 20% of backbone xylosyl residues. Some
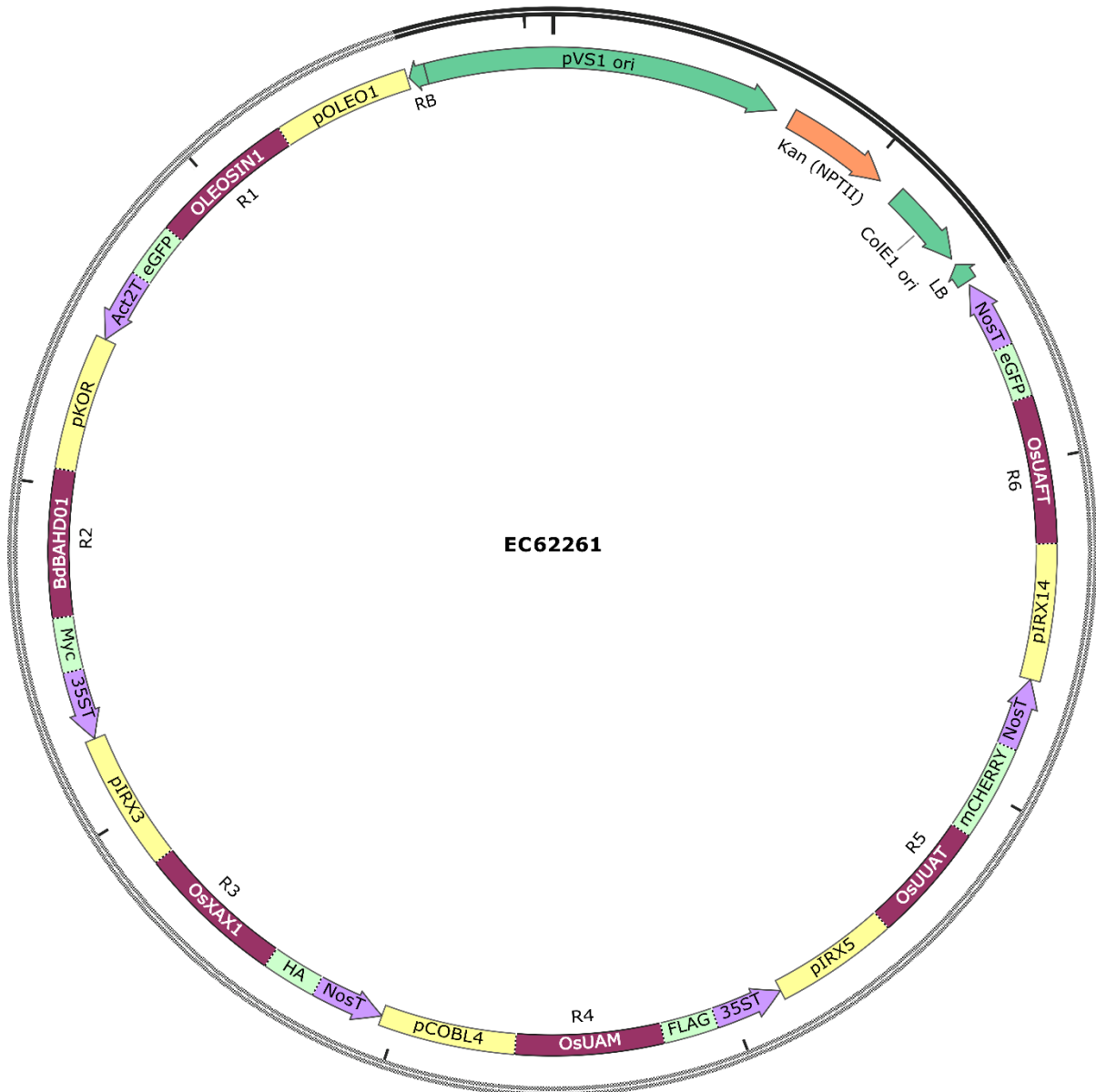
Figure 5.2: *EC62261* **plasmid map.** The promoters, CDS, tag and terminator for expression of the ferulation genes in the secondary cell wall are shown. The size of each DNA part is not shown to scale, but to enable the name of each part to be read.

GT61s have been shown to require this additional space on the xylan backbone to produce substantial amounts of arabinosylation (Temple-Sanchez, personal communication). The *cad2 cad6* and *gux1 gux2* mutants were also transformed. These genotypes may produce higher levels of ferulation than wild type. Like *esk1 kak*, *gux1 gux2* has a reduced substitution level compared to wild type. *cad2 cad6* may have elevated levels of ferulyol CoA, as some *cad* mutants accumulate feruloyl CoA derived metabolites (Van Acker et al., 2017). Feruloyl CoA is the substrate for BAHD enzymes. In addition, the transformants in the *cad2 cad6* and *gux1 gux2* background could make it possible to test if ferulation restores the recalcitrance phenotype of these mutants.

Ten T_1 lines per genotype were selected by seed fluorescence and sown on soil. T_2 seeds were sown on plates (by Dr Temple-Sanchez), and three mono-insertional lines were selected per genetic background (except for *esk1* due to a low transformation and germination rate), based on ~3:1 segregation of T_2 seed fluorescence. At least 3 homozygous T_2 plants per line were selected based on 100% T_3 seed fluorescence, and background genotype plants were selected based on the absence of seed fluorescence, to act as negative controls. Each line was named *genotype.61.n* where n corresponds to the line number.

5.2.2 The new selected secondary cell wall promoters drive expression of the proteins in secondary thickening stems

To test if the transgenic proteins are expressed in secondary thickening infloresence stems, total proteins were extracted from the thickening stems between the second and third internode of T_3 homozygous transgenic plants. Proteins were extracted from matching plants of the same genetic background as negative controls. The samples were then separated by SDS-PAGE. The gels were Stain-Free™ imaged to check that proteins had been extracted from each genotype, Figure 5.3. Although the exact amount of protein varied, the presence of bands in all samples indicated successful extraction of proteins with similar efficiency.

To test if the KORRIGAN and COBL4 promoters drove expression of the cytosolic proteins, BdBAHD01-Myc and OsUAM-FLAG, the proteins in the SDS-PAGE gel were transferred to a nitrocellulose membrane. α-Myc and α-FLAG Western blots were performed on the membranes and the chemiluminsecent and colourimetric images of the blots are

overlaid and shown in Figure 5.4. In the α-Myc Western blot, there are no bands in the negative controls, despite a similar loading of protein each lane, see Figure 5.3. In each transgenic line, except for *cad2 cad6.61.1*, there is a band that migrates between 55-70kDa, this migration is consistent with the predicted molecular weight of BdBAHD01-Myc at 62.3 kDa. In the *WT.61.n* lines there is an additional band between 35-55kDa, which could be a degradation product of BdBAHD01–Myc. This result confirms BdBAHD01-Myc was expressed in the inflorescence stems of these transgenic lines, except for *cad2 cad6.61.1*. In the anti-FLAG Western blot, there is also an absence of bands in the negative controls, but in all transgenic lines, except for *cad2 cad6.61.1*, a band that migrates between 35-55 kDa can be seen, consistent with the 41.5 kDa predicted molecular weight of OsUAM-FLAG. This shows OsUAM-FLAG was also expressed in inflorescence stems. The T_2 seeds of *cad2 cad6.61.1* plants were fluorescent, demonstrating successful transformation. It is possible that the construct has been silenced in the mature stems, leading to no protein production, this could be due to the large number of genes introduced, a multi-insertional line, or their genomic insertion location. These results show that the selected part of the Arabidopsis promoters of *KORRIGAN* and *COBL4* are effective for driving expression of monocot cytosolic proteins in this tissue.

To test the expression of the Golgi-localised OsXAX1, the same protocol was performed, however no expression was detected (data not shown). To test if this was due to inefficient protein extraction, a more efficient protein extraction method using ball-milling of flash-frozen stems, was used on two samples, *gux1 gux2* and *gux1 gux2.61.1*. The new protein extracts, and an increased amount of the previous protein extracts were separated by SDS-PAGE, and transferred to a nitrocellulose membrane, and expression of OsXAX1-HA was detected by an α-HA Western blot, Figure 5.5. In the first lane, a positive control of a tobacco microsome extract containing HA-tagged IRX14 has a weak band at the expected size. In the negative control lane, *gux1 gux2* high efficiency extract, there is a non-specific band between 25-35 kDa. In the *gux1 gux2.61.1* high efficiency extract lane there is an additional band between 100-130 kDa. The predicted molecular weight of OsXAX1-HA is 62 kDa, so this protein is approximately twice the molecular weight of the OsXAX1-HA monomer. There is also a faint band at the predicted molecular weight. Glycosyltransferases frequently dimerise in the Golgi, and this dimerisation has been observed to remain after denaturation in our lab, even with the use of reducing agents (Lyczakowski et al., 2017). Therefore, it is likely that this band corresponds to an
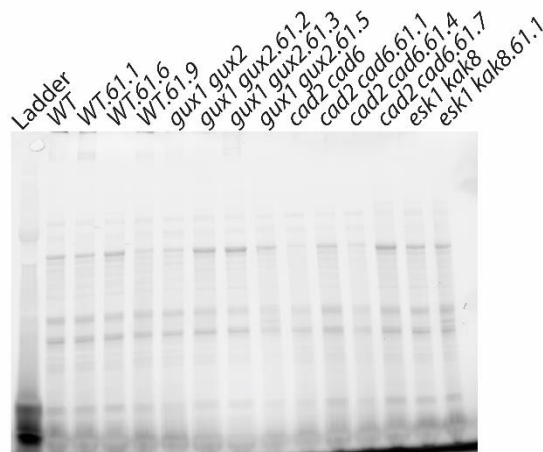
Figure 5.3: Similar amounts of protein were extracted from thickening Arabidopsis stems. The approximately 4cm section of stem between the second and third internode of two plants per line were crushed in Laemmli buffer to extract proteins. The proteins were denatured at 100°C for ten minutes. The extracted proteins were separated on a Stain-Free SDS-PAGE gel.
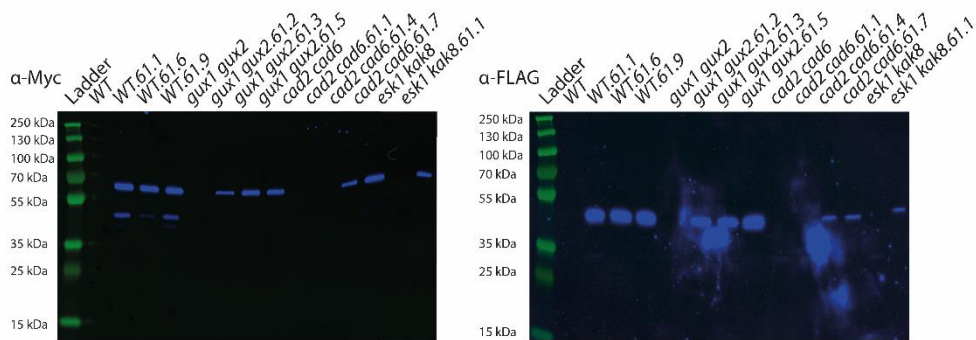
Figure 5.4: The cytosolic proteins BdBAHD01-Myc and OsUAM-FLAG are expressed in most transgenic lines. The proteins in the SDS-PAGE gel shown in Figure 5.8 were transferred to a nitrocellulose membrane and Western blots were performed, using either α-Myc or α-FLAG antibodies. HRP-linked secondary antibodies were used for visualisation and the chemiluminescent and colourimetric (to visualise the ladder) images are overlaid. The blots were imaged with a BioRad ChemiDoc imaging system. Predicted molecular weights: BdBAHD01 – 62 kDa, OsUAM – 42 kDa.

OsXAX1-HA dimer. In the lower efficiency extracts, when 5x more protein was loaded on the gel, there are also OsXAX1 bands in the transgenic lines, except for *cad2 cad6.61.1* and the negative controls. The *WT.61.1* sample is absent from this Western blot. There is a stronger monomeric band in the lower efficiency protein extracts. Therefore, the transgenic lines, other than *cad2 cad6.61.1*, express OsXAX1-HA.

To test if the other newly selected promoter, IRX14, was successful at driving expression of OsUUAT, expression of OsUUAT-eGFP was detected by Western blot with an α-GFP antibody, Figure 5.6. While the high efficiency extracts produced similar results to the other proteins in the construct, with bands of the expected size in the transgenic line but not the negative control, expression was not detected in the low efficiency extracts. This is likely due to the decreased amount of microsomal proteins, as can be seen in Figure 5.5. Unfortunately no more protein extract was available to perfom mCherry Western bots to test OsUAFT expression. To be able to provide stronger evidence that the two nucleotide sugar transporters are expressed, additional T_3 plants should be grown and protein expression should be tested by Western blot.

To provide additional evidence for the expression of the fluorescent protein-tagged nucleotide sugar transporters, confocal microscopy was performed on sections of thickening stem, with Dr Temple-Sanchez, Figure 5.7. For the wild type, *gux1 gux2* and *esk1 kak* stems, there was low GFP fluorescence and virtually undetectable mCherry fluorescence in the vasculature. In the transgenic lines of these three genotypes the vascular bundles have notably increased mCherry fluorescence. There also seems to be a small increase in the eGFP fluorescence of the vascular bundles which is consistent with the WB results, but there is significant background signal. In the *cad2 cad6* lines there is high background eGFP and mCherry fluorescence. This is likely due to the unique hydroxycinnamaldehyde composition of the lignin which results in greatly increased carbon-carbon double bonds, which can autofluoresce. There may be slightly increased mCherry and eGFP fluorescence in both the *cad2 cad6.61.4* and *cad2 cad6.61.7* lines, which expressed the other genes in the construct, as detected by Western blot.

Collectively, these results suggest that all lines, except for *cad2 cad6.61.1* express all five transgenes in secondarily thickening stem tissues.
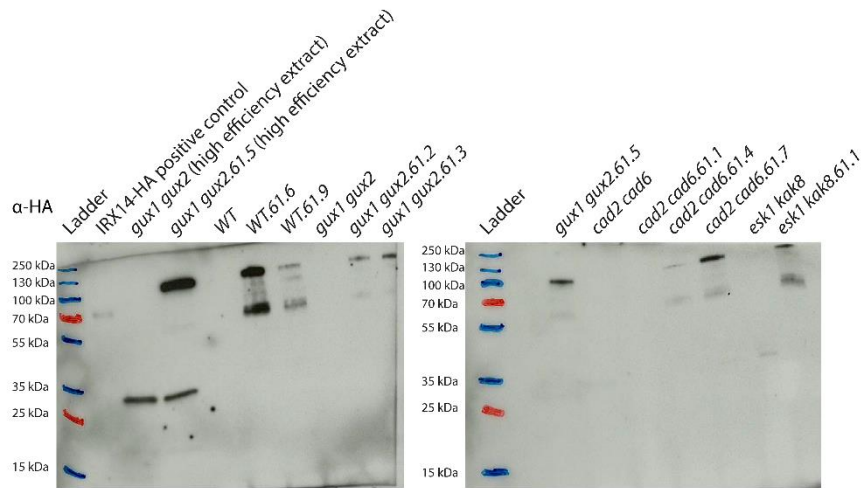
Figure 5.5: OsXAX1-HA is expressed in the stems of most transgenic lines. The proteins from an SDS-PAGE gel were transferred to a nitrocellulose membrane and an α-HA Western blot was performed. A HRP-linked secondary antibody was used for visualisation and the photographic film from an overnight exposure is shown. Predicted molecular weight: OsXAX1 – 62.0 kDa (monomeric), 124.0 kDa (dimeric).
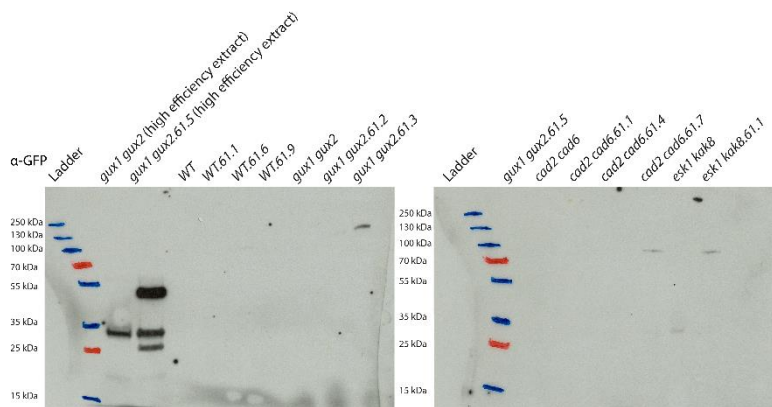
Figure 5.6: OsUUAT-GFP is expressed in *gux1 gux2.61.1*. The proteins from an SDS-PAGE gel were transferred to a nitrocellulose membrane and an α-GFP Western blot was performed. A HRP-linked secondary antibody was used for visualisation and the photographic film from an overnight exposure is shown. Predicted molecular weight: OsUUAT-GFP – 65.8 kDa.
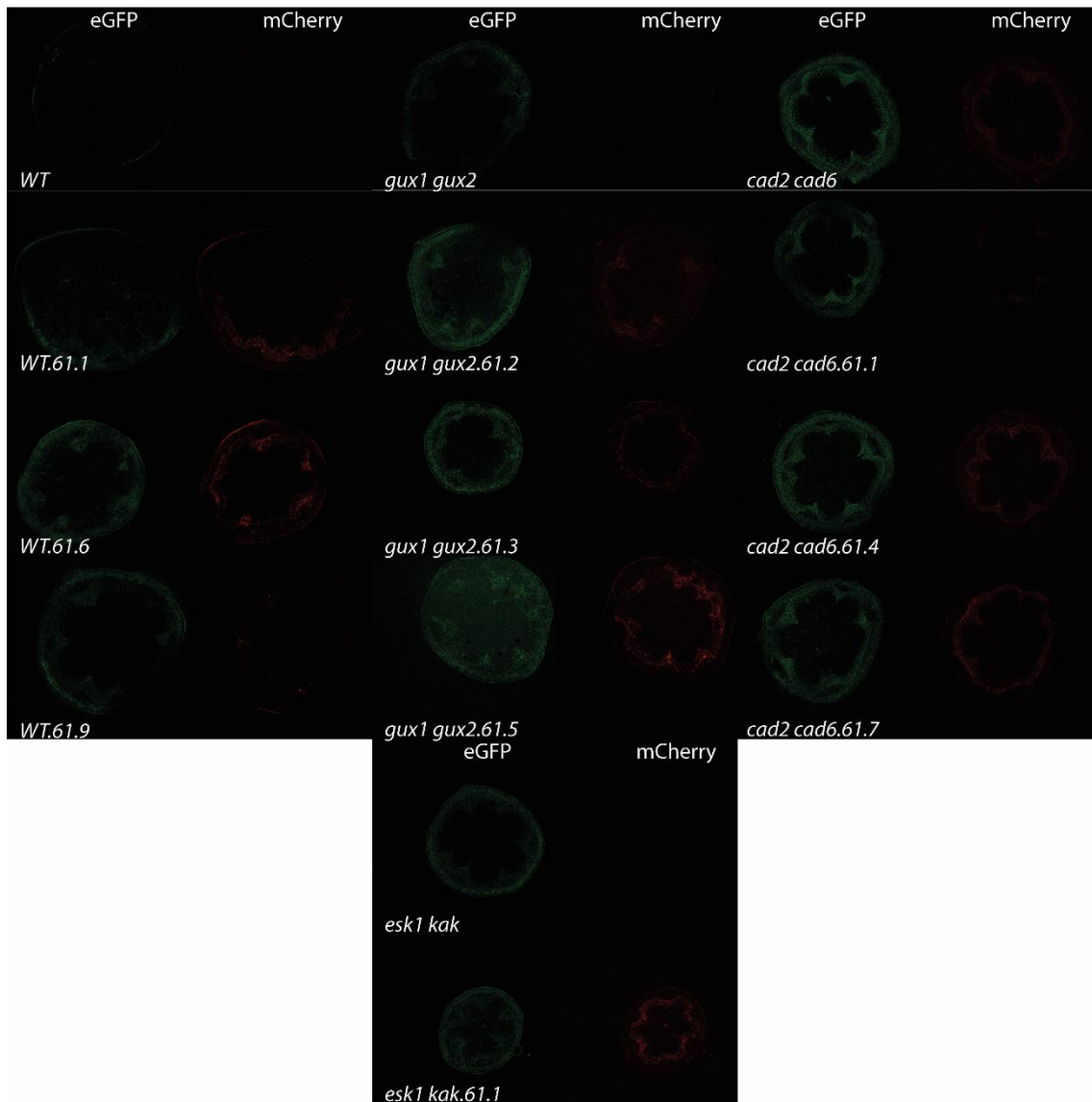
Figure 5.7: OsUUAT-GFP and OsUAFT-mCherry are expressed in vascular bundles. Confocal microscopy was used to image hand-sectioned stem slices that were between 5-10 μm thick. Each picture shows a single 0.52 μm slice from the z-stack. All the images were taken with identical parameters. Each image has been contrast enhanced identically in Adobe Photoshop. For the eGFP, the excitation laser had a wavelength of 488nm and transmission occurred in the range 500-520nm. For mCherry the excitation laser wavelength was 587nm and the transmission occurred in the range 600-620nm. Each image was sequentially acquired to prevent additional background signal from the other fluorescent proteins.

5.2.3 There are no xylan structural changes detectable in T$_2$ homozygous plants

The ferulation genes were successfully expressed in the transgenic lines. To test if these proteins were sufficient to cause ferulation of xylan, the xylan from homozygous T$_2$ plants was first analysed for arabinosylation. Alkali extraction would remove any ferulation, but the arabinosyl substitutions would remain after this extraction.

To analyse xylan arabinosylation, xylan can be digested with glycosyl hydrolases (GH) (Anders et al., 2012). The *Neocallimastix patriciarum* xylanase from family GH11 can digest the xylan backbone with one unsubstituted xylosyl residue in the +1 position, and two unsubstituted xylose residues in the -2 and -1 position (Vardakou et al., 2008). Thus, when acting on de-acetylated xylan the enzyme produces xylose (X), xylobiose (XX) and xylotetraose substituted with glucuronic acid or arabinose (XUXX or XAXX). Longer oligosaccharides will be produced if substitutions are separated by less than three xylosyl residues, e.g. XUXUXX, XAXUXX. As the only reported non-alkali labile substitutions on Arabidopsis secondary cell wall xylan are glucuronic acid, GH11 will normally produce X, XX and XUXX from alkali-extracted Arabidopsis AIR. In addition to GH11, an α-arabinofuranosidase from family GH62 was also used in a simultaneous digestion. The GH62 enzyme acts synergistically with GH11, by removing arabinosyl substitutions, GH62 converts longer oligosaccharides such as XAXX and XUXAXX to GH11-digestible fragments which can be further cleaved to X, XX and XUXX, (in these two examples XXXX and XUXXXX) (Kellett et al., 1990). By comparing the GH11 digestions alone to GH11 and GH62 simultaneous digestions it should be possible to identify arabinosylated oligosaccharides (bands) in the GH11 digestions, and determine if the transgenic plants have additional arabinosylation over their genetic background.

To test if expression of the transgenes from *EC62261* resulted in arabinosylation of xylan in wild type plants, xylan was alkali-extracted from AIR of homozygous T$_2$ plants. The xylan was digested with GH11, with and without GH62, and the produced oligosaccharides were labelled with the ANTS fluorophore and separated and visualised by PACE. The main oligosaccharide products of the GH11 digests were X, XX and XUXX, see Figure 5.8. X and XX were identified by their co-migration with the xylo-oligosacchride ladder, while XUXX migrates between X4 and X3, due to the additional negative charge of the glucuronic acid. Unexpectedly, XUX and XUXXX were also produced. Other members of the Dupree group have reported that this batch of GH11 xylanase also

produces these products, possibly due to contamination by other enzymes. Importantly, simultaneous digestion with GH11 and GH62 showed no significant differences from the GH11 alone digestion, suggesting Arabidopsis wild type secondary cell wall xylan is not arabinosylated, consistent with previous reports. The transgenic lines in the wild type background produce similar oligosaccharides to the negative control, with X, XX, XUX, XUXX and XUXXX being the main products, and even at higher exposures there seem to be no GH62-sensitve products, suggesting there is no arabinosylation introduced by the construct.

It could be the case that XAX1 activity is significantly inhibited by the high substitution level of wild type Arabidopsis xylan. To test this, xylan was alkali extracted from *esk1* and *esk1.61.1* AIR and digested with GH11, with and without GH62. The resultant oligosaccharides were ANTS labelled and visualised by PACE, Figure 5.8. Similarly to wild type, the main oligosaccharides produced by GH11 from *esk1* were X, XX, XUX, XUXX and XUXXX. Comparison of the oligosaccharides produced by GH11 and GH62 from *esk1* clearly shows a low intensity GH62-sensitive band, marked by an asterisk. This is consistent with reports of an increase in endogenous xylan arabinosylation in this mutant (Temple-Sanchez,, personal communication). There were no significant differences between the oligosaccharides produced from *esk1* and the *esk1.61.1* line. This suggests any XAX1 activity is not boosted by a reduction in xylan acetylation, and that there is no significant additional arabinosylation in *esk1.61.1*. These conclusions would be strengthened by the analysis of additional independent insertional lines in the *esk1 kak* background.

To test if there was additional arabinosylation in the reduced recalcitrance mutants, *gux1 gux2* and *cad2 cad6*, xylan was alkali-extracted and digested with GH11, with and without GH62. The resultant oligosaccharides were ANTS labelled and visualised by PACE, Figure 5.9. In *cad2 cad6* and its transgenic lines, the oligosaccharides produced by GH11 were similar to wild type, with X, XX, and XUXX being the main products. The transgenic lines had no additional bands in the GH11 digests only, and there were no obvious differences in the oligosaccharide profile with the addition of GH62.

Due to the lack of secondary cell wall glucuronidation in the *gux1 gux2* mutant the main products from the GH11 digest were X, XX, XXX and XXXX, see Figure 5.9. In the *gux1 gux2* mutant, the band that co-migrates with XXXX is GH62 sensitive, and there are some higher DP bands, marked with an asterisk, that are GH62 sensitive, suggesting there may

131
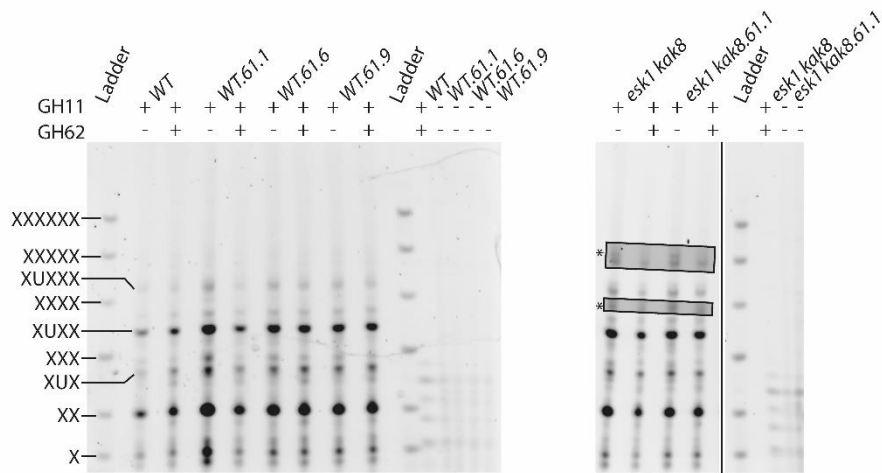
Figure 5.8: PACE analysis of GH11 and GH62 digests of wild type and *esk1 kak EC62261* lines shows no additional arabinosylation. Alkali-extracted xylan from homozygous T₂ plants expressing *EC62261* and the background genotype was digested with GH11 alone or in conjunction with GH62. The digestions were technically replicated twice and a representative gel is shown. The gels are shown at the same exposure. Areas shown at a higher exposure are highlighted with a black box. Putative arabinosylated bands are marked with an asterisk.
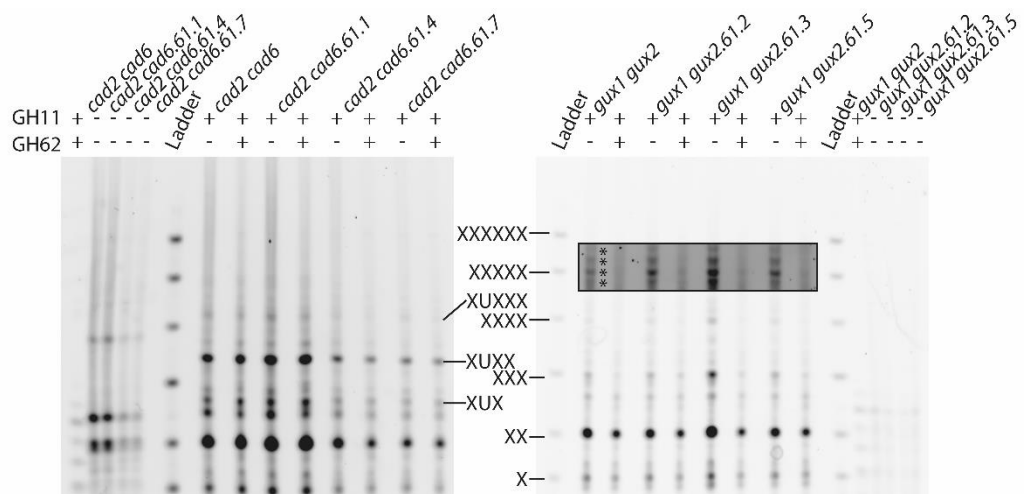
Figure 5.9: PACE analysis of GH11 and GH62 digests of *gux1 gux2* and *cad2 cad6 EC62261* lines shows no additional xylan arabinosylation. Alkali-extracted xylan from homozygous T₂ plants expressing *EC62261* and the background genotype was digested with GH11 alone or in conjunction with GH62. The digests were technically replicated twice and a representative gel is shown. The gels are shown at the same exposure. Areas shown at a higher exposure are highlighted with a black box. Arabinosylated bands are marked with an asterisk.

be some xylan arabinosylation in this genetic background. This could occur due to the reduction in xylan substitution in this mutant, as in *esk1 kak*. There is a weak XUXX band in the digestion, which is likely to arise from primary cell wall xylan decorated by GUX3 enzyme (Mortimer et al., 2015). The oligosaccharides produced from the *gux1 gux2* background are similar to those produced from the transgenic lines, suggesting no additional arabinosylation was produced.

To confirm there were no additional ferulated arabinosyl substitutions on the xylan of some transgenic lines, AIR from *gux1 gux2* and the relevant transgenic lines was subjected to mild TFA hydrolysis. The mild acid preferentially hydrolyses the unstable α glycosidic bonds that attach xylan arabinosyl substitutions to the backbone, though the xylan backbone is also cleaved. Previously, this method has been used to release ferulate arabinosyl susbtitutions for detection by PACE and demonstrate their reduction in the *xax1* rice mutant (Feijão, 2016). As positive and negative controls for the detection of ferulated arabinosyl substitutions on the PACE gel, both wild type and *xax1* rice AIR were hydrolysed alongside the Arabidopsis *gux1 gux2* transgenic lines. Due to the requirement for a large mass of AIR, most genotypes were excluded. After mild TFA hydrolysis, the supernatant was ANTS labelled and the oligosaccharides were separated and visualised by PACE, Figure 5.10. The mild TFA hydrolysis worked in all samples, as demonstrated by the presence of xylo-oligosaccharide bands due to hydrolysis of the backbone glycosidic bonds. The wild type rice has four bands that previously have been characterised as ferulated and coumarylated arabinose (A5f3 and A5c3) as well as the same structures with a β-1,2 xylosyl substitution on the arabinose (D2,5f3 and D2,5c3). These structures were previously characterised using enzyme digestions and mass spectrometry and are here labelled based only on co-migration (Feijão, 2016). In the *xax1* mutant the intensity of these bands are signifcantly reduced, as previously reported. The intensity of the bands in the wild type rice and *xax1* mutant lanes was measured by densitometry, using ImageJ, and peaks were scaled to the same peak that migrates around XXXX. Relatively speaking, it is clear that the esterified arabinosyl structures are all less intense in *xax1* than wild type rice, as previously reported. The mild TFA hydrolysates of the background genotypes and *EC62261* transgenic lines are similar, and lack additional intensity where A5f3, A5c3, D2,5f3 and D2,5c3 migrate. These results suggest that the xylan in the transgenic lines is not detectably ferulated.
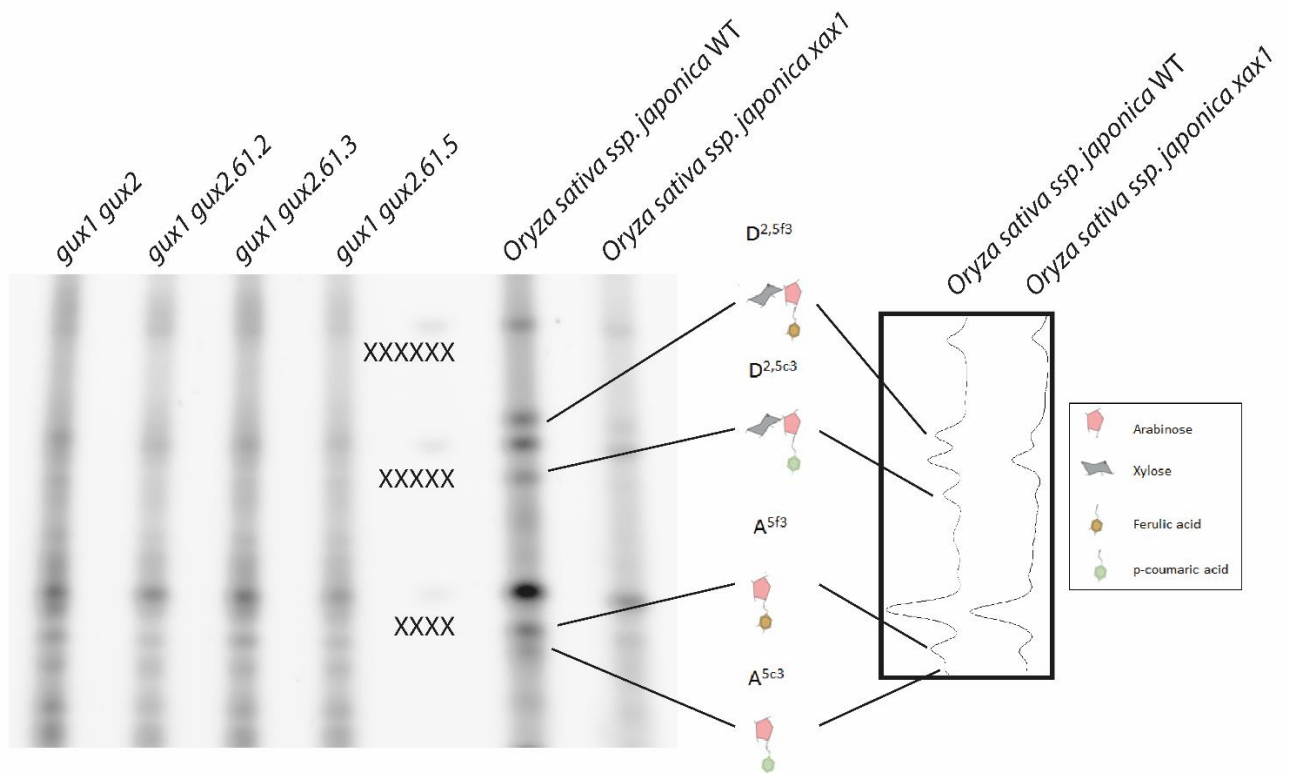
Figure 5.10: Ferulated arabinose residues were not detected in transgenic lines. AIR of transgenic *gux1 gux2* plants, *gux1 gux2* negative control, wild type rice and *xax1* mutant rice were subjected to mild acidolysis. The esterified arabinose substitutions were previously structurally characterised, and are labelled based on co-migration after (Feijão 2016).

5.3 Discussion

In previous chapters it was proposed that glucuronic acid mediated xylan:lignin cross-linking is crucial for recalcitrance of the cell wall to enzymatic digestion. In monocots xylan:lignin cross-linking occurs through the participation of xylan ferulate substitutions in lignification. In this chapter the objective was to reconstitute the pathway for synthesis of ferulate substitutions in Arabidopsis. This reconstitution enables testing of ferulation synthesis models. If successful, this could enable testing of the functionality of ferulate mediated xylan:lignin cross-linking in Arabidopsis. To this end, five genes that were hypothesised to be sufficient for the synthesis of ferulate substitutions were introduced into multiple genetic backgrounds. Expression of the proteins was confirmed by Western blot and confocal microscopy experiments. Xylan structural changes were investigated using enzymatic hydrolysis and mild acid hydrolysis, but no changes were observed. These results provide the lab with a platform for engineering secondary cell wall synthesis with an expanded range of secondary cell wall-specific promoters. In addition, the lack of ferulation in the transgenic lines has interesting implications for our current model of xylan ferulation.

5.3.1 Despite the use of a previously successful strategy, no xylan structural changes were observed in transgenic plants

Previously, expressing GT61 enzymes in Arabidopsis secondary cell walls has been a powerful tool for understanding their functions in xylan synthesis. The first demonstration of this was the identification of the xylan arabinosylation activity of clade A GT61s. Bioinformatic analysis had identified clade A GT61s as candidates for xylan arabinosylation (Mitchell et al., 2007). RNAi silencing of two GT61s that are highly expressed in wheat endosperm resulted in reductions in xylan arabinosylation (Anders et al., 2012). To decisively demonstrate their activity, the GT61 genes were ectopically expressed in Arabidopsis, which lacks xylan arabinosylation in secondary cell walls. The α-1,3 arabinosylation activity of TaXAT1/2 was detected by digestions with the GH11 and GH62 enzymes described in this chapter, followed by PACE analysis of the oligosaccharide products. In addition to proving clade A GT61 activity, expression of conifer GT61s in Arabidopsis has helped us understand previously unknown aspects of xylan synthesis (Temple-Sanchez, personal communication). Expression of conifer GT61s has produced novel substitutions in Arabidopsis, that have not been identified in conifers.

135

This confirms the feasibility of this approach for understanding xylan substitution synthesis.

In this chapter a similar approach was taken with the goal of engineering ferulate substitutions onto xylan. This approach enables us to understand if our current model of xylan ferulation synthesis is correct and may help us to investigate the function of xylan:lignin cross-linking in cell wall recalcitrance. Xylan ferulation is thought to require GT61 enzymes to perform the final step of transferring the ferulated arabinofuranose to the xylan (Feijão, 2016). However, ferulation is thought to require more components than a single GT61, see Figure 5.1 and the introduction Section. UDP-Ara$p$, from the Golgi, is converted to UDP-Ara$f$ on the cytosolic face of the Golgi by a UAM (Delgado et al., 1998). There is strong *in vivo* RNAi evidence that the UDP-Ara$f$ is then ferulated by BAHD acyltrasnferases (Mitchell et al., 2007; Bartley et al., 2013; de Souza et al., 2018). The ferulated UDP-arabinofuranose, which would presumably be cytosolic, would need to be transported back into the Golgi, by a UAFT (Rautengarten et al., 2017).

To test this model of ferulated xylan synthesis, OsUUAT, OsUAFT and OsUAM genes, along with the critical OsXAX1 and BdBAHD01 enzymes were cloned into a genetic construct for secondary cell wall expression, transformed into Arabidopsis. The plants were studied for evidence of expression, and all the genes were successfully expressed. Evidence for the expression of UUAT and UAFT is the weakest and would be strengthened by additional Western blot analysis. The xylan structure of homozygous T_2 plants was investigated using PACE analysis of GH11 and GH62 digestions of the xylan, along with mild acidolysis digestions. Despite the effectiveness of this strategy for investigating xylan structural changes introduced by GT61s (Anders et al., 2012; Temple-Sanchez, personal communication), there were no observed changes in xylan structure in multiple lines in four different Arabidopsis genotypes.

5.3.2 Synthesis of ferulated xylan is probably more complex than current models suggest

It could be the case that some of the proteins are inactive in Arabidopsis due to misfolding or incorrect Golgi trafficking, for example, and this could explain the lack of ferulation in the transgenic lines obtained in this work. However, xylan synthesis enzymes from phylogenetically distant species from multiple GT families have been successfully

expressed in Arabidopsis, so this explanation seems unlikely (Anders et al., 2012; Chiniquy et al., 2013; Lyczakowski et al., 2017; Temple-Sanchez, personal communication). It is more likely that the ferulation pathway is missing significant components in Arabidopsis, suggesting our current model is incomplete.

There are three main ideas that could explain the lack of ferulation in the transgenic plants: the absence of necessary protein-protein interactions, or absence of correct biosynthetic components, and a lack of sufficient metabolite concentrations.

Specific protein-protein interactions may be required for the function of the ferulation pathway. In wheat, it has been reported that the xylan synthase complex contains putative UAM genes from GT75 (Zeng et al., 2010; Jiang et al., 2016), in addition to the IRX10/14 proteins that are found in the asparagus XSC (Chiniquy et al., 2013; Zeng et al., 2016). Mass-spectrometric (MS) analysis of trypsin-digested semi-purified XSC contained peptides unique to GT43 and GT47 proteins, and several GT75s (Zeng et al., 2010). In addition, the immunoprecipitated complexes could produce glucuronosylated and arabinosylated xylo-oligosaccharides, suggesting that these complexes also contained GUX and XAT (GT61) enzymes, though these were not identified in the MS analysis. This could suggest that protein-protein interactions between backbone synthesis enzymes, substitution enzymes such as GUX and XAT as well as accessory proteins such as UAMs, could be important for xylan synthesis.

These protein-protein interactions may be important for the normal activity of the xylan synthase complex, and may be specific to different species due to protein sequence divergence. For instance, complementation of Arabidopsis *irx9* and *irx14* mutants with rice orthologues led to mixed degrees of complementation (Chiniquy et al., 2013). While growth was complemented, different rice orthologues affected the length and amount of xylan differently, suggesting their protein interactions, or expression, are insufficient to completely complement the biosynthesis of xylan in Arabidopsis. It could be the case that in the transgenic lines described in this chapter any protein-protein interactions between AtIRX proteins and OsUAM and OsXAX1 are weaker than native interactions with OsIRX proteins, due to sequence divergence, with effects on the synthesis of ferulate substitution of xylan. Since there is some evidence that the ferulated UDP-Ara*f* precursor is assembled on UAM proteins (Obel et al., 2003), this could extend to hypothetical interactions between BdBAHD01 and OsUAM proteins, and it may have been better to use OsBAHD01 as the BAHD candidate in the construct. To test these ideas, OsIRX14

137

and OsBAHD01 could be included in a similar genetic construct, and any xylan structural changes could be evaluated using PACE.

Another possible explanation for the lack of ferulation is the absence of critical components from the construct. To test this idea the rice co-expression database, RICEFREND, was examined for other candidate genes involved in ferulation of xylan, using XAX1 as bait. The most highly co-expressed gene, after IRX9, was BS1, a GDSL esterase that can remove acetyl substitutions from xylose (Zhang et al., 2019) and has been proposed to act on the xylan backbone in rice (Zhang et al., 2017). Interestingly, a survey of the six partially characterised clade A GT61s from rice demonstrates a clear pattern of reported activity and co-expression with BS1-like GDSL esterases (Sato et al., 2013). The enzymes that are thought to add unesterified substitutions to the backbone, XAT2/3 and XYXT, do not co-express well with BS1-like GDSL esterases. On the other hand, XAX1 (Chiniquy et al., 2012), SAC1 (Marriott et al., 2014) and mutant 5 (Feijão, 2016), which all have some evidence suggesting they could act as ferulated UDP-arabinosyl transferases, co-express with BS1 or a homologue from the same clade (Zhang et al., 2017), with a mutual rank of 12 or less. This could suggest that the ferulated UDP-arabinosyl transferases require BS1-like activity.

How could BS1 be involved in ferulation? BS1 has been reported to remove acetate substitutions from the xylan backbone (Zhang et al., 2017). The activity of some xylan substitution enzymes (GUXs from GT8) has been reported to be strongly affected by the amount of xylan acetylation and its pattern (Grantham et al., 2017). It could be the case that BS1 generates small de-acetylated regions of the xylan backbone that XAX1 and related GT61s can act on, and without this, XAX1 cannot substitute xylan. In support of this, the *bs1* mutant is dwarfed and brittle (Zhang et al., 2017), like *xax1*, suggesting mechanical strength defects, consistent with a reduction in ferulation-mediated xylan:lignin cross-linking (Chiniquy et al., 2012). To test these ideas, mild acid hydrolysis to detect ferulated structures in *bs1* and matching wild type rice AIR could be performed, as well as PACE analysis of the degree of xylan arabinosylation, which would both be hypothesised to be reduced in the *bs1* mutant. In terms of reconstituting the xylan ferulation pathway in Arabidopsis the BS1 gene could be included in a construct with the genes already described here, and the transgenic lines produced could be tested for increases in xylan arabinosylation and ferulation, using the methods described in this chapter.

The final consideration could be an insufficient supply of feruloyl CoA or UDP-Ara*f*. Other GT61s, including from rice, produce reasonably high levels of arabinosylation in Arabidopsis secondary cell walls (Temple-Sanchez, personal communication). This suggests UDP-Araf supply is not a significant issue during xylan synthesis in Arabidopsis. Feruloyl CoA is synthesised from caffeoyl CoA by the caffeoyl coenzyme A o-methyl transferase (CCoAOMT) enzyme (Wagner et al., 2011) or from ferulic acid by the 4-hydroxycinnamate:CoA ligase (4CL) (Barros et al., 2019). CoA derivatives of hydroxycinnamates are thought to be the substrates for BAHD acyltransferases, and thus the supply of this metabolite may significantly affect the amount of ferulation in the transgenic Arabidopsis plants. Despite this, there does not seem to be an expansion of CCoAOMT or 4CL genes in monocots. However, overexpression of a Sorghum CCoAOMT gene specifically increased the amount of wall-bound hydroxycinnamates, but did not affect the amount of lignin (Tetreault et al., 2018). This could suggest that feruloyl CoA supply could regulate xylan ferulation, or that specific CCoAOMT genes supply feruloyl CoA for ferulation, while others direct feruloyl CoA supply to lignification, possibly through substrate channelling. To test this idea monocot CCoAOMT or 4CL genes could be included in a construct designed to introduce ferulate substitutions into Arabidopsis.

Chapter 6: Solid-state NMR investigation of the interactions between cell wall components in *Picea abies*

6.1 Introduction

Conifer species are of great economic and ecological importance: they are the dominant clade in the boreal forests which cover 30% of global forest area (Bradshaw et al., 2009); and the majority of commercially planted trees are conifers (Del Lungo et al., 2006). The cell walls of conifers play an important role in softwood material properties, the ease of processing to non-wood products and the resistance of the plant to pathogens. The softwood they produce, which is composed of cell walls, can be used for a variety of purposes including paper, construction and food technology (Devaraj et al., 2019). It is poorly understood how cell wall components and their interactions determine softwood properties. In order to design genetic or chemical strategies for modification of cell walls to improve conifers for any one purpose it is important to describe the native interactions of cell wall components in conifer cell walls.

The main component, by weight, of conifer cell walls is cellulose. Cellulose is formed from linear β-1,4 glucan chains which crystallise together to form a microfibril (Jarvis, 2018). The microfibril is thought to contain 18-24 chains in spruce (Fernandes et al., 2011), but this can vary between species (Thomas et al., 2014), and multiple microfibrils might crystallise together (Jarvis, 2018). Solid state NMR has identified two structural domains of cellulose (Atalla and Vanderhart, 1984), distinguishable by the chemical shifts of the glucosyl residues at carbons 4, 5 and 6. Historically these have been referred to as surface and interior cellulose chains, but here they are referred to as domain 1 and 2 after (Dupree et al., 2015), since their exact location in the microfibril remains to be precisely determined. There is some evidence supporting the assignment of these domains as surface/interior chains. For instance the ratio of domain 1:2 does scale with the width of microfibrils between non-plant species with differently sized microfibrils (Earl and VanderHart, 1981), and the domain 1 chains are more inaccessible to deuteration and excitation by water protons (Wang et al., 2016b).

However, Density Functional Theory (DFT) modelling and direct measurement of the proximity between the glucosyl residue carbon 6 protons and other carbons have suggested that the key structural difference between domains 1 and 2 is the

conformation of the carbon 6 hydroxymethyl, thought to be *tg* and *gg/gt* respectively (Wang et al., 2016b; Yang et al., 2018; Phyo et al., 2018). In the *tg* conformation of domain 1, see Figure 1, the hydroxyl of carbon 6 forms a hydrogen bond with the carbon 2 hydroxyl of the preceding residue, and the carbon 3 hydroxyl of a glucosyl residue in an adjacent chain (Jarvis, 2018). In the *gg/gt* conformation of domain 2 the cross-chain hydrogen bonding does not occur. It is logical to conclude that the proposed interior location of domain 1 chains stabilises a *tg* conformation due to the hydrogen bonding of adjacent glucan chains, while the surface domain 2 chains remain in a *gg/gt* conformation, because they lack these hydrogen bonds. However, the proximity (determined by NMR) of domain 1 chains to pectic and hemicellulosic polysaccharides seems to be inconsistent with their proposed location in the interior of the microfibril (Wang et al., 2015; White et al., 2014; Simmons et al., 2016). Thus, the location of the two domains of cellulose in the microfibril and their interactions with hemicelluloses and lignin are still substantial questions in the cell wall field. Resolution of these issues would enable more detailed interpretation of solid-state NMR data.
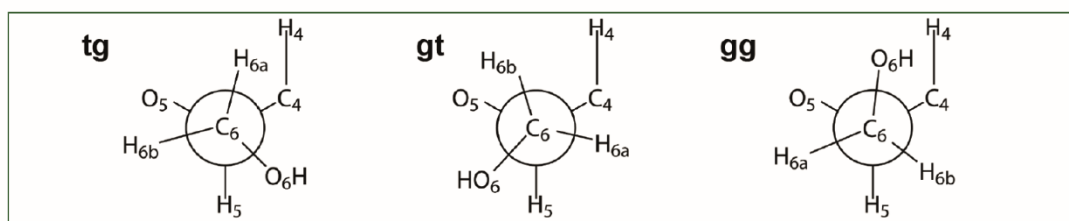
Figure 6.1: The different conformations of the carbon 6 hydroxymethyl of the glucosyl units of cellulose. In the *tg* conformation, the carbon 6 hydroxyl is oriented towards the carbon 4 of the same residue. This conformation is stabilised by hydrogen bonding from carbon 2 and 3 of adjacent glucan chains. The other two conformations occur without hydrogen bonding from adjacent glucan chains. Adapted from (Wang et al., 2016b).

Conifer xylan is composed of β-1,4 linked xylosyl residues, with α-1,2 glucuronosyl and α-1,3 arabinosyl substitutions and no acetylation (Busse-Wicher et al., 2016). The majority of the substitutions are patterned such that there is an even number of backbone xylosyl residues between consecutive substitutions. There are some lower frequency substitutions which are not patterned in this way (Martínez-Abad et al., 2017), notably consecutive glucuronic acid substitutions have been observed.

The presence of evenly patterned substitution of xylans throughout eudicots and gymnosperms led to the hypothesis that these patterns facilitate a specific interaction between xylan and cellulose, where xylan binds to the hydrophilic face of cellulose as a

two-fold screw (Busse-Wicher et al., 2014; Bromley et al., 2013). This interaction would be facilitated by the structural similarity of the xylan and glucan chains, as both polysaccharides have equatorial hydroxyl groups. These evenly spaced patterns are thus termed "compatible", to distinguish from the highly substituted areas of xylan which are "incompatible" with binding to the hydrophilic face due to the steric hindrance of the substitutions. However the negative effects of incompatible substitutions on xylan binding to the hydrophilic face have been disputed by some molecular dynamics simulations (Martínez-Abad et al., 2017). DFT simulations suggest that the two-fold screw conformation of xylan significantly alters the chemical shifts of carbons 1 and 4 relative to the three-fold xylan that is found in solution (Simmons et al., 2016). Solid-state NMR showed that this new conformation of xylan exists in Arabidopsis secondary cell walls; that it is physically close to cellulose, and dependent both on the presence of cellulose and compatible xylan patterns (Grantham et al., 2017). Two-fold xylan has also been observed in the secondary cell walls of three monocot species (Kang et al., 2019), though any relationship to substitution patterning has yet to be determined. Multidimensional solid-state NMR has yet to be used to investigate the conformation of xylan in gymnosperm cell walls. Previously, studies utilising electron microscopy and FT-IR have suggested that xylan does not bind to cellulose in Gymnosperms (Terashima et al., 2009; Simonović et al., 2011). These studies proposed that xylan associates with lignin away from the cellulose surface, though the evidence is not robust since these techniques cannot discriminate well between different polysaccharides. The mode of xylan-binding to cellulose in gymnosperms remains an unresolved question, which could be elucidated with solid-state NMR.

Galactoglucomannan (GGM) is the predominant hemicellulose in conifers. GGM has a β-1,4 linked backbone of mannosyl and glucosyl residues, in a 4:1 ratio (Hannuksela and Hervé du Penhoat, 2004). The GGM backbone is synthesised by the Cellulose Synthase Like A (CSLA) proteins of the GT2 family (Goubet et al., 2009). Some mannosyl residues can be α-1,6 galactosylated, by Mannan α-Galactosyl Transferases (MAGTs), or acetylated, by Trichome Birefringence-Like (TBL) proteins at carbons 2 or 3 (Yu et al., 2018; Zhong et al., 2018; Hannuksela and Hervé du Penhoat, 2004). Though patterned mannans have been observed in nature, such as the glucose-mannose repeating glucomannan synthesised by CSLA2 in Arabidopsis (Yu et al., 2018), spruce GGMs have a backbone that is around 80% mannosyl residues, which are randomly arranged in the backbone, along with randomly patterned substitutions (Hannuksela and Hervé du

Penhoat, 2004). The axial position of the carbon 2 hydroxyl of mannosyl residues may prevent the binding of GGM to the cellulose surface. *In silico* experiments suggest that the mannosyl residues reduce binding of GGM to cellulose if the cellulose cross-section is hexagonal, and if the mannose carbon 2 hydroxyl is directed at the cellulose surface (Yu et al., 2018). In addition, manno-oligosaccharides form less stable two-fold screws than gluco-oligosaccharides (Berglund et al., 2016). However, *in silico* cellulose microfibrils with "grooves" in the surface, such as the square shape, can trap GGM chains (Yu et al., 2018), see Figure 6.2. *In vitro*, spruce GGM can bind to the surface of cellulose microfibrils, and this binding is strengthened by the removal of acetyl groups (Whitney et al., 1998; Hannuksela et al., 2002). Samples from such *in vitro* binding studies have not been examined using multi-dimensional solid-state NMR, and thus it is unclear if the mannans are as closely associated with cellulose as xylan is the Arabidopsis cell walls, or if such "binding" corresponds to the kind of intimate interactions which have been investigated *in silico*. *In vivo* electron microscopy and FT-IR has been used to suggest that GGM binds to the surface of cellulose in spruce (Terashima et al., 2009; 2004; Simonović et al., 2011). Multidimensional solid-state NMR has been very useful for investigating the xylan:cellulose interaction and should enable investigation of any *in vivo* GGM:cellulose interaction.
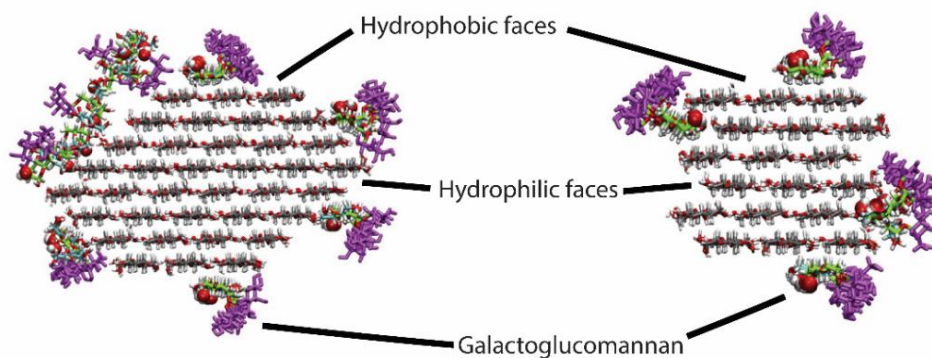
Figure 6.2: The stability of galactoglucomannan binding to cellulose is affected by microfibril shape and hydrophilicity *in silico*. The image of a 250 ns molecular dynamics simulation is shown, adapted from (Yu et al., 2018). GGM binds stably to the hydrophobic faces of both cellulose microfibrils. On the hydrophilic face, binding is stabilised when the GGM becomes trapped in "grooves" on the cellulose surface.

The lignin polymer forms up to 30% of conifer cell walls by weight (Yue et al., 2016). Secreted monolignols are oxidised by laccase and peroxidase enzymes (Ralph et al., 2004). In conifers only coniferyl and small amounts of *p*-coumaroyl alcohol are synthesised, making the lignin composition significantly different from angiosperms,

143

where large amounts of sinapyl alcohol are produced. The produced radicals combinatorially couple to form a phenolic polymer. The unique composition of gymnosperm lignin results in more carbon-carbon linkages than in angiosperm lignin (Wagner et al., 2015).

The most common linkage formed between monolignols is the β-O-4 linkage (Yue et al., 2016). During the production of a β-O-4 linkage a quinone methide intermediate is formed. To form the mature linkage, nucleophilic attack must occur at the α position, and this nucleophile may be water or a polysaccharide nucleophile (Terrett and Dupree, 2019). In the latter case a covalent bond between the emerging lignin polymer and the polysaccharide is formed. In Japanese red pine the carbon 6 of a mannosyl residue in GGM has been shown to from a covalent linkage to lignin in this way, anchoring GGM and lignin together (Nishimura et al., 2018). Theoretically this could also occur with xylan (or any cell wall polysaccharide), and may be more likely to occur with acidic sugars such as glucuronic acid, as the carboxylate ion is a stronger nucleophile than a hydroxyl. The frequency of such lignin:polysaccharide linkages and their importance for cell wall assembly remain unknown. Their role in cell wall recalcitrance was investigated in chapter 4. *In silico* studies have also suggested that monolignols and lignins can bind to cellulose non-covalently and covalent linkages may be important for these non-covalent interactions (Houtman and Atalla, 1995; Besombes and Mazeau, 2005a; 2005b).

Recently, the interactions of lignin with polysaccharides have been investigated in Arabidopsis and grass species using solid-state NMR (Kang et al., 2019). It was found that there are cross-peaks between lignin and three-fold xylan and a distorted two-fold xylan, as well as domain 2 cellulose, which has been proposed to be on the surface of the cellulose microfibril. Therefore in principle, multidimensional solid-state NMR can enable us to determine the extent of associations between polysaccharides and lignin in spruce cell walls, and investigate if the unique composition of conifer lignin affects these interactions.

Previously, the solid state NMR ^{13}C CP-INADEQUATE and CP-PDSD experiments have been used to identify the chemical shifts of polysaccharides in the cell wall and their spatial proximity (Dick-Pérez et al., 2011; Wang et al., 2014; Dupree et al., 2015; Simmons et al., 2016). The CP-INADEQUATE experiment utilises scalar coupling (or through-bond transfer of magnetisation) to identify pairs of carbons that are covalently bonded together. Magnetisation is transferred to the carbons by first magnetising the

protons they are bonded to. While enhancing sensitivity, this cross-polarisation (CP) transfer only occurs efficiently when the protons and carbons have low mobility. Therefore, only the more immobile components of the cell wall will be visible in this experiment. In the spectra each carbon has a single quantum (SQ) chemical shift, and the sum of these chemical shifts, the double quantum (DQ), appears in the second dimension. In this way, covalently bonded pairs of carbons have different SQ shifts (on the x axis), but appear at the same DQ shift (on the y axis), providing verification that they are covalently connected, and enabling resolution of carbons by their DQ shifts, even if they have similar SQ shifts.

The CP-PDSD experiment utilises dipolar coupling (or through-space transfer of magnetisation) to identify pairs of carbons that are close to each other in space (Dick-Pérez et al., 2011; Wang et al., 2016b; Simmons et al., 2016). At short mixing times the transfer will only occur between carbons in the same molecule, but at longer mixing times transfer will occur between carbons that are a maximum of 5-10 Å apart. 10 Å is approximately the length of two β-1,4 linked glycosyl residues. This resolution is high, considering the heterogeneous nature of the material.

These multidimensional ^{13}C solid-state NMR experiments were performed on ^{13}C enriched material. The Arabidopsis plants were grown with Dr Jan J. Lyczakowski and the spruce plants were purchased from IsoLife. Solid-state NMR experiments were performed by Professor Ray Dupree, in collaboration with myself, Dr Jan L. Lyczakowski, Dinu Iuga and Dr William T. Franks, at the University of Warwick on an 850MHz or 700MHz spectrometer.

In this chapter, the solid-state NMR experiments described above will be used to investigate the *in muro* interactions of GGM, xylan, cellulose and lignin in spruce. The specific aims are:

- To identify GGM chemical shifts in spruce wood and Arabidopsis
- To define the chemical shift diversity of cellulose environments within spruce wood
- To determine the *in muro* proximity between xylan, GGM and cellulose sub-domains
- To determine the *in muro* proximity of lignin to cell wall polysaccharides

6.2 Results

6.2.1 Identification of galactoglucomannan chemical shifts in Arabidopsis and *Picea abies* and analysis of xylan conformation in spruce

To use genetics to identify the chemical shifts of GGM carbons, CP-INADEQUATE experiments were performed on two biological replicates of wild type Arabidopsis and the *csla2/3/9* mutant, which, due to the absence of active CSLA proteins, lacks detectable GGM in the stem (Goubet et al., 2009). The neutral carbohydrate regions of representative spectra are overlaid in Figure 6.1. The correlations of polysaccharide carbons 1 and 2 can be observed at DQ values of between 170-180 ppm, due to the high chemical shift of carbon 1. There are two pairs of carbon 1-2 correlations in the wild type that are absent in the *csla2/3/9* mutant, suggesting these correspond to the chemical shifts of GGM carbons 1 and 2. They are labelled mannose carbon 1 and 2 (M1= 101.9 and M2=71.7ppm) and mannose X carbon 1 and 2 (M1^X=101.0ppm and M2^X=72ppm). The differences between M and M^X will be identified later in this Section.

The chemical shifts of these carbons are consistent with the reported values for GGM in solution NMR studies, see table 6.1 for a comparison of the solution NMR values for spruce GGM and the assignments made here. Correlations between M4$^{[X]}$-M3/5$^{[X]}$, M5$^{[X]}$, M6$^{[X]}$ and M2$^{[X]}$-M3$^{[X]}$ are not easily identifiable in this spectra, however these areas are congested and dominated by strong correlations from cellulose and xylan.

To compare the chemical shifts of GGM in Arabidopsis with spruce, CP-INADEQUATE experiments were performed on spruce wood. The spruce spectrum is overlaid with the wild type Arabidopsis CP-INADEQUATE in Figure 6.4. The M1$^{[X]}$-M2$^{[X]}$ correlations identified in Arabidopsis are much more intense, matching the increased amount of GGM in spruce relative to Arabidopsis (Scheller and Ulvskov, 2010). This suggests these signals are attributable to GGM in both species.

Interestingly, the three-fold and two-fold xylan carbon 4 and 5 signals (Xn4^{3f}=77.7 ppm, Xn5^{3f}=63.9 ppm, Xn4^{2f}=82.4 ppm, Xn5^{2f}=64.3 ppm) that previously were identified in Arabidopsis are clearly visible in the spruce spectra, Figure 6.4. As was previously hypothesised, this suggests two-fold xylan formation is a common feature of plant cell walls, that may be enabled by the evenly patterned xylan substitutions that have been identified in eudicots and gymnosperms (Busse-Wicher et al., 2016). It will be important
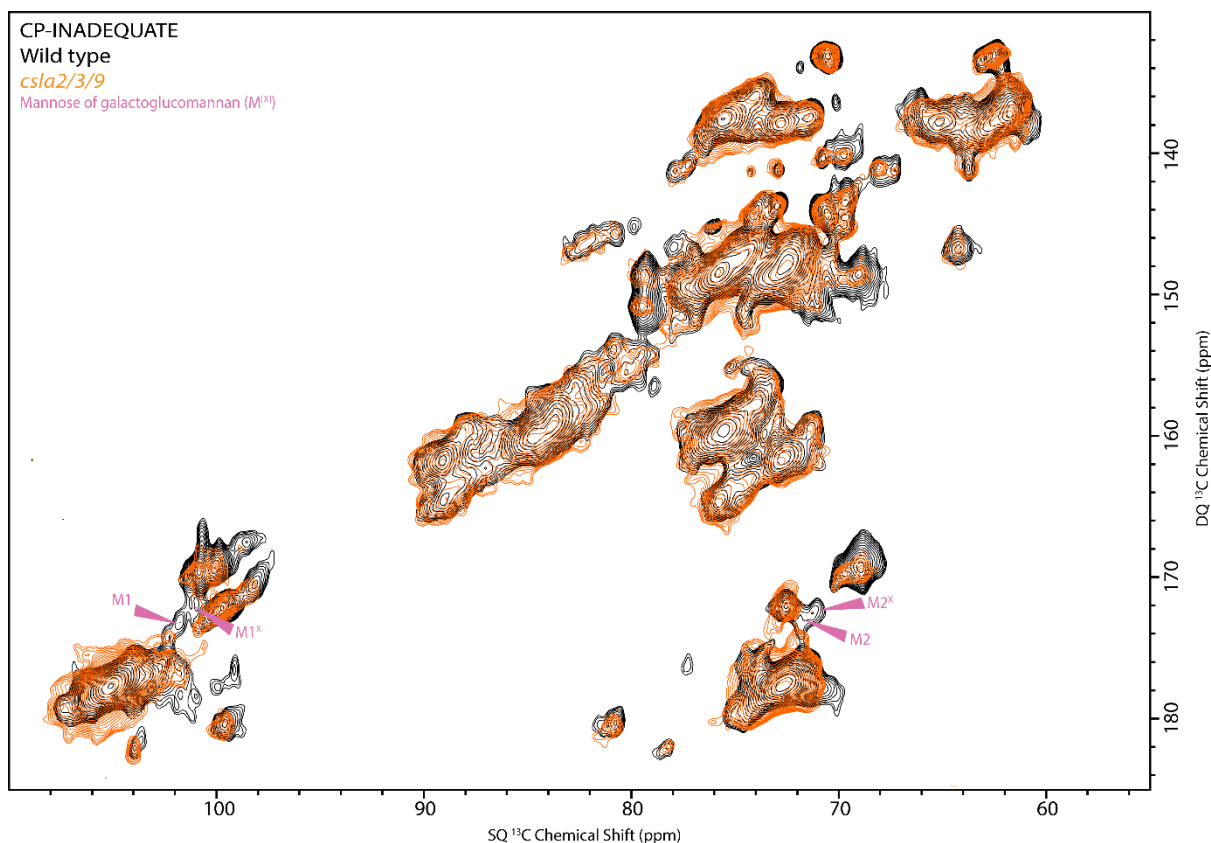
Figure 6.3: Genetic identification of GGM chemical shifts in Arabidopsis. The overlaid carbohydrate regions of refocussed ^{13}C CP-INADEQUATE spectra of the wild type and *csla2/3/9* mutant Arabidopsis. There are representative spectra, performed on two biological replicates of undried basal stem material from 16 plants. The proposed peaks corresponding to M1[X] and M2[X] are labelled. The superscript X enables distinction between the two pairs of carbons.

Table 6.1 Carbon 1 and 2 chemical shifts for GGM in the solid- and solution- state are similar. Carbon 1 and 2 chemical shifts for the two mannose residues (M and M^X) are shown. In the upper table, the chemical shift assignments made here are shown. Solution-state NMR chemical shifts taken from (Hannuksela and Hervé du Penhoat, 2004) are shown, in grey, in the lower table.

GGM	1	2
M	101.9	71.7
M^X	101.0	72.0
M	101.2	71.1-71.5
M^{Ac}	99.8-100.6	69.2-72.7
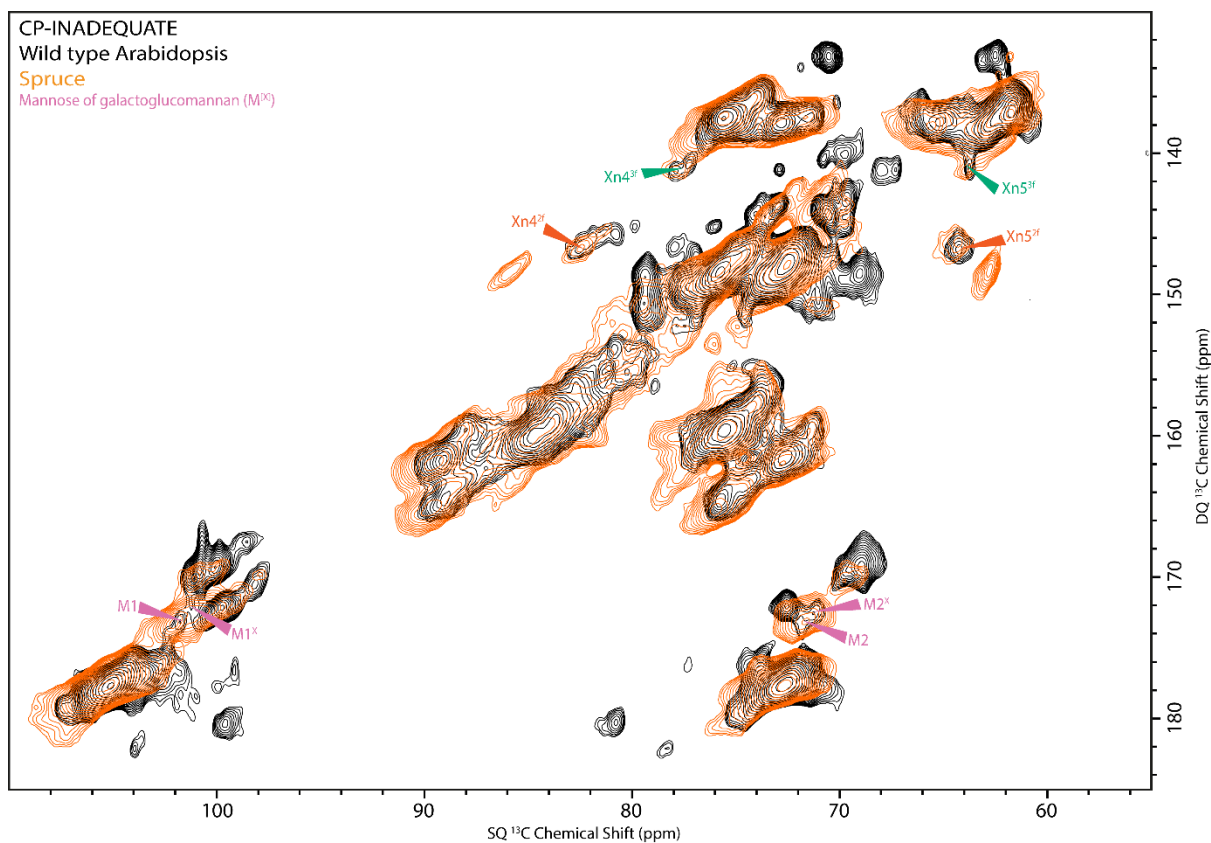
Figure 6.4: The GGM carbon 1 and 2 peaks are significantly stronger in spruce than Arabidopsis. The carbohydrate regions of refocussed ^{13}C CP-INADEQUATE of spruce and wild type Arabidopsis are overlaid. This experiment has been performed on one biological replicate of four spruce plants. The stems were never-dried and were ground in liquid nitrogen, while the Arabidopsis stems were undried, frozen in liquid nitrogen, but not ground. The proposed peaks corresponding to M1[X] and M2[X] are labelled. The two conformers of xylan, previously identified in (Simmons et al., 2016), are labelled also. The superscript X enables distinction between the two pairs of carbons.

to verify if the two-fold xylan is bound to cellulose in spruce.

To fully assign the chemical shifts of the two GGM residues, especially the missing M4/3/5/6 assignments, 30ms CP-PDSD experiments were performed on spruce. The spectrum is shown in Figure 6.5. In the 30ms CP-PDSD spectrum, M1[X]>M2[X] correlations were identified to corroborate the assignments made from the CP-INADEQUATE spectrum. Mannose carbon 4 (M4[X]=80.4ppm) and 6 (M6[X]=61.7ppm) were identified by correlations with M1 and M1[X]. The M4[X] and M6[X] chemical shifts were corroborated by M4[X]>M6[X] correlations. The chemical shifts of M3[X]/M5[X] were identified by correlations with M1 and M1[X]. Correlations with the acetate methyl carbon (Ac[Me]=21.4ppm) were used to identify the chemical shifts of carbons 2 and 3 in one of the GGM residues.

As GGM mannosyl residues can be acetylated or galactosylated, one of the two residues is likely to be modified in this way. Galactosylation is relatively rare, therefore it is likely that M or M[X] are acetylated (Hannuksela and Hervé du Penhoat, 2004). To determine which residue is acetylated, correlations between the acetate methyl (Ac[Me]=21.4 ppm) and GGM carbons were identified in the 100ms CP-PDSD spectrum, see Figure 6.6. The Ac[Me]> M1[X] correlation clearly indicates that M[X] is an acetylated mannose residue (M[Ac]), while M is likely to be an unacetylated mannose residue.

The chemical shifts of the two identified GGM residues, M and M[X/Ac], are summarised in table 6.2. They have similar chemical shifts at carbons 2, 4 and 6, but differ significantly at carbon 1. This could be due to the effect of acetylation, either directly by altering the local chemical environment, or indirectly by altering the glycosidic bond angle of acetylated residues. Acetylation has been shown to affect solution-state NMR assignments of GGM as well. Interestingly, the chemical shift of carbon 4 is 2-6 ppm higher than for GGM in solution, while those of carbon 6 are similar to solution NMR assignments (Hannuksela and Hervé du Penhoat, 2004; Nishimura et al., 2018), the carbon 1 chemical shift of M is around 1 ppm higher than solution, but M[Ac] only differs by about 0.5 ppm. The chemical shifts of carbons in the glycosidic bond can be affected by the angle of the glycosidic bond. For instance, the conformational transition of xylan from a three to a two-fold screw causes a ~4ppm increase in the carbon 4 shift and a 2-3 ppm increase in the carbon 1 chemical shift (Simmons et al., 2016). The chemical shifts of carbons 1 and 4 can also be affected by the carbon 6 hydroxymethyl conformation in β-1,4 glycans (Wang et al., 2016b; Yang et al., 2018; Phyo et al., 2018).
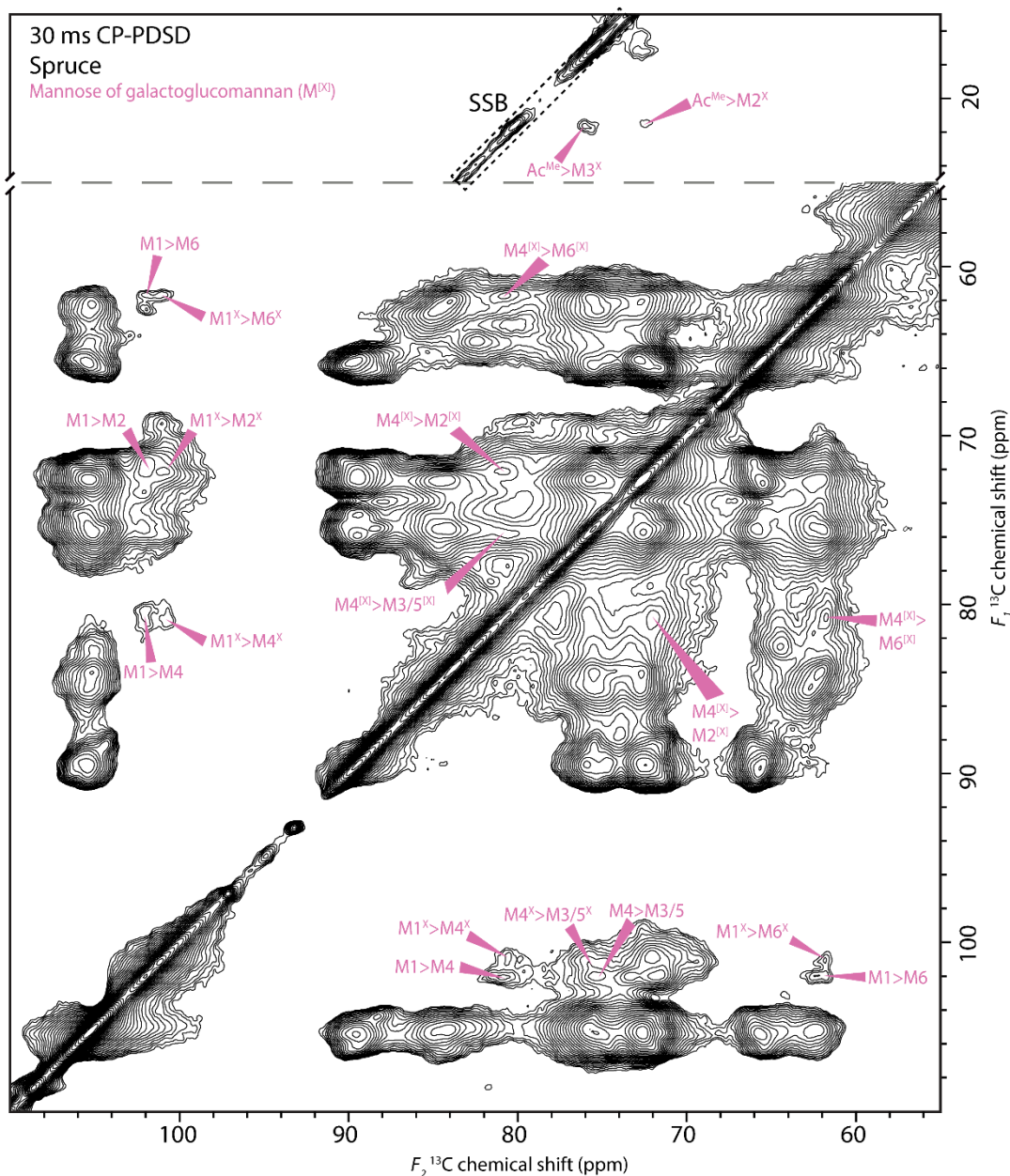
Figure 6.5: Further assignment of the chemical shifts of GGM in spruce. The acetate methyl (Ac^{Me}) and carbohydrate regions of a 30 ms mixing time CP-PDSD spectrum of spruce are shown. The intramolecular peaks of two GGM mannose backbone residues are labelled. Spinning side bands are surrounded by black dotted lines and are marked SSB. The superscript X enables distinction between the two pairs of carbons.
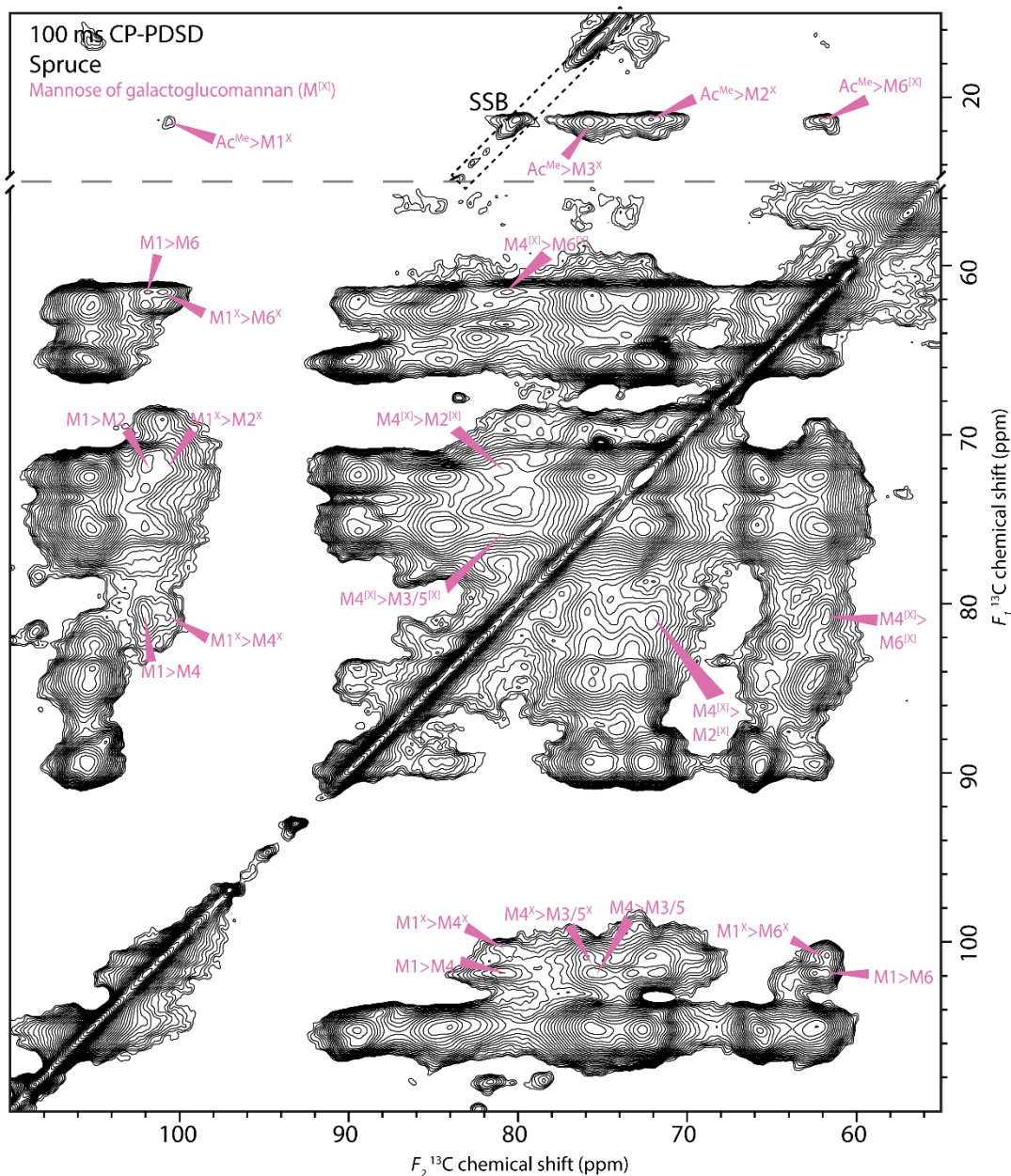
Figure 6.6: M^X is the acetylated mannose residue of GGM. The acetate methyl and carbohydrate regions of a 100 ms mixing time CP-PDSD spectrum of spruce are shown. Spinning side bands are surround by black dotted lines and are marked SSB. The superscript X enables distinction between the two pairs of carbons.

Table 6.2 Spruce galactoglucomannan chemical shifts in solid- and solution-state NMR. Chemical shifts for non-acetylated mannosyl (M) residues and acetylated mannosyl residues (M^{Ac}) are shown. In the upper table, the chemical shift assignments made in this chapter for galactoglucomannan are summarised. Solution-state NMR chemical shifts taken from (Hannuksela and Hervé du Penhoat, 2004) are shown, in grey, in the lower table.

GGM

	1	2	3	4	5	6
M	101.9	71.7	-	80.4	75.8	61.6
M^{Ac}	101.0	72.0	75.9	80.4	75.8	61.6

	1	2	3	4	5	6
M	101.2	71.1-71.5	72.6-73.8	67.6-78.5	76.1-77.5	61.5-61.9
M^{Ac}	99.8-100.6	69.2-72.7	71.2-74.4	74.2-78.5	76.1-76.3	-

The differences in GGM chemical shifts between the solution and solid state are likely to be the result of one or both of these factors. Since the hydroxymethyl conformation alterations also affect the chemical shift of carbon 6, if this were the causal factor, we would expect the M6[Ac] chemical shift to be significantly different from (Hannuksela and Hervé du Penhoat, 2004), but it is similar to these reported values. Thus, it is likely there is a conformational shift of the GGM backbone that causes the differences in chemical shift between the solution and solid-state. Indeed, unbranched two-fold screw mannan in the mannan I crystal allomorph have carbon 1, 4 and 6 chemical shifts similar to the GGM assigned here (Daud and Jarvis, 1992; Heux et al., 2005). In the case of xylan, such a conformational shift is induced by binding to cellulose, and this could be the case with GGM. It is interesting to note that the chemical shift differences between solution and *in muro* GGM are not as large as for xylan (Simmons et al., 2016; Grantham et al., 2017). However, while xylan has been reported to be three-fold in solution, the presence of a carbon 6 in GGM may cause the solution conformation to be intermediate between a two and three-fold screw, reducing the significance of shift differences between solution and *in muro*. Indeed *in silico*, the backbone linkages of GGM have been observed to be less flexible than the backbone linkages of xylan, but still more flexible than the glucan chains of cellulose (Berglund et al., 2016). Now that GGM chemical shifts have been identified in spruce it is possible to investigate its interactions with cellulose. To enable this, the sub-domains of cellulose in spruce need to be annotated.

6.2.2 Identification of cellulose environment diversity in *Picea abies*

To begin to investigate the interactions between hemicelluloses and cellulose it is important to describe the chemical shift diversity of the cellulose environments in spruce. In Arabidopsis primary cell walls, at least 7 cellulose environments have been identified (Wang et al., 2016b), and this level of diversity is probably also present in spruce. Spruce may have a different cellulose structure to eudicots (Fernandes et al., 2011; Thomas et al., 2014).

The CP-INADEQUATE spectrum was used to partially assign cellulose environments in spruce wood, see Figure 6.7. While correlations between carbons 1 and 2 (C1=105ppm, C2=72ppm) and carbons 2 and 3 (C3= 75ppm) are similar between the sub-domains, or in substantially congested areas of the spectrum, there is chemical shift diversity at carbons 4, 5 and 6. The cellulose environments have been previously assigned as
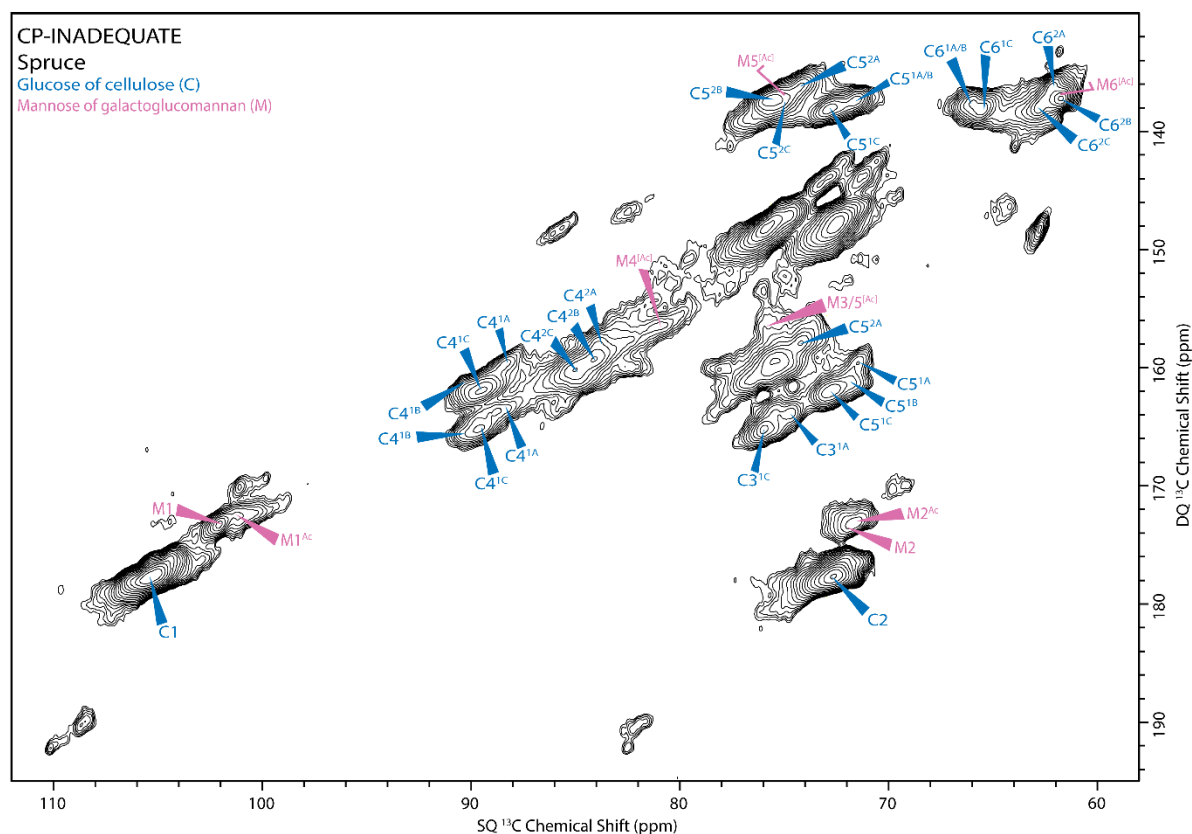
Figure 6.7: Assigning cellulose sub-domains in spruce. The carbohydrate region of a ^{13}C refocussed CP-INADEQUATE spectrum of spruce is shown. The cellulose environments have been split into two domains, 1 and 2, based on (Dupree et al. 2015). The high resolution enables domains 1 and 2 to be split into three sub-domains each, labelled $C^{1A/B/C}$ and $C^{2A/B/C}$. For GGM the acetylated mannose residue carbons are labelled Ac, where they share the same chemical shift as the unacetylated mannose carbons, they are labelled [Ac]. The positions of GGM carbons 5 to 6 are obscured by cellulose peaks in the same region; their theoretical positions, calculated from table 6.2 are labelled with an unfilled arrow. The region SQ=70-80 ppm, DQ=142-152ppm is unlabelled due to overlapping peaks from multiple polysaccharides, but cellulose carbon 2-3 correlations occur in this region.

domain 1 and 2 (C^1 and C^2), based on whether their C4 chemical shift is around 89ppm or 84ppm respectively (Dupree et al., 2015). The subdomains are then labelled $C^{1A, B, C}$ etc. The three C^1 sub-domains are distinguished by their C4-C5 correlations, for C^{1A} and C^{1B} the chemical shifts for C5 and C6 are overlapping, at ~71.7ppm and ~65.5 ppm respectively. The three C^2 sub-domains are distinguished by their C5-C6 correlations, while for C^{2A} and C^{2B} the chemical shift of C4 is similar, at ~83.9 ppm.

To corroborate the assignments of the 6 cellulose sub-domains, and identify the chemical shifts of some unassigned carbons, the 30ms CP-PDSD spectrum was examined for intramolecular cellulose correlations, see Figure 6.8. In this way it was possible to identify specific chemical shifts for C1, 2, 4, 5 and 6 for all 6 sub-domains. In addition, a lower intensity seventh sub-domain, C^{1D}, was identified. For many sub-domains the C3 shares chemical shift similarity with either carbons 2 or 5, thus it was not possible to confidently assign a chemical shift to this carbon, using either the CP-INADEQUATE or 30ms CP-PDSD spectra. The chemical shifts of each cellulose sub-domain are summarised in table 6.3. It is likely that there are more sub-domains, for instance in maize 8 sub-domains were identified (Wang et al., 2016b). To resolve them would require a higher frequency magnetic field or different processing parameters. It could also be the case that freezing and grinding the samples reduces resolution and broadens spectral lines (Atalla and VanderHart, 1999). These physical treatment choices were made for practical reasons, to enable transport and stable sample spinning.

To determine which of the 7 sub-domains of cellulose are adjacent to each other within the microfibril, the 100 ms CP-PDSD spectra was examined for cross peaks between the sub-domains. There are cross peaks between $C4^{2C}$>$C6^{1C}$, $C6^{2C}$>$C4^{1A}$ and $C4^{1B}$, $C4^{2C}$>$C4^{1C}$ and $C4^{1B}$. These cross-peaks show that sub-domain C^{2C} is close to sub-domain $C^{1A, B}$ and ^{C}. Sub-domains C^{2A} and ^{B} are impossible to resolve at C4 so are considered together here. There are cross peaks between $C4^{2A/B}$>$C6^{1C}$, $C6^{2A/B}$>$C6^{1C}$ and $C4^{2A/B}$>$C4^{1C}$. This suggests that subdomains $C^{2A/B}$ are only adjacent to C^{1C} and not C^{1A} or ^{B}. The intermolecular cross peaks between sub-domains within the same domain are too close to intramolecular cross-peaks to make strong statements about their proximity, but it has previously been reported that some sub-domains within the same domain are close to each other (Wang et al., 2016b). The short mixing time of the CP-PDSD and the strong inter-domain cross peaks suggests these contacts are close in space. Such cross peaks likely represent transfer between neighbouring glucan chains or chains in adjacent

155
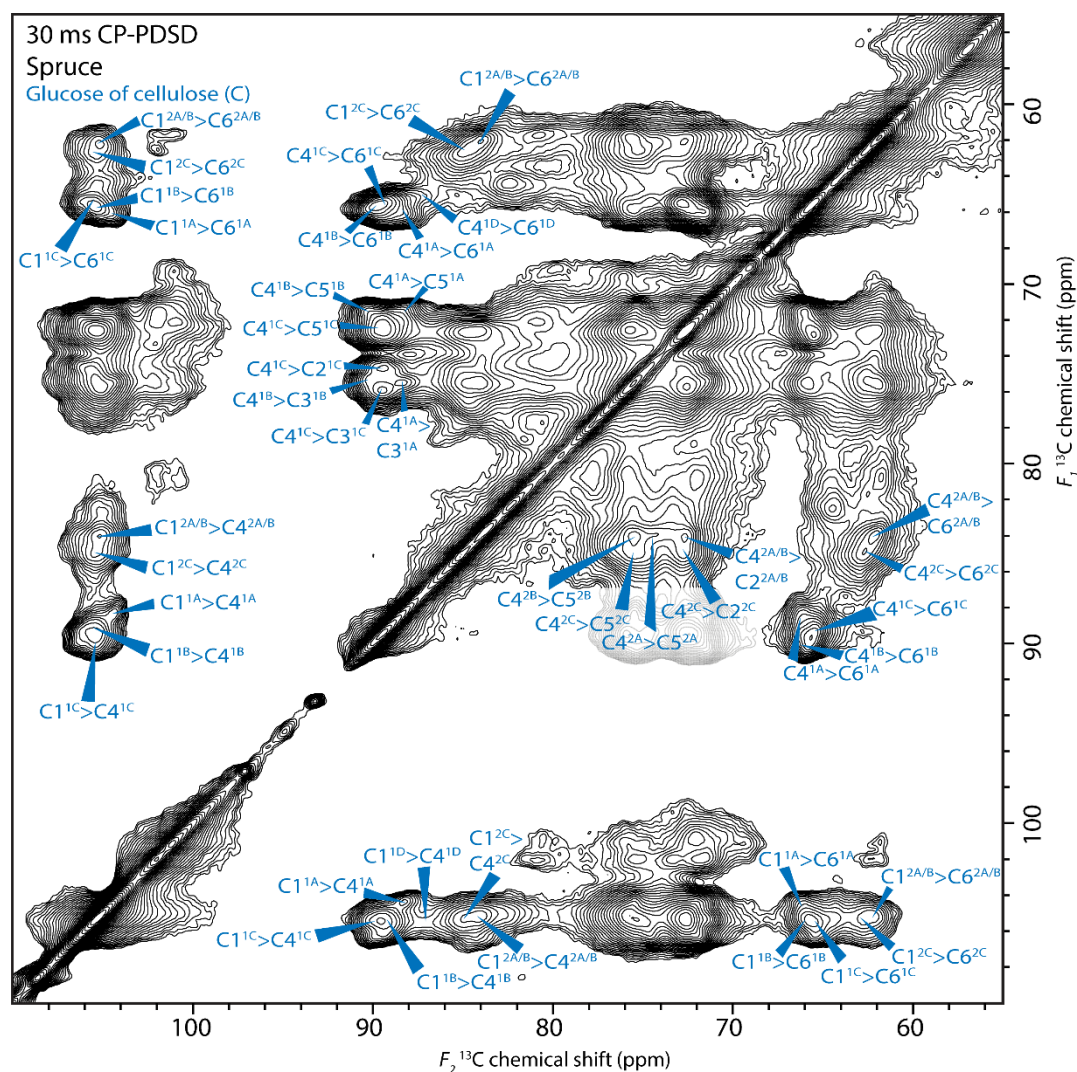
Figure 6.8: Further assignment of the chemical shifts of cellulose in spruce. The carbohydrate region of a 30 ms mixing time CP-PDSD spectrum of spruce is shown. The intramolecular cross-peaks between the carbons of different cellulose sub-domains are labelled. A small portion of the spectrum (~F2=68-80ppm, ~F1=87-92ppm) has been made translucent to create space for annotations.

Table 6.3: Chemical shifts of the 6 dominant cellulose sub-domains. Assignments made from 30 ms CP-PDSD and CP-INADEQUATE spectra of spruce.

Cellulose

	1	2	3	4	5	6
C^{1A}	104.6	–	–	88.7	71.7	65.2
C^{1B}	105.4	71.6	75.3	90.2	71.7	65.8
C^{1C}	105.4	72.4	75.9	89.5	72.8	65.5
C^{2A}	105.3	72.6	75.4	84.1	74.3	62.2
C^{2B}	105.2	–	74.2	83.7	75.6	61.9
C^{2C}	105.3	72.7	75.6	84.8	75.2	62.6
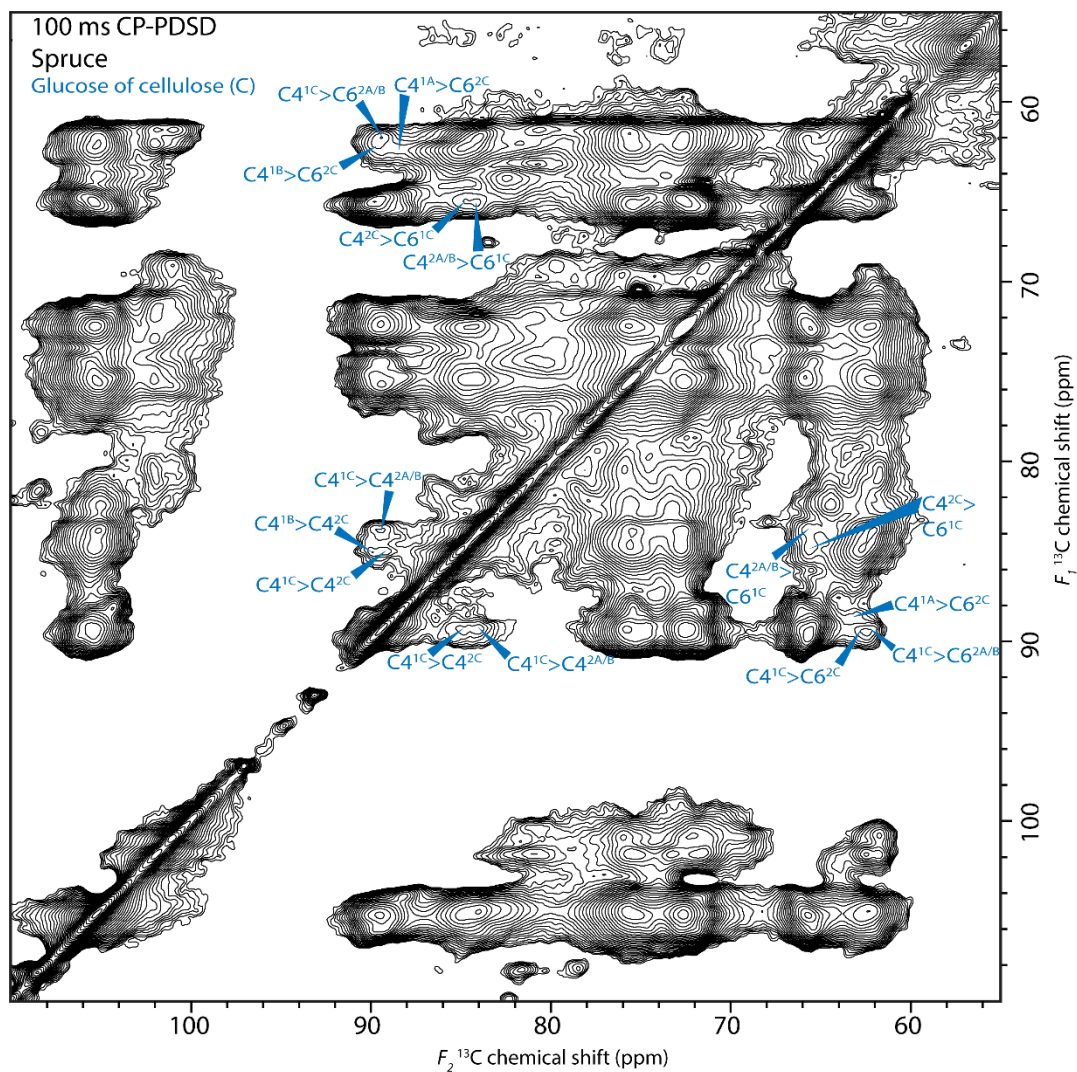
Figure 6.9: Determining the spatial proximity of cellulose sub-domains. The carbohydrate region of a 100 ms mixing time CP-PDSD spectrum of spruce is shown. Cross-peaks between different cellulose sub-domains are labelled.

sheets in the microfibril. Interestingly, while $C2^{2C}$ is close to $C1^{A-C}$, $C2^{2A/B}$ are only close to $C1^{1C}$. This suggests these sub-domains have subtly different locations within the microfibrils, and they may interact differently with hemicelluloses and lignin.

6.2.3 Interactions between cellulose and the hemicelluloses xylan and galactoglucomannan

Having identified the chemical shifts of GGM and the interactions of cellulose sub-domains, the interactions of GGM and xylan with cellulose sub-domains could next be investigated. No cross peaks between hemicelluloses and cellulose were visible in the 100 ms spectra so a longer mixing time (400 ms) CP-PDSD experiment was performed on spruce.

Previously it was shown that two-fold xylan binds to cellulose in Arabidopsis (Simmons et al., 2016), and in Section 6.2.1 two-fold xylan signals were identified in CP-INADEQUATE spectra of spruce. To investigate if two-fold xylan is bound to cellulose in spruce, the 400 ms CP-PDSD spectrum was examined for cross-peaks between xylan and cellulose, Figure 6.10. At 400 ms there are $Xn4^{2f}$>$C4^{1C}$ and $C6^{1C}$ cross-peaks. There is also an $Xn4^{2f}$>$C6^{2C}$ cross-peak, and a poorly resolved $Xn4^{2f}$>$C6^{2A/B}$ cross-peak. These observations are similar to the cross-peaks observed in Arabidopsis, where two-fold xylan was spatially close to both domains 1 and 2 of cellulose in 1 s CP-PDSD spectra (Simmons et al., 2016). This suggests no obvious differences in xylan:cellulose binding between these species are detectable by solid-state NMR. Interestingly, it also suggests that the $C1$ sub-domains of cellulose, which have been proposed to be localised in the interior of the microfibril, can occur on the cellulose surface or that xylan chains may intercalate microfibrils.

In Section 6.2.1 two GGM residues were identified in CP-INADEQUATE and CP-PDSD experiments. The similarity of these chemical shifts to the two-fold screw mannans in mannan I crystals could suggest that spruce GGM binds to cellulose, in a similar fashion to xylan (Daud and Jarvis, 1992). To discover if this GGM binds to cellulose microfibrils, the 400 ms CP-PDSD spectrum was examined for cross-peaks between GGM and cellulose, Figure 6.10. In the carbohydrate region there are $M1^{[Ac]}$>$C4^{1C}$ and $C6^{1C}$ cross-peaks. There are also $M1^{[Ac]}$>$C4^{2A/B}$ cross-peaks, and weaker cross-peaks to $C4^{2C}$. This could suggest GGM can bind a similar set of cellulose sub-domains as xylan, but there is
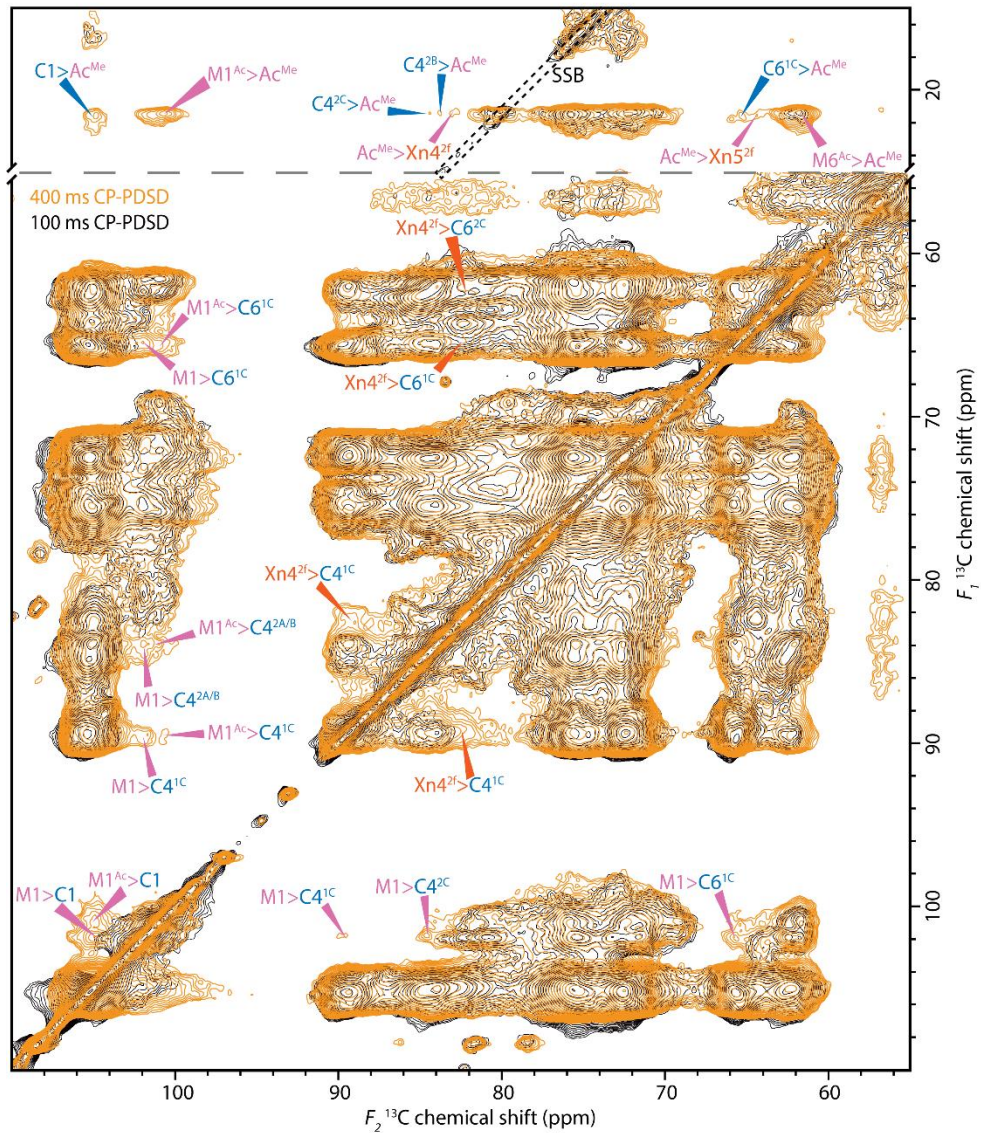
Figure 6.10: Xylan and GGM are close to cellulose sub-domains. The acetate methyl and carbohydrate regions of 100 ms and 400 ms mixing time CP-PDSD spectra are overlaid. Cross-peaks between cellulose and xylan/GGM are labelled. Spinning side bands are surround by black dotted lines and are marked SSB.

a greater preference for GGM to bind $C2^{A/B}$, and for xylan to bind $C2^C$. Generally, the cross-peaks from $M1^{Ac}$ to cellulose sub-domains are weaker than M1 to cellulose. To examine if there is stronger evidence for interactions between M^{Ac} residues and cellulose the Ac^{Me} region was examined for cross-peaks to cellulose, Figure 6.10. There is a strong Ac^{Me}>C1 cross-peak, as well as weaker Ac^{Me}>$C4^{2A/B}$ and $C4^{2C}$ and $C6^{1C}$ cross-peaks, corroborating the interaction of M^{Ac} residues with these sub-domains of cellulose.

Interestingly, there are weak Ac^{Me}>$Xn4/5^{2f}$ cross-peaks in the 400 ms mixing time CP-PDSD spectra. This could suggest that the xylan and GGM that are bound to cellulose are bound to the same microfibrils, as xylan is not acetylated in conifers (Busse-Wicher et al., 2016). To investigate if the cellulose-bound xylan and GGM are bound to the same cellulose microfibrils, a longer mixing time (1500 ms) CP-PDSD spectrum was examined for xylan-GGM cross-peaks, Figure 6.11. At this mixing time there are Ac^{Me}> $Xn4^{2f}$ and M1>$Xn4^{2f}$ cross-peaks. These cross-peaks suggest that at least some of the two-fold xylan and GGM are within 10 Å of each other.

6.2.4 The spatial relationship between lignin and cell wall polysaccharides

Recently it has been shown that there are extensive non-covalent interactions between lignin and xylan in the cell walls of grasses, such as maize, and the eudicot, Arabidopsis (Kang et al., 2019). The proportions of xylan and GGM, as well as the composition of lignin in conifers is significantly different to Angiosperms, thus it is interesting to determine if these compositional differences affect lignin:polysaccharide interactions.

To determine if there are interactions between lignin and polysaccharides in spruce, the 1500 ms mixing time CP-PDSD spectrum were examined for cross-peaks from the lignin methoxyl carbon (lignin CH_3O=56.6 ppm) to polysaccharide carbons. The lignin CH_3O is a strong, well-resolved peak, facilitating analysis. Figure 6.12a shows the carbohydrate region of a 1500ms CP-PDSD spectrum, where extensive cross-peaks between the lignin CH_3O and carbons with chemical shifts between 60-90 ppm (possibly polysaccharide carbons) are clearly visible. The cross-peaks are broad, suggesting they are intramolecular, rather than the sharper peaks of the polysaccharide carbons. Thus, there are likely to be intramolecular cross-peaks between lignin carbons and intermolecular cross-peaks to polysaccharides in this region.

To examine the buildup of magnetisation over time and thereby help determine if there
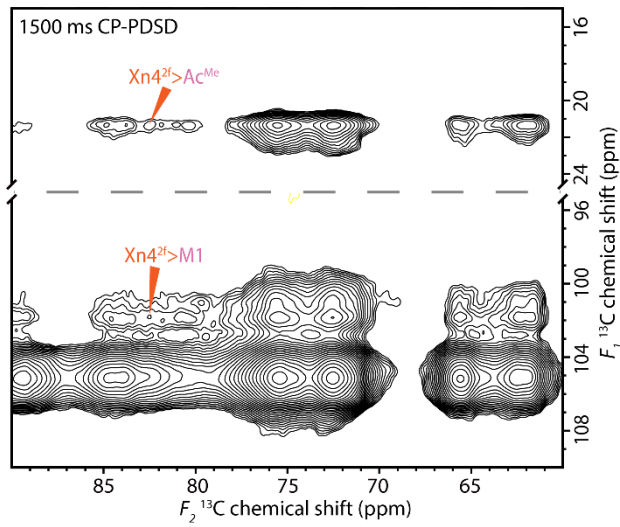
Figure 6.11: Some two-fold xylan and GGM are close to each other. Small sections of the acetate methyl and carbohydrate regions of a 1500 ms mixing time CP-PDSD spectrum of spruce are shown. Cross-peaks between xylan and GGM carbons are labelled.
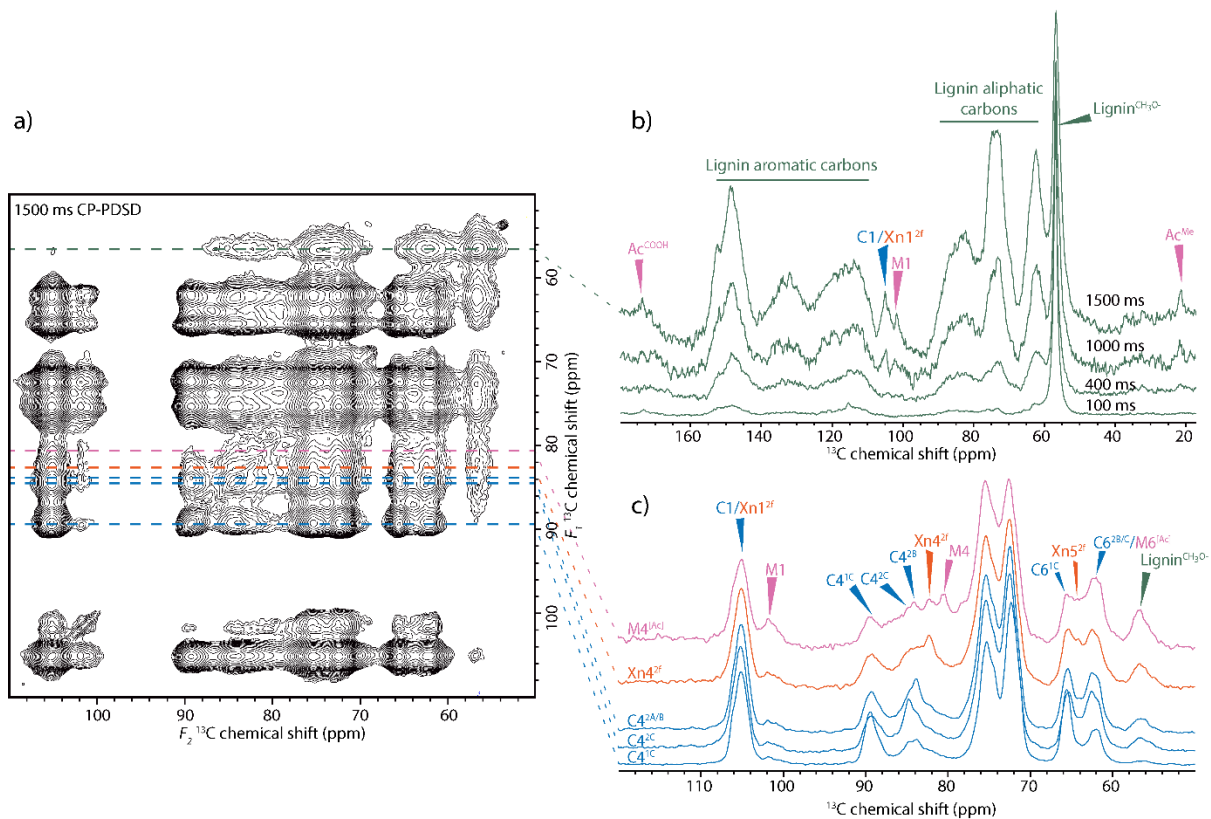
Figure 6.12: Lignin is close to the hemicelluloses and some cellulose sub-domains. a) The carbohydrate region of a 1500 ms CP-PDSD spectrum of spruce is shown. Dotted lines mark the chemical shift where slices have been taken for b) and c). b) The 56.6 ppm slice at different mixing times is shown. Cross-peaks to polysaccharides are labelled. c) The slices of several polysaccharide carbon 4 are shown. Cross-peaks to other polysaccharides and lignin are labelled.

is any lignin-polysaccharide proximity, the 56.6 ppm slices of PDSD spectra at different mixing times are shown in Figure 6.9B. All slices are normalised to the self-peak at 56.6 ppm. In the 100 ms and 400 ms slices, broad cross-peaks to aromatic carbons in the range of 110-50 ppm appear, as well as broad cross-peaks in the range 60-90 ppm. Their appearance at short mixing times and broad line widths are all consistent with their assignment as intramolecular cross-peaks within the lignin polymers, across a range of linkage types (Kim and Ralph, 2010). This suggests lignin is strongly associated with itself. At 1000 ms, weak but sharp cross-peaks from the lignin CH$_3$O->Ac^{Me}, C1/Xn1^{2f} and M1$^{[Ac]}$ all appear. At 1500 ms these cross-peaks become stronger, and a new cross-peak from the lignin CH$_3$O->Ac$^{COO^-}$ appears. These strongly suggest that M$^{[Ac]}$ and cellulose or xylan (or both) are associated with lignin in spruce.

To investigate if both xylan and cellulose are close to lignin, and if there are any differences in cellulose sub-domain associations with lignin, slices were taken at M4$^{[Ac]}$, Xn4^{2f} and C4$^{1C, 2A/B, 2C}$ and examined for cross-peaks to the lignin CH$_3$O- carbon. All slices are normalised to the height of the carbon 4 self-peaks. Confirming the association of M$^{[Ac]}$ with lignin, there is a strong cross-peak to the lignin CH$_3$O- carbon. For xylan, there was also a strong cross-peak to the lignin CH$_3$O- carbon, but this was relatively weaker than the M4$^{[Ac]}$ cross-peak. Interestingly, different cellulose sub-domains had different strength cross-peaks to lignin. Both sub-domains C4$^{2A/B}$ and C4^{2C} had a strong cross-peak to the lignin CH$_3$O- carbon, but this cross-peak is noticeably weaker than that from either xylan or GGM. The C4^{1C} has virtually no cross-peak to the lignin CH$_3$O- carbon, suggesting there is no association between this sub-domain of cellulose and lignin.

6.3 Discussion

In this chapter, the primary aim was to use solid-state NMR to identify the chemical shifts of GGM and xylan and investigate their spatial proximity to cellulose and lignin in Norway spruce, a commercially and scientifically interesting species (Del Lungo et al., 2006; Bradshaw et al., 2009). The focus was on understanding the cell wall molecular architecture, since these interactions are thought to affect cell wall properties such as recalcitrance to enzymatic digestion and mechanical strength (Burton et al., 2010; Terrett and Dupree, 2019). Interestingly, the results challenge the existing models of softwood molecular architecture significantly. Both xylan and GGM bind to cellulose microfibrils in

a similar fashion, and interact with lignin. This new model of softwood molecular architecture may help improve industrial processes.

6.3.1 Xylan interacts with cellulose in spruce, and may have a preference for the hydrophilic face

Previously, xylans have been proposed to bind to the hydrophilic face of cellulose. The oligosaccharides produced by GH30 xylanases have been used to demonstrate that, in eudicots and gymnosperms, glucuronic acid substitutions are predominantly placed on the xylan backbone with an even number of xylosyl residues between consecutive substitutions (Bromley et al., 2013; Busse-Wicher et al., 2016; Martínez-Abad et al., 2017). The even patterning of xylan substitutions extends to acetyl branches in eudicots and arabinosyl substitutions in most gymnosperms (Busse-Wicher et al., 2014; 2016; Martínez-Abad et al., 2017). It was proposed that the evenly spaced substitutions enable xylan to form a two-fold screw on the hydrophilic face of cellulose (Busse-Wicher et al., 2014). In a two-fold screw, the even spacing would mean that the branches project from one side of the xylan molecule, with the other side having unsubstituted hydroxyls. This unsubstituted face imitates a glucan chain of cellulose and could thus hydrogen bond to the hydrophilic face of cellulose. Solid-state NMR studies have shown that two-fold xylan does exist in the cell walls of Arabidopsis, as well as several grasses (Simmons et al., 2016; Kang et al., 2019). This two-fold xylan is within 10 Å of the cellulose surface and does form in the absence of cellulose (Simmons et al., 2016). The *eskimo1* mutant synthesises xylans without evenly spaced substitutions, due to the absence of an acetyltransferase, which regulates the glucuronic acid pattern produced by GUX enzymes (Grantham et al., 2017). This mutant lacks two-fold xylan, providing supporting evidence for the model of evenly patterned xylans binding to the hydrophilic face as a two-fold screw, since hydrophobic face binding would have no requirement for even patterning, and in fact would be enhanced by it (Martínez-Abad et al., 2017). However there have been concerns that the reduced xylan substitution level of *eskimo1* xylan may mean it is partially insoluble, and this could also prevent the xylan from binding to cellulose (Xiong et al., 2015).

The model of xylan binding to the hydrophilic face has been controversial in softwoods. Molecular dynamics simulations by (Martínez-Abad et al., 2017) have suggested that xylan binding to cellulose hydrophilic faces may be stabilised by consecutively spaced

glucuronic acid residues, rather than inhibited by them. In addition, evidence from FT-IR and electron microscopy suggest that xylan in gymnosperms is not bound to cellulose but associates with lignin away from the microfibril surface (Terashima et al., 2009; Simonović et al., 2011). In this chapter it was found that two-fold xylan exists in spruce cell walls, as was predicted in Busse-Wicher et al., (2016). The two-fold xylan peaks identified here are four times as strong as the three-fold xylan peaks in CP-INADEQUATE spectra, similar to wild type Arabidopsis (data not shown). Although multi-dimensional NMR spectra are not directly quantitative, this does suggest that the majority of the xylan exists as the two-fold conformer in spruce. Since the consecutively spaced glucuronic acid residues are relatively rare, but the evenly spaced substitutions are very common, it seems likely that at least some of the evenly spaced substitutions occur on xylan that is in the two-fold conformation (Martínez-Abad et al., 2017). To provide evidence that two-fold screw xylan formation in spruce is dependent on evenly patterned substitutions, spruce could be engineered to synthesise xylans with mostly consecutive glucuronic acid substitutions. This could be achieved by over-expressing clade 2 GUX enzymes from spruce (Łyczakowski, 2018), which produce consecutive glucuronic acid substitutions when over-expressed in Arabidopsis. We would predict that a high amount of consecutive glucuronic acid substitutions would reduce two-fold screw xylan formation in spruce.

There could be some evidence that the two-fold screw xylan in spruce binds to the hydrophilic face of cellulose. Here, in the CP-PDSD spectrum we found strong dipolar cross-peaks between two-fold xylan and cellulose sub-domain C^{1C}. This domain has been proposed to exist in the interior of cellulose microfibrils, and there is some evidence for this location in the microfibril, though this is contradicted by its binding to xylan and other hemicelluloses (Simmons et al., 2016; Dick-Pérez et al., 2011). More recently it has been proposed that the carbon 6 of domain C^1 is in the *tg* conformation, stabilised by hydrogen bonding from adjacent glycan chains (Phyo et al., 2018; Jarvis, 2018). In the interior of the microfibril, the adjacent chains can be other glucan chains, but on the hydrophilic surfaces of cellulose, other polysaccharides, such as xylan, could provide these hydrogen bonds. Therefore, the proximity of two-fold xylan to sub-domain C^{1C} could provide evidence for the formation of two-fold xylan on the hydrophilic face of cellulose, where it converts the adjacent C^{2A-C} glucan chain to sub-domain C^{1C} by stabilising the *tg* conformation of carbon 6, see Figure 6.13. There is some evidence in Arabidopsis supporting this hypothesis. The *eskimo1* mutant, which has barely any two-fold screw xylan, has a lower proportion of domain C^1 (data not shown), though there could be other

165

explanations for this observation such as potential changes in the cellulose microfibril deposition, aggregation or the dwarfing phenotype. As discussed in chapter 4, the *gux1 gux2* and *cad2 cad6* mutants, have a higher proportion of domain C[1], despite significant differences in the proportions of two-fold xylan in these two mutants. In chapter 4 it was suggested that the conformation of glucan chains may also be affected by lignin:polysaccharide interactions.
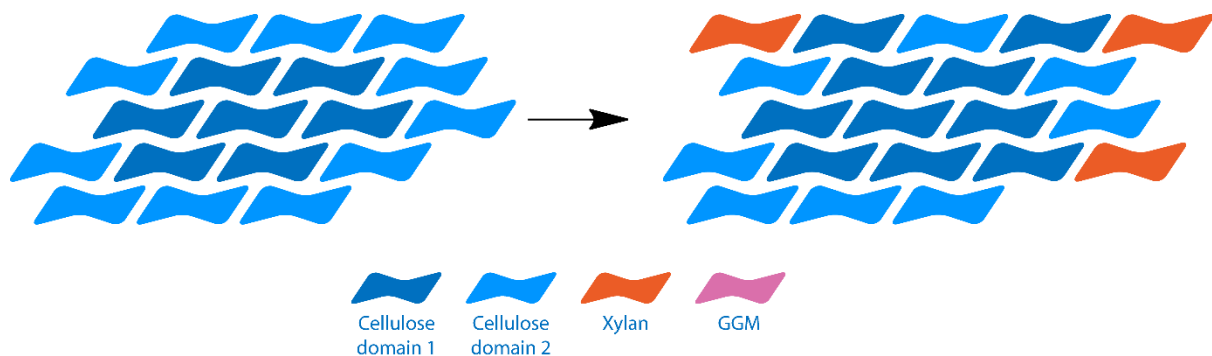
Figure 6.13: Xylan-binding to the hydrophilic face of cellulose could convert domain 2 glucan chains to domain 1. Cellulose is shown with domain 1 glucan chains mainly in the interior of the microfibril. The number of glucan chains of each domain is based on quantification of C4 to C6 cross-peaks in 30 ms CP-PDSD spectra (data not shown). When xylan binds to the hydrophobic face, the hydrogen bonding from xylosyl carbons 2 and 3 stabilises the carbon 6 of the surface glucan chains in the *tg* conformation, converting them from domain 2 to domain 1, see Figure 6.1.

However, with this understanding of the significance of domain 1 versus 2 proximity to hemicellulose chains, there is some evidence that the two-fold xylan binds to the hydrophobic face of cellulose. There are some weak cross-peaks between two-fold xylan and cellulose sub-domains C[2A-C], though these are significantly obscured by intramolecular cross-peaks, possibly from terminal arabinose residues. Sub-domains C[2A-C] are thought to have their carbon 6 in the *gt/gg* conformation, due to a lack of hydrogen bonding to adjacent glycan chains (Phyo et al., 2018; Jarvis, 2018). These sub-domains may exist on the hydrophobic faces of cellulose, indeed water-polarisation experiments suggest these sub-domains occur on the surface of cellulose in spruce (data not shown). The cross-peaks between two-fold xylan and C[2A-C] could suggest that some two-fold xylan binds to the hydrophobic face of cellulose, see Figure 6.14. This hydrophobic-face bound two-fold xylan may be stabilised by the consecutive glucuronic acid substitutions on xylan, as proposed in (Martínez-Abad et al., 2017).
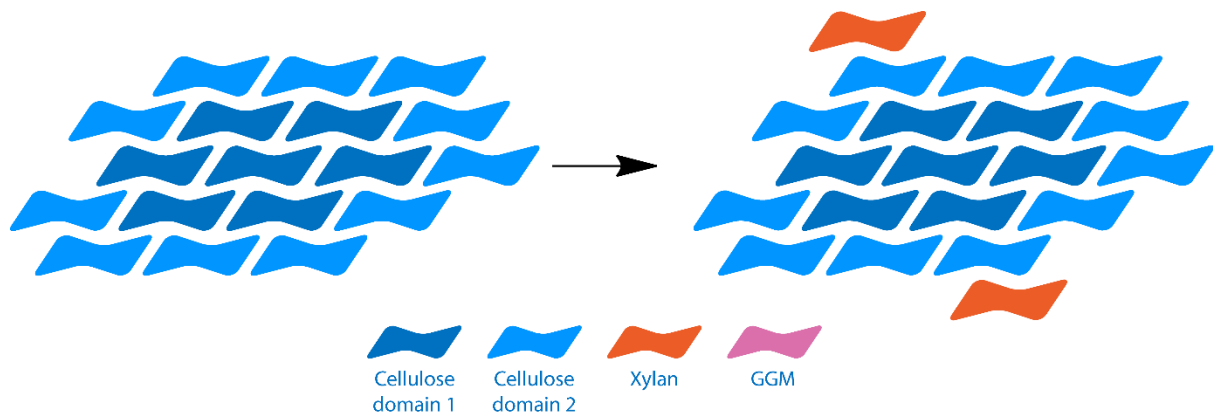
Figure 6.14: Xylan binding to the hydrophobic face may not convert glucan chains from domain 2 to domain 1. Cellulose is shown with domain 1 glucan chains mainly in the interior of the microfibril. When xylan binds to the hydrophobic surface of the microfibril, the interaction is not stabilised by hydrogen bonding, thus the surface glucan chains may stay in their domain 2 conformation.

6.3.2 Galactoglucomannan binds to cellulose, and may bind to both hydrophilic and hydrophobic faces

In vitro, branching of GGM negatively affects binding to cellulose. Binding studies have suggested that lower degrees of GGM galactose and acetyl substitution and higher proportions of glucose are more optimal for mannan binding to cellulose (Whitney et al., 1998; Hannuksela et al., 2002). This suggests that unsubstituted stretches of mannan, or areas of the backbone with many glucose residues bind to cellulose, while branched mannans or high mannose mannans bind less or not at all. This cellulose binding is not necessarily equivalent to the *in muro* sub-nanometre associations that have been observed between xylan and cellulose with solid-state NMR (Simmons et al., 2016; Grantham et al., 2017).

Based on the *in vitro* binding studies the structure of spruce GGM would seem to be incompatible with binding to cellulose. Around half of the mannosyl residues are acetylated, which can occur at carbon 3 as well as 2, and one in ten mannosyl residues can be galactosylated (Hannuksela and Hervé du Penhoat, 2004). In spruce GGM, there are also many consecutive mannose residues, and thus there will always be axial hydroxyls and branches directed at cellulose surfaces. This might lead one to hypothesise that spruce GGM would bind poorly to cellulose (Yu et al., 2018).

167

Despite this hypothesised incompatibility with cellulose binding, the evidence presented here clearly indicates that spruce GGM binds to cellulose *in muro*. There were multiple cross-peaks to cellulose from both identified GGM residues, as well as from the acetate methyl carbon of GGM. These cross-peaks clearly demonstrate that GGM is binding to cellulose, including in acetylated areas of the backbone. As cross-peaks from cellulose to xylan and GGM appear at the same time, it is also clear that the two hemicelluloses are a similar distance from the cellulose surface.

The spruce GGM may form a two-fold screw on the cellulose surface, in a similar fashion to xylan-binding to cellulose (Simmons et al., 2016). The chemical shifts of the cellulose-bound GGM are similar to the mannan I crystal allomorph at carbon 1, 4 and 6 (Daud and Jarvis, 1992; Heux et al., 2005). Mannans that are unbranched, such as ivory nut, or mannans that are debranched, such as alkali-extracted konjac glucomannan, can crystallise in this allomorph, where each mannan chain is in a twofold screw. Branched mannans, such as the spruce GGM here, do not crystallise in this way. Despite this, the similarity of the *in muro* GGM chemical shifts to the mannan I crystal could suggest that the spruce GGM is forming a mannan I-like two-fold screw on the cellulose surface. To investigate the dependency of this conformation on the presence of cellulose, the chemical shifts of GGM in Arabidopsis *irx3* mutants could be investigated. The low content of GGM in Arabidopsis means it is currently not possible to see GGM cross-peaks in 30 ms CP-PDSD spectra, so longer experiments with a higher signal to noise ratio could be performed, or *irx3* could be transformed to produce more GGM.

The structure of mannans may prevent hydrophilic face binding. One of the features of GGM is the axial position of the carbon 2 hydroxyl of backbone mannosyl residues. *In silico* studies have suggested that if the carbon 2 hydroxyl is directed at the cellulose surfaces, the axial position destabilises interactions with the cellulose surface, but only on hydrophilic faces (Yu et al., 2018), and this may explain the *in vitro* observation that high substitution levels and a high proportion of mannosyl residues reduce cellulose-binding (Whitney et al., 1998; Hannuksela et al., 2002). GGMs may be binding to the hydrophobic face in *in vitro* systems. Although proximities between the mannans and cellulose couldn't be investigated, due to the lack of ^{13}C enrichment, there were interesting cellulose structural changes observed by 1D NMR (Whitney et al., 1998). There was an increase in the proportion of domain 2 in the cellulose. This could suggest a reduction in the size of the cellulose fibrils, though this was not observed. Alternatively,

the mannan-binding could alter the conformation of the carbon 6 hydroxymethyl of the cellulose. This is similar to the changes in the proportions of domain 1 and 2 that were observed in the *eskimo1* mutant and discussed in Section 6.3.1, though the change occurs in the reverse direction. The binding of mannan to cellulose may cause the conversion of domain 1 chains to domain 2 in this *in vitro* system. Alternatively, the axial position of the carbon 2 hydroxyl may mean that GGM bound to the hydrophilic face does not cause domain 1 formation, due to insufficient hydrogen bonding.

Unexpectedly, spruce GGM may bind to both hydrophilic and hydrophobic faces. Here it was shown that spruce GGM residues has strong cross-peaks to both domain 1 and domain 2 cellulose. As discussed in Section 6.3.1, proximity to domain 1 cellulose could be indicative of hydrophilic face binding, due to an alteration in the carbon 6 hydroxymethyl conformation of surface cellulose chains. Thus, it could be the case that spruce GGM is binding to the hydrophilic and hydrophobic faces of cellulose, in contrast to the discussion above, based on *in silico* and *in vitro* binding data (Yu et al., 2018; Whitney et al., 1998).

There is a significant disagreement between the *in muro* proximity data presented here, which shows spruce GGM binding to cellulose, and the *in silico* and *in vitro* binding studies which suggest cellulose-binding should be prevented by high degrees of substitution or low amounts of glucosyl residues in the backbone. To test how GGM structure affects cellulose-binding, the structure of Arabidopsis GGM could be altered using synthetic biology, and the *in muro* cellulose-binding properties could be investigated by solid-state NMR. The CSLA enzymes from GT2 synthesise GGM backbones (Goubet et al., 2009). CSLA2 has been reported to synthesise a backbone with a glucosyl-mannosyl repeating unit (Yu et al., 2018). The CSLA9 enzyme is thought to synthesise GGM with some consecutive mannosyl units (Yu, personal communication). In addition, conifers are known to synthesise high mannose GGMs in wood, and thus wood-expressed conifer CSLAs are good candidates for synthesis of high mannose GGMs in Arabidopsis (Hannuksela and Hervé du Penhoat, 2004). Furthermore, the MSR proteins have been reported to modulate the glucose:mannose ratio of GGM produced by CSLA proteins in *Pichia pastoris* (Voiniciuc et al., 2019). The MAGT and TBL substitution enzymes add galactose and acetyl substitutions to mannosyl residues, respectively (Yu et al., 2018; Zhong et al., 2018). Collectively this enzymatic toolbox could be used to synthesise many different GGM structures (e.g. high mannosyl residue content, high galactosyl

substitution) in Arabidopsis secondary cell walls, and their proximities to different cellulose sub-domains could be investigated using solid-state NMR. It could be the case that different GGM structures prevent binding to cellulose or cause a preference for binding to one sub-domain versus another.

It would also be possible to evaluate the biological function of these different GGMs and their binding to cellulose in Arabidopsis. Previously, *csla2/3/9* mutants (lacking detectable GGM) were thought to have no phenotype (Goubet et al., 2009). Recently it has been shown that GGM may have roles in regulating cellulose orientation, and determining the direction of growth in the gravitropism response (Somssich et al., 2019). It is unknown if this function is mediated by binding of mannans to cellulose, but it would be very interesting to use a panel of Arabidopsis transgenic plants, with different GGM structures and potentially different cellulose-binding properties, to determine how GGM binding to different sub-domains of cellulose affects cellulose orientation in the gravitropism response, and thus GGM biological function.

6.3.3 Hemicelluloses and some cellulose domains interact with lignin in spruce cell walls

It has been reported that there are covalent interactions between lignin and hemicelluloses. As discussed in chapter 4, hemicelluloses can cross-link into lignin by re-aromatising intermediates in the formation of β-O-4 linkages between monolignols, resulting in ether and ester bonds to lignin (Terrett and Dupree, 2019). This has most convincingly been demonstrated in the identification of ether linkages between the carbon 6 of mannosyl residues in GGM (Nishimura et al., 2018), but some evidence in chapter 4 and elsewhere suggests that glucuronic acid substitutions of xylan may also be involved in cross-linking between lignin and xylan. These covalent linkages may mediate extensive non-covalent interactions between lignin and hemicelluloses. In grasses, where xylan and lignin are cross-linked by ferulate substitutions (Ralph et al., 1994), solid-state NMR has been used to identify many sub-nanometre contacts between lignin and two-fold xylan, three-fold xylan and cellulose domains that are equivalent to sub-domains C[2A-c] (Kang et al., 2019).

It is not clear if the non-covalent lignin:polysaccharide interactions are mediated by covalent lignin:polysaccharide interactions. Kang et al., (2019) concluded that in grasses

the covalent linkages may not be important for non-covalent interactions, since there were few lignin-ferulic acid cross-peaks. However, these rare linkages may be important for facilitating non-covalent contacts between lignin and xylan. In addition, their scarcity might prevent easy detection in solid-state NMR. Here, we also see multiple cross-peaks between the lignin methoxyl group and xylan and GGM. It would be fascinating to investigate how important covalent cross-linking is for these non-covalent interactions in spruce. In chapter 4, it was shown that the *gux1 gux2* and *cad2 cad6* mutants had weaker lignin:xylan interactions than wild type Arabidopsis. These results suggest that covalent cross-linking is important for non-covalent interactions. To test the idea that ether and ester hemicellulose:lignin linkages are important for these non-covalent interactions pine RNAi CCoAOMT lines could be analysed for proximity between lignin and polysaccharides (Wagner et al., 2011). The additional catechyl units of the lignin in these CCoAOMT-silenced plants would reduce hemicellulose:lignin cross-linking by re-aromatisation of lignification intermediates (Terrett and Dupree, 2019), and they would be expected to have weaker polysaccharide-lignin cross-peaks.

The conformation of GGM and xylan that is bound to cellulose seems to be unaffected by non-covalent interactions with lignin in spruce. Kang et al., (2019) identify a cross-peak from lignin to xylan in an intermediate fold between three-fold and two-fold, as well as the two-fold and three-fold xylan conformers. They suggest the lignin-binding may alter the xylan conformation, though there could also be a contribution to these cross-peaks from the arabinosyl substituents of xylan, which have carbons with similar chemical shifts to the xylan with an "altered" conformation. Here, we find a cross-peak between the lignin and two-fold xylan and GGM. The shifts of the hemicellulose seem to be consistent with the cellulose-bound forms, suggesting that the lignin in spruce does not alter the xylan or GGM conformation in the same way as in grasses. However, the exact experiment performed by Kang et al., (2019) was not replicated here. It would be interesting to perform their lignin-polarisation experiment, to determine which polysaccharides are most easily polarised by lignin, to test if the conformation of a small amount of GGM and xylan is affected by lignin binding.

Kang et al., (2019) identified cross-peaks mostly between S lignin and xylan and suggest that the hydrophobicity of the additional methoxyl groups on S lignin aids association with xylan. However, their assignments of S lignin carbons are likely to be incorrect. The authors assign 148 ppm to the carbon 3 of G units of lignin, and the 153 ppm to the

carbons 3 and 5 of S units of lignin. Here, in spruce, both peaks are visible. It is well established that conifers lack S lignin (Ralph et al., 2004; Wagner et al., 2015; Yue et al., 2016). It is therefore likely that the signals at both 148 ppm and 153 ppm represent methoxylated carbons in both S and G units that are in different linkages, for instance β-O-4 or β-β. Therefore, they are likely observing a preference of certain lignin linkages to associate with polysaccharides. For instance, the β-O-4 linkages can participate in hemicellulose:lignin cross-linking, and their α and γ hydroxyls may allow hydrogen bond mediated associations with polysaccharides in the β-O-4 linkages that are not covalently linked to polysaccharides (Terrett and Dupree, 2019). Here, we observe that the G lignin of conifers associates with both xylan and GGM, with the strongest cross-peaks to GGM. Based on modelling data, it seems unlikely that polysaccharide hydroxyl-lignin methoxyl interactions are important for non-covalent polysaccharide:lignin interactions. *In silico*, the aromatic rings of lignin associate well with polysaccharides via stacking effects rather than via methoxyl-polysaccharide interactions (Besombes and Mazeau, 2005a; 2005b; Houtman and Atalla, 1995). It would be interesting to investigate the hydrophobicity hypothesis of Kang et al., (2019) by comparing the lignin:polysaccharide interactions between wild type pine and the ferulate 5- hydroxylase and caffeic acid O-methyltransferase transgenic pine, which makes S lignin (Wagner et al., 2015). If their model is correct, the pine with S lignin should have polysaccharide-lignin cross-peaks at shorter mixing times in CP-PDSD experiments.

There are interesting differences in the cross-peaks from lignin to different cellulose environments. There are cross-peaks from lignin to C^{2A-C} but a very weak cross-peak to C^{1C}. As has been discussed in Sections 6.3.1 to 6.3.2 C^1 was thought to occur only in the interior of cellulose microfibrils (Jarvis, 2018). Data presented here suggests that C^{1C} can also occur on the surface of cellulose, shown by cross-peaks to both xylan and GGM. It was speculated that these hemicellulose-C^{1C} cross-peaks could indicate that hemicelluloses binding to the hydrophilic face induce the "surface" C^{2A-C} glucan chains to adopt an "interior"-like carbon 6 hydroxymethyl conformation, converting them to C^{1C}. The fact that lignin does not have cross-peaks to C^{1C} could support this idea, as it suggests that C^{1C} is not exposed on the cellulose surface. Water polarisation evidence also suggests that C^{1C} is not exposed on the cellulose surface in spruce (data not shown). This is consistent with the idea that hemicellulose binding to the hydrophilic face induces a change in conformation of the carbon 6. If the C^{1C} only occurs on the microfibril surface

when bound by hemicelluloses, surface C^{1C} will be partially concealed from lignin/water by hemicellulose chains.

6.3.4 Solid-state NMR data contradicts several existing models of gymnosperm cell walls

Existing studies of the molecular architecture of conifer cell walls are based on methods such as electron microscopy, AFM and diffraction (of x-rays or neutrons) which do not provide chemical information, or methods like FT-IR with ambiguous chemical information (Keplinger et al., 2014; Terashima et al., 2004; 2009; Simonović et al., 2011; Salmén, 2015). The models produced by these studies have several features that contradict the data presented here, discussed below. Multi-dimensional solid-state NMR provides structural and chemical information at sub-nanometre resolution. This combination means that solid-state NMR conclusions about proximities between different cell wall components are more robust than many other biophysical methods.

One conclusion of the models produced by FT-IR and EM-based studies is that while mannan is bound to the surface of cellulose, xylan associates with lignin, away from the cellulose surface (Terashima et al., 2009; Simonović et al., 2011; Salmén, 2015). Here, we show that this aspect of these models may be incorrect, Figure 6.15. A large amount of both xylan and GGM binds to the cellulose surface, meaning that much of the cellulose microfibril surface is bound by both hemicelluloses. In addition, both hemicelluloses have strong cross-peaks to lignin, suggesting that lignin-binding is not restricted to xylan and cellulose-binding is not restricted to GGM.

Any significant hemicellulose binding to cellulose seems to be inconsistent with some non-NMR experimental data (Fernandes et al., 2011). SANS experiments measure the distance between the centre of adjacent cellulose microfibrils, while x-ray diffraction can measure the width of the microfibril. As the SANS spacing is only slightly larger than the diffraction width of microfibrils in spruce, it has been suggested that most of the cellulose microfibrils are too close together to allow significant hemicellulose binding. To be exact, on average there is space for one layer of hemicellulose between 24 chain microfibrils. However, hemicellulose binding may extend the apparent microfibril size, due to binding to the microfibril surface in a way that is similar to the binding of glucan chains to each other (Jarvis, 2018). This interpretation is more consistent with the observed size of

173

cellulose synthase complexes (Turner and Kumar, 2018). CesA complexes are thought to contain 18 CesA monomers, and thus presumably make microfibrils with 18 chains. The microfibril extension by hemicelluloses may make the 18 chain microfibril size appear to be 24 chains, consistent with the apparently large amount of cellulose-bound hemicelluloses. The remaining space between these hemicellulose:cellulose complexes may contain some lignin or hemicelluloses in other conformations.
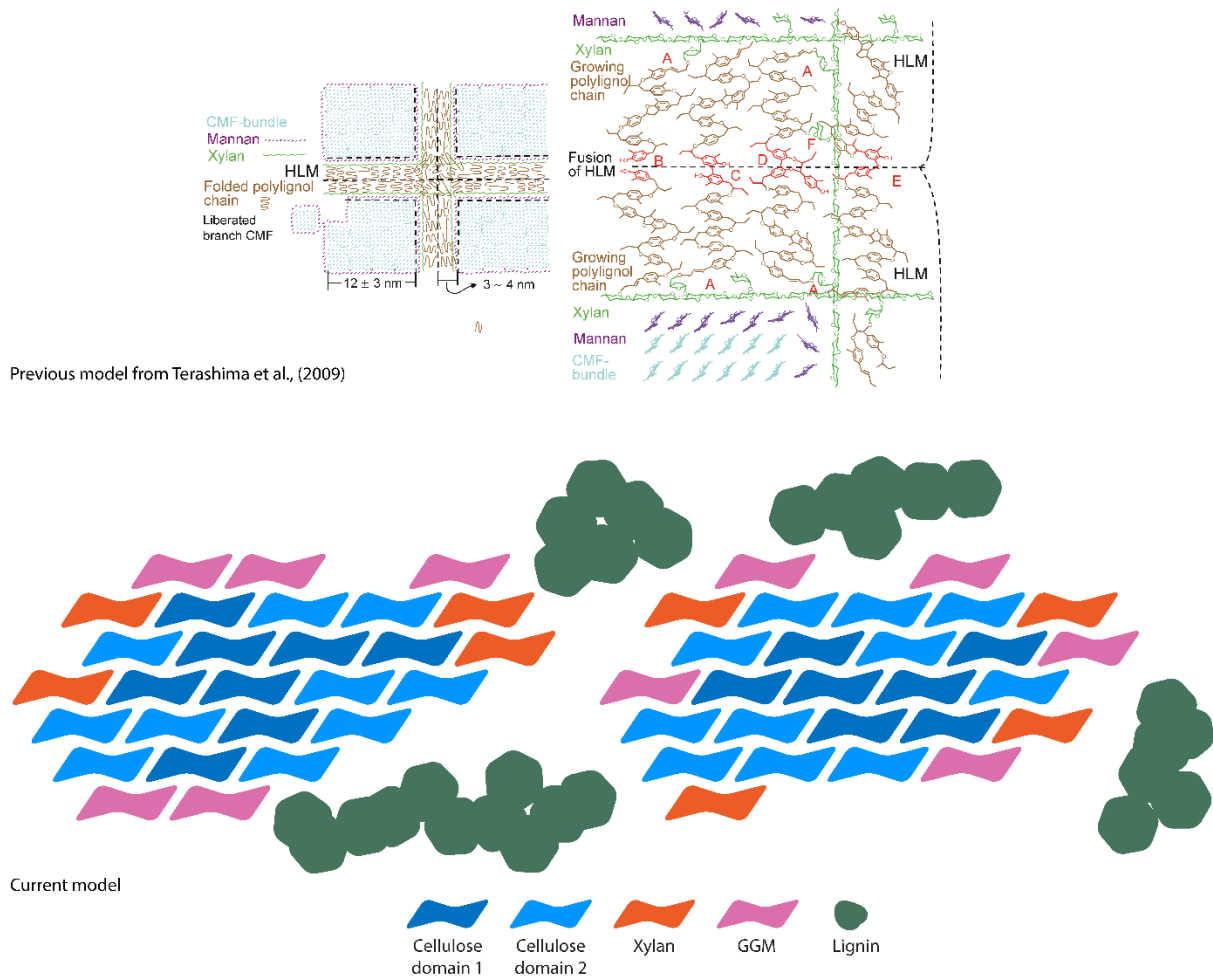
Figure 6.15: Previous models of gymnosperm cell walls and a model based on the data presented in this chapter. Previous models, taken from Terashima et al., (2009), of gymnosperm cell walls show cellulose microfibrils mostly binding to each other with a surface layer of bound mannan/GGM. The xylan associates with lignin slightly away from the cellulose macrofibril surface. In our current model, the relative amounts of cellulose domain 1 and 2 as well as xylan and GGM are based on the volume of carbon 4-6 and 4-5 cross-peaks in 30ms CP-PDSD exeriments. The cellulose sub-domains are grouped into domains due to the inaccuracy of quanitifying the volumes of overlapping peaks. The xylan and GGM are shown bound to the cellulose microfibril surface, and lignin is shown in proximity to xylan, mannan and domain 2 cellulose. As conclusive evidence for binding

to specific cellulose faces has not been found here, GGM and xylan are shown on both faces. The exact distribution of domain 1 and 2 in the microfibril is unknown, but here domain 1 is only shown on the microfibril surface when covered by a hemicellulose chain. The exact shape of the microfibril is unknown, but one has been selected based on evidence from a recent publication (Kubicki et al., 2018).

Chapter 7: Discussion

7.1 The problem of cell wall molecular architecture

Expanded utilisation of lignocellulosic biomass could be an important aspect of building a more sustainable economy (Masson-Delmotte et al., 2018). Lignocellulose may replace fossil fuels used for materials, fuels and pharmaceuticals and could also become more dominant as a construction material (McCann and Carpita, 2015). The properties of lignocellulose are directly related to its composition and molecular architecture (Burton et al., 2010). To inform genetic or chemical strategies to alter these properties for specific applications, it is crucial to understand how cell wall molecular architecture leads to lignocellulose properties such as mechanical strength and recalcitrance to enzymatic digestion (McCann and Carpita, 2015).

The principal aims of this thesis were to investigate the role and synthesis of xylan:lignin cross-links, and to investigate the interactions of hemicelluloses with other cell wall components in conifer lignocellulose.

7.1.1 The role of xylan:lignin cross-linking in recalcitrance to enzymatic digestion

Previously, it was reported that the *gux1 gux2* mutant, lacking glucuronic acid on secondary cell wall xylan, has a dramatic reduction in the recalcitrance of the cell wall to enzymatic digestion (Lyczakowski et al., 2017). In addition, it was found that lignin and xylan are more extractible from this mutant using DMSO, suggesting that the strength of interactions between cell wall components are affected (Busse-Wicher, personal communication). In this thesis, and elsewhere, it was hypothesised that the glucuronic acid could act as a nucleophile that re-aromatises the quinone methide intermediates of lignification (Mottiar et al., 2016; Terrett and Dupree, 2019) and that the cross-linking between xylan and lignin, formed in this process, is important for cell wall recalcitrance.

To investigate this hypothesis, the *gux1 gux2* mutant was compared with the *cad2 cad6* mutant. The hydroxycinnamaldehyde lignin of *cad2 cad6* does not require nucleophilic attack to re-aromatise quinone methides, and thus should lack glucuronic acid mediated lignin:xylan cross-links (Kim et al., 2003). In chapters 3 and 4 it was found that the *gux1 gux2* mutant has an increased volume of enzyme-accessible nanoscale

pores. Solid-state NMR showed that there are weaker *in muro* associations between xylan and lignin. This probably leads to an increased enzyme accessibility of *gux1 gux2* xylan. Interestingly, the *cad2 cad6* mutant shares many of these phenotypes. Although the porosity of this mutant has not been assessed yet, it was found that *cad2 cad6* also has weaker associations between xylan and lignin, and an increase in the enzyme accessibility of xylan. This effect was notable when compared to other lignin mutants, which have similar changes in lignin content or in monolignol ratios, but do not share the phenotypes of *gux1 gux2* and *cad2 cad6*. These results strongly suggest that lignin-glucuronate esters have important effects on the assembly of xylan and lignin. The coating of xylan by lignin, due to these cross-links, seems to have important functional consequences on the digestibility of the cell wall.

Some progress was made in analysing lignin-glucuronate esters using PACE and NMR. Unfortunately, it was not yet possible convincingly to verify their existence with through-space correlations in NOESY NMR experiments (Nishimura et al., 2018). This appears to currently be the result of inefficient NOE transfer in these experiments. Adapting the solvent conditions, such as the inclusion of deuterated pyridine (Kim and Ralph, 2010), may facilitate more efficient NOE transfer and therefore complete analysis of these samples in the future.

There are many future directions for this area of cell wall biology. It may be important to characterise the frequency of lignin-glucuronate esters, relative to the amount of glucuronic acid substitutions on xylan. It is currently unknown if the number of such cross-links varies with the different lignin compositions found in different phylogenetic groups, such as the G lignin of conifers, or the more heavily acylated lignins of grasses. It would be interesting to identify the relative position of such cross-links within the lignin polymer. The sensitivity of lignin-glucuronate esters to carbohydrate esterases could also be investigated. Together this further characterisation of lignin-glucuronate cross-linking could inform novel strategies aimed at breeding biomass with desired properties or using genetic engineering to alter lignocellulose characteristics.

7.1.2   Molecular architecture of conifer wood

Conifers are notable for their economic and ecological importance (Bradshaw et al., 2009; Del Lungo et al., 2006). Their cell wall composition is considerably different to

177

other higher plants (Scheller and Ulvskov, 2010; Busse-Wicher et al., 2016) in particular, unlike in other plants, conifer wood is rich in GGM. Analysis of conifer lignocellulose by solid-state NMR provides the opportunity to compare the interactions of GGM and xylan with other cell wall components in the same species. In addition, the unique composition of conifer lignin, composed of 95% G units, provides the opportunity to investigate the interactions of this unique lignin with other cell wall components (Ralph et al., 2004).

Conifer xylan is unacetylated, unlike monocots and eudicots, but also exhibits an even number of xylosyl residues between most consecutive substitutions (Busse-Wicher et al., 2016; Martínez-Abad et al., 2017). The even spacing of xylan substitutions has been suggested to be an important structural feature that enables xylan to bind to the hydrophilic face of cellulose as a two-fold screw (Busse-Wicher et al., 2014; 2016). In Arabidopsis, xylan has been found to exist as a two-fold screw (Simmons et al., 2016). Importantly, the *esk1* mutant, which lacks evenly spaced substitutions does not have two-fold xylan (Grantham et al., 2017). In chapter 6 of this thesis it was found that xylan in conifers binds to cellulose as a two-fold screw. Along with the recent paper by (Kang et al., 2019), this confirms that two-fold xylan binding to cellulose is a widespread phenomenon in Angiosperms and conifers.

The major hemicellulose in conifers is GGM (Scheller and Ulvskov, 2010). The structure of conifer GGM is strikingly different from conifer xylan (Hannuksela and Hervé du Penhoat, 2004; Busse-Wicher et al., 2016) with 80% of the GGM backbone formed from mannosyl residues which, unlike xylosyl or glucosyl resdiues, has an axial hydroxyl on carbon 2. In addition, mannosyl carbon 2 and 3 can be acetylated and carbon 6 galactosylated (Scheller and Ulvskov, 2010). Unlike in xylan, these GGM substitutions are arranged randomly (Hannuksela and Hervé du Penhoat, 2004), and therefore one might hypothesise that the binding of GGM to cellulose would differ significantly from xylan binding. Here, the *in muro* GGM chemical shifts were identified, with the assistance of Arabidopsis molecular genetics. The GGM was found to be equally distant from the cellulose surface as xylan, suggesting their behaviour *in muro* may be similar. There were convincing cross-peaks between GGM and both domains of cellulose, while xylan also probably has cross-peaks to both domains of cellulose. Thus it is currently unclear how the cellulose-binding properties of mannan and xylan differ in the wall.

Conifer lignin contains no S units, and is mainly G units. It was concluded that the lignin:xylan non-covalent interactions in monocots mainly occur between xylan and S units of lignin (Kang et al., 2019). As discussed in chapter 6, this interpretation is incorrect, as the assignment of S and G subunits was not correct. However, the composition of lignin affects the relative amounts of interunit linkages (Wagner et al., 2015; Shi et al., 2016), and thus the different structures of conifer lignin may affect lignin:polysaccharide interactions. In chapter 6 it was found that some lignin is close to both hemicelluloses, and one domain of cellulose. These non-covalent interactions may be mediated by covalent linkages, such as the glucuronic acid mediated xylan:lignin linkages that were investigated in chapters 3 and 4, or the GGM linkages published in (Nishimura et al., 2018).

The results discussed in this Section and the previous Section may have important industrial implications. For instance, lignin:xylan cross-linking through glucuronic acid is important for the recalcitrance of the cell wall to enzymatic digestion. Increasing the frequency of such cross-links could improve the resistance of crops to pathogens (Reem et al., 2016; Bellincampi et al., 2014). Decreasing such cross-links could improve the economic viability of biofuels or reduce the environmental impact of xylan removal in paper pulping (McCann and Carpita, 2015; Dimmel et al., 2001). The investigations on the molecular architecture of conifer wood suggest that both xylan and GGM are important targets in softwood for modifying lignocellulose properties.

7.2   Further work

The results described here may need further work to fortify the conclusions described. Specifically, it will be important to further characterise *cad2 cad6* mutants, and structurally characterise the potential LCCs using solution-state NMR. However, the work described here also suggests new research questions, which are discussed below.

7.2.1 Are there functional differences between different types of xylan:lignin cross-linking?

There seems to be good evidence that lignin:polysaccharide cross-linking is important for cell wall recalcitrance (Chiniquy et al., 2012; Lyczakowski et al., 2017). This cross-linking can proceed via two different mechanisms. The re-aromatisation of quinone methides

179

by polysaccharide nucleophiles discussed in chapter 4, but also the participation of xylan ferulate substitutions in lignification, discussed in Section 1.3.2 and chapter 5 (Terrett and Dupree, 2019). Ferulate substitution is restricted to commelinid monocots (Buanafina, 2009). However re-aromatisation of quinone methides by polysaccharides can occur in any plant that has lignin mainly composed of conventional monolignols (Ralph et al., 2019).

Could there be a functional difference between these two modes of lignin:xylan cross-linking? Ferulate cross-links were originally proposed to be lignification nucleation sites (Ralph et al., 1995), suggesting that lignins can only extend away from a ferulate substitution. When polysaccharide nucleophiles re-aromatise quinone methide intermediates, the lignin unit formed can likely be extended in two directions (Ralph et al., 2019). Thus lignin that is ferulate cross-linked to xylan may align poorly with the xylan chain it is attached to, while lignin that is cross-linked via glucuronic acid (or some other polysaccharide nucleophile) may be oriented in a more linear fashion along the xylan chain. Thus these two types of cross-linking could have different effects on cell wall assembly. In commelinid monocots, the phylogenetic group that contains most crops, both of these types of cross-linking could occur. Investigation of the functional differences between these two types of cross-linking may facilitate deeper understanding of the ecological success of this group (Linder et al., 2018), or improvement of their agronomic properties.

It would be interesting to investigate to what extent these substitutions play equivalent roles in the cell well. This idea could be tested by increasing the glucuronic acid substitution level of commelinind monocot plants with reduced ferulation, such as *xax1 Oryza sativa* (Chiniquy et al., 2012) or RNAi-BAHD01 *Setaria viridis* (de Souza et al., 2018). Overexpression of GUX enzymes in Arabidopsis has been an effective approach to altering xylan substitution levels (Lyczakowski et al., 2017). The *gux1 gux2* and *cad2 cad6* phenotypes described in this thesis (nanoscale porosity, enzyme accessibility of xylan, and *in muro* lignin:xylan interactions) could all be evaluated in these mutants and their *GUX*-overexpression lines. Conversely, the approach tested in chapter 5, of introducing ferulate substitutions into *gux1 gux2* or *cad2 cad6* plants could be continued. Modifications to this strategy could involve using other GT61 candidates instead of XAX1, or overexpressing CCoAOMT genes to enhance the feruloyl CoA supply (Tetreault et al., 2018). The recalcitrance phenotypes, enzyme accessibility, and

lignin:xylan interactions of *cad2 cad6* or *gux1 gux2* mutants complemented in this way could be tested.

7.2.2   What are the structural determinants of GGM binding to cellulose?

In Arabidopsis, the unique evenly patterned structures of xylan seem to be important for xylan function. In wild type Arabidopsos, the xylan binds to cellulose as a two-fold screw (Simmons et al., 2016), but in *eskimo1*, xylan is only found in the three-fold conformation and the plant has collapsed xylem vessels (Grantham et al., 2017; Xiong et al., 2013). This suggests that a specific pattern of xylan substitutions is important for cell wall mechanical properties.

In contrast to xylan, the substitutions of conifer GGM seem to be unpatterned, and the backbone residues are also not patterned (Hannuksela and Hervé du Penhoat, 2004). It could be the case that there is some type of substitution patterning which has not yet been observed, for instance in the distribution of acetyl groups. It would be extremely interesting to assess what structural features of GGM are important for binding to cellulose, by producing Arabidopsis plants with differently structured GGM polymers, and monitoring the binding to cellulose with solid-state NMR. This analysis could be coupled with molecular dynamics studies to explain the cellulose-binding behaviour of differently structured GGMs (Busse-Wicher et al., 2014; 2016; Yu et al., 2018).

In addition, it would be interesting to determine whether GGM and xylan are functionally equivalent. To test this, reduced xylan Arabidopsis mutants could be transformed to produce more GGM. The extra GGM may bind to cellulose, effectively replacing the xylan. If there are differences in the function of cellulose-binding between xylan and GGM, such plants may still have collapsed xylem vessels and a dwarfed phenotype (Brown et al., 2007). Further analysis of GGM function is of particular importance for the industrial application of softwood where this polysaccharide is the dominant wood hemicellulose and understanding its role could enable breeding of varieties more resistant to pathogens or with timber more amendable to chemical or enzymatic processing.

Bibliography

Alejandro, S., Lee, Y., Takayuki, T., Sudre, D., Osorio, S., Park, J., Bovet, L., Lee, Y., Geldner, N., Fernie, A. R., Martinoia, E. 2012. "AtABCG29 Is a Monolignol Transporter Involved in Lignin Biosynthesis." *Current Biology: CB* 22 (13): 1207–12.

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., Shrewry, P. R., Dupree, P., 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): 989–93.

Anderson, N. A., Tobimatsu, Y., Ciesielski, P. N., Ximenes, E., Ralph, J., Donohoe, B. S., Ladisch, M., Chapple, C. 2015. "Manipulation of Guaiacyl and Syringyl Monomer Biosynthesis in an Arabidopsis Cinnamyl Alcohol Dehydrogenase Mutant Results in Atypical Lignin Biosynthesis and Modified Cell Wall Structure." *The Plant Cell* 27 (8): 2195–2209.

Andlar, M., Rezić, T., Marđetko, N., Kracher, D., Ludwig, R., Šantek, B. 2018. "Lignocellulose Degradation: An Overview of Fungi and Fungal Enzymes Involved in Lignocellulose Degradation." *Engineering in Life Sciences* 18 (11): 768–78.

Arnling Bååth, J., Giummarella, N., Klaubauf, S., Lawoko, M., Olsson, L. 2016. "A Glucuronoyl Esterase from Acremonium Alcalophilum Cleaves Native Lignin-Carbohydrate Ester Bonds." *FEBS Letters*, 590 (16): 2611-8.

Arnling Bååth, J., Mazurkewich, S., Knudsen, R. M., Poulsen, J-C. N., Olsson, L., Leggio, L. L., Larsbrink, J. 2018. "Biochemical and Structural Features of Diverse Bacterial Glucuronoyl Esterases Facilitating Recalcitrant Biomass Conversion." *Biotechnology for Biofuels* 11 (1): 213.

Assis, T. d., Huang, S., Driemeier, C. E., Donohoe, B. S., Kim, C., Kim, S. H., Gonzalez, R., Jameel, H., Park, S. 2018. "Toward an Understanding of the Increase in Enzymatic Hydrolysis by Mechanical Refining." *Biotechnology for Biofuels* 11 (1): 289.

Atalla, R. H., VanderHart, D. L. 1999. "The Role of Solid State ^{13}C NMR Spectroscopy in Studies of the Nature of Native Celluloses." *Solid State Nuclear Magnetic Resonance* 15 (1): 1–19.

Atalla, R. H., Vanderhart, D. L. 1984. "Native Cellulose: A Composite of Two Distinct Crystalline Forms." *Science* 223 (4633): 283–85.

Atmodjo, M. A., Hao, Z., Mohnen, D. 2013. "Evolving Views of Pectin Biosynthesis." *Annual Review of Plant Biology* 64 (1): 747–79.

Atmodjo, M. A., Sakuragi, Y., Zhu, X., Burrell, A. J., Mohanty, S. S., Atwood, J. A., Orlando, R., Scheller, H. V., Mohnen, D. 2011. "Galacturonosyltransferase (GAUT)1 and GAUT7 Are the Core of a Plant Cell Wall Pectin Biosynthetic Homogalacturonan:Galacturonosyltransferase Complex." *Proceedings of the National Academy of Sciences* 108 (50): 20225–30.

Aznar, A., Chalvin, C., Shih, P. M., Maimann, M., Ebert, B., Birdseye, D. S., Loqué, D., Scheller, H. V. 2018. "Gene Stacking of Multiple Traits for High Yield of Fermentable Sugars in Plant Biomass." *Biotechnology for Biofuels* 11 (1): 2.

Barbosa, I. C. R., Rojas-Murcia, N., Geldner, N. 2019. "The Casparian Strip—One Ring to Bring Cell Biology to Lignification?" *Current Opinion in Biotechnology* 56 (April): 121–29.

Bar-On, Y. M., Phillips, R., Milo, R. 2018. "The Biomass Distribution on Earth." *Proceedings of the National Academy of Sciences* 115 (25): 6506–11.

Barros, J., Escamilla-Trevino, L., Song, L., Rao, X., Serrani-Yarce, J. C., Palacios, M. D., Engle, N., Choudhury, F. K., Tschaplinski, T. J., Venables, B. J., Mittler, R., Dixon, R. A. 2019. "4-Coumarate 3-Hydroxylase in the Lignin Biosynthesis Pathway Is a Cytosolic Ascorbate Peroxidase." *Nature Communications* 10 (1): 1994.

Barros, J., Serrani-Yarce, J. C., Chen, F., Baxter, D., Venables, B. J., Dixon, R. A. 2016. "Role of Bifunctional Ammonia-Lyase in Grass Cell Wall Biosynthesis." *Nature Plants* 2 (6): 16050.

Barros, J., Temple, S., Dixon, R. A. 2019. "Development and Commercialization of Reduced Lignin Alfalfa." *Current Opinion in Biotechnology* 56 (April): 48–54.

Bartley, L. E., Peck, M. L., Kim, S. R., Ebert, B., Manisseri, C., Chiniquy, D. M., Sykes, R., Gao, L., Rautengarten, C., Vega-Sánchez, M. E., Benke, P. I., Canlas, P. E., Cao P., Brewer, S., Lin, F., Smith, W. L., Zhang, X., Keasling, J. D., Jentoff, R. E., Foster, S. B., Zhou, J., Ziebell, A., An, G., Scheller, H. V., Ronald, P. C. 2013. "Overexpression of a BAHD

Acyltransferase, OsAt10, Alters Rice Cell Wall Hydroxycinnamic Acid Content and Saccharification." *Plant Physiology* 161 (4): 1615–33.

Basu, D., Liang, Y., Liu, X., Himmeldirk, K., Faik, A., Kieliszewski, M., Held, M., Showalter, A. M. 2013. "Functional Identification of a Hydroxyproline-O-Galactosyltransferase Specific for Arabinogalactan Protein Biosynthesis in Arabidopsis." *Journal of Biological Chemistry* 288 (14): 10132–43.

Bellincampi, D., Cervone, F., Lionetti, V. 2014. "Plant Cell Wall Dynamics and Wall-Related Susceptibility in Plant–Pathogen Interactions." *Frontiers in Plant Science* 5: 228.

Bennici, A. 2008. "Origin and Early Evolution of Land Plants." *Communicative & Integrative Biology* 1 (2): 212–18.

Bensussan, M., Lefebvre, V., Ducamp, A., Trouverie, J., Gineau, E., Fortabat, M-N., Guillebaux, A., Baldy, A., Naquin, D., Herbette, S., Lapierre, C., Mouille, G., Horlow, C., Durand-Tardif, M. 2015. "Suppression of Dwarf and Irregular Xylem Phenotypes Generates Low-Acetylated Biomass Lines in Arabidopsis." *Plant Physiology* 168 (2): 452–63.

Berglund, J., d'Ortoli, T. A., Vilaplana, F., Widmalm, G., Bergenstråhle-Wohlert, M., Lawoko, M., Henriksson, G., Lindström, M., Wohlert, J. 2016. "A Molecular Dynamics Study of the Effect of Glycosidic Linkage Type in the Hemicellulose Backbone on the Molecular Chain Flexibility." *The Plant Journal* 88 (1): 56-70.

Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Bris, P. L., Borrega, N., Hervé, J., Blondet, E., Balzergue, S., Lapierre, C., Jouanin, L. 2011. "Disruption of LACCASE4 and 17 Results in Tissue-Specific Alterations to Lignification of *Arabidopsis thaliana* Stems." *The Plant Cell* 23 (3): 1124–37.

Besombes, S., Mazeau, K. 2005a. "The Cellulose/Lignin Assembly Assessed by Molecular Modeling. Part 1: Adsorption of a Threo Guaiacyl β-O-4 Dimer onto a Iβ Cellulose Whisker." *Plant Physiology and Biochemistry: PPB* 43 (3): 299–308.

Besombes, S., Mazeau, K. 2005b. "The Cellulose/Lignin Assembly Assessed by Molecular Modeling. Part 2: Seeking for Evidence of Organization of Lignin Molecules at the Interface with Cellulose." *Plant Physiology and Biochemistry: PPB* 43 (3): 277–86.

Bethke, G., Grundman, R. E., Sreekanta, S., Truman, W., Katagiri, F., Glazebrook, J. 2014. "Arabidopsis PECTIN METHYLESTERASEs Contribute to Immunity against *Pseudomonas syringae." Plant Physiology* 164 (2): 1093–1107.

Bhandari, S., Fujino, T., Thammanagowda, S., Zhang, D., Xu, F., Joshi, C. P. 2006. "Xylem-Specific and Tension Stress-Responsive Coexpression of KORRIGAN Endoglucanase and Three Secondary Wall-Associated Cellulose Synthase Genes in Aspen Trees." *Planta* 224 (4): 828–37.

Biswal, A. K., Atmodjo, M. A., Li, M., Baxter, H. L., Yoo, C. G., Pu, Y., Lee, Y-C., Mazarei, M., Black, I. M., Zhang, J. Y., Ramanna, H., Bray, A. L., King, Z. R., LaFayette, P. R., Pattathil, S., Donohoe, B. S., Mohanty, S. S., Ryno, D., Yee, K., Thompson, O. A., Rodriguez, M. Jr., Dumitrache, A., Natzke, J., Winkeler, K., Collins, C., Yang, X., Tan, L., Sykes, R. W., Gjersing, E. L., Ziebell, A., Turner, G. B., Decker, S. R., Hahn, M. G., Davison, B. H., Udvardi, M. K., Mielenz, J. R., Davis, M. F., Nelson, R. S., Parrott, W. A., Ragauskas, A. J., Neal Stewart C. Jr., Mohnen D. 2018. "Sugar Release and Growth of Biofuel Crops Are Improved by Downregulation of Pectin Biosynthesis." *Nature Biotechnology* 36 (3): 249–57.

Boija, E., Lundquist, A., Edwards, K., Johansson, G. 2007. "Evaluation of Bilayer Disks as Plant Cell Membrane Models in Partition Studies." *Analytical Biochemistry* 364 (2): 145–52.

Bonawitz, N. D., Kim, J. I., Tobimatsu, Y., Ciesielski, P. N., Anderson, N. A., Ximenes, E., Maeda, J., Ralph, J., Donohoe, B. S., Ladisch, M., Chapple, C. 2014. "Disruption of Mediator Rescues the Stunted Growth of a Lignin-Deficient Arabidopsis Mutant." *Nature* 509 (7500): 376–80.

Bouvier d'Yvoire, M., Bouchabke-Coussa, O., Voorend, W., Antelme, S., Cézard, L., Legée, F., Lebris, P., Legay, S., Whitehead, C., McQueen-Mason, S. J., Gomez, L. D., Jouanin, L., Lapierre. C/, Sibout. R. 2013. "Disrupting the Cinnamyl Alcohol Dehydrogenase 1 Gene (BdCAD1) Leads to Altered Lignification and Improved Saccharification in Brachypodium Distachyon." *The Plant Journal* 73 (3): 496–508.

Bradshaw, C. J. A., Warkentin, I. G., Sodhi, N. S. 2009. "Urgent Preservation of Boreal Carbon Stocks and Biodiversity." *Trends in Ecology & Evolution* 24 (10): 541–48.

Bragatto, J., Segato, F., Cota, J., Mello, D. B., Oliveira, M. M., Buckeridge, M. S., Squina, F. M., Driemeier, C. E. 2012. "Insights on How the Activity of an Endoglucanase Is Affected

by Physical Properties of Insoluble Celluloses." *The Journal of Physical Chemistry B* 116 (21): 6128–36.

Bromley, J. R., Busse-Wicher, M., Tryfona, T., Mortimer, J. C., Zhang, Z., Brown, D. M., Dupree, P. 2013. "GUX1 and GUX2 Glucuronyltransferases Decorate Distinct Domains of Glucuronoxylan with Different Substitution Patterns." *The Plant Journal* 74 (3): 423–34.

Brown, D. M., Goubet, F., Wong, V. W., Goodacre, R., Stephens, E., Dupree, P., Turner, S. R. 2007. "Comparison of Five Xylan Synthesis Mutants Reveals New Insight into the Mechanisms of Xylan Synthesis." *The Plant Journal* 52 (6): 1154–68.

Brown, D. M., Zhang, Z., Stephens, E., Dupree, P., Turner, S. R. 2009. "Characterization of IRX10 and IRX10-like Reveals an Essential Role in Glucuronoxylan Biosynthesis in Arabidopsis." *The Plant Journal* 57 (4): 732–46.

Burget, E. G., Verma, R., Mølhøj, M., Reiter, W. D. 2003. "The Biosynthesis of L-Arabinose in Plants: Molecular Cloning and Characterization of a Golgi-Localized UDP-D-Xylose 4-Epimerase Encoded by the MUR4 Gene of Arabidopsis." *THE PLANT CELL* 15 (2): 523–31.

Burton, R. A., Gidley, M. J., Fincher, G. B. 2010. "Heterogeneity in the Chemistry, Structure and Function of Plant Cell Walls." *Nature Chemical Biology* 6 (10): 724–32.

Burton, R. A., Wilson, S. M., Hrmova, M., Harvey, A. J., Shirley, N. J., Medhurst, A., Stone, N. A., Newbigin, E. J., Bacic, A., Fincher, G. B. 2006. "Cellulose Synthase-Like CslF Genes Mediate the Synthesis of Cell Wall (1,3;1,4)-ß-d-Glucans." *Science* 311 (5769): 1940–42.

Busse-Wicher, M. personal communication.

Busse-Wicher, M., Gomes, T. C. F., Tryfona, T., Nikolovski, N., Stott, K., Grantham, N. J., Bolam, D. M., Skaf, M. S., 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.

Busse-Wicher, M., Li, A., Silveira, R. L., Pereira, C. S., Tryfona, T., Gomes, T. C. F., Skaf, M. S., Dupree, P. 2016. "Evolution of Xylan Substitution Patterns in Gymnosperms and Angiosperms: Implications for Xylan Interaction with Cellulose." *Plant Physiology* 171 (4): 2418-31.

Busse-Wicher, M., Wicher, K. B., Kusche-Gullberg, M. 2014. "The Extostosin Family: Proteins with Many Functions." *Matrix Biology*, Proteoglycan Biology, 35 (April): 25–33.

Caffall, K. H., Pattathil, S., Phillips, S. E., Hahn, M. G., Mohnen, D. 2009. "*Arabidopsis thaliana* T-DNA Mutants Implicate GAUT Genes in the Biosynthesis of Pectin and Xylan in Cell Walls and Seed Testa." *Molecular Plant* 2 (5): 1000–1014.

Campbell, W. G., Bryant, S. A. 1941. "Determination of pH in Wood." *Nature* 147 (3725): 357.

Carmona, C., Langan, P., Smith, J. C., Petridis, L. 2014. "Why Genetic Modification of Lignin Leads to Low-Recalcitrance Biomass." *Physical Chemistry Chemical Physics* 17 (1): 358–64.

Cavalier, D. M., Lerouxel, O., Neumetzler, L., Yamauchi, K., Reinecke, A., Freshour, G., Zabotina, O. A., Hahn, M. G., Burgert, I., Pauly, M., Raikhel, N. V., Keegstra, K. 2008. "Disrupting Two *Arabidopsis thaliana* Xylosyltransferase Genes Results in Plants Deficient in Xyloglucan, a Major Primary Cell Wall Component." *The Plant Cell* 20 (6): 1519–37.

Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., Rance, M., Skelton, N. J. 2007. *Protein NMR Spectroscopy: Principles and Practice*. Second Edition. Elsevier Academic Press.

Chanliaud, E., Saulnier, L., Thibault, J. F. 1995. "Alkaline Extraction and Characterisation of Heteroxylans from Maize Bran." *Journal of Cereal Science* 21 (2): 195–203.

Chapple, C., Vogt, T., Ellis, B. E., Somerville, C. R. 1992. "An Arabidopsis Mutant Defective in the General Phenylpropanoid Pathway." *The Plant Cell* 4 (11): 1413–24.

Chen, F., Tobimatsu, Y., Havkin-Frenkel, D., Dixon, R. A., Ralph, J. 2012. "A Polymer of Caffeyl Alcohol in Plant Seeds." *Proceedings of the National Academy of Sciences* 109 (5): 1772–77.

Chiniquy, D., Sharma, V., Schultink, A., Baidoo, E. E., Rautengarten, C., Cheng, K., Carroll, A., Ulvskov, P., Harholt, J., Keasling, J. D., Pauly, M., Scheller, H. V., Ronald, P. C. 2012. "XAX1 from Glycosyltransferase Family 61 Mediates Xylosyltransfer to Rice Xylan." *Proceedings of the National Academy of Sciences* 109 (42): 17117–22.

Chiniquy, D., Varanasi, P., Oh, T., Harholt, J., Katnelson, J., Singh, S., Auer, M., Simmons, B., Adams, P. D., Scheller, H. V., Ronald, P. C. 2013. "Three Novel Rice Genes Closely Related to the Arabidopsis IRX9, IRX9L, and IRX14 Genes and Their Roles in Xylan Biosynthesis." *Frontiers in Plant Science* 4 (April): 83.

Chong, S-L., Derba-Maceluch, M., Koutaniemi, S., Gómez, L. D., McQueen-Mason, S. J., Tenkanen, M., Mellerowicz, E. J. 2015. "Active Fungal GH115 α-Glucuronidase Produced in *Arabidopsis thaliana* Affects Only the UX1-Reactive Glucuronate Decorations on Native Glucuronoxylans." *BMC Biotechnology* 15: 56.

Chou, Y-H., Pogorelko, G., Zabotina, O. A. 2012. "Xyloglucan Xylosyltransferases XXT1, XXT2, and XXT5 and the Glucan Synthase CSLC4 Form Golgi-Localized Multiprotein Complexes." *Plant Physiology* 159 (4): 1355–66.

Clough, S. J., Bent, A. F. 1998. "Floral Dip: A Simplified Method for Agrobacterium-Mediated Transformation of *Arabidopsis thaliana*." *The Plant Journal: For Cell and Molecular Biology* 16 (6): 735–43.

Cocuron, J-C., Lerouxel, O., Drakakaki, G., Alonso, A. P., Liepman, A. H., Keegstra, K., Raikhel, N., Wilkerson, C. G. 2007. "A Gene from the Cellulose Synthase-like C Family Encodes a β-1,4 Glucan Synthase." *Proceedings of the National Academy of Sciences* 104 (20): 8550–55.

Cosgrove, D. J. 2015. "Plant Expansins: Diversity and Interactions with Plant Cell Walls." *Current Opinion in Plant Biology* 25 (June): 162–72.

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

Coutinho, P. M., Deleury, E., Davies, G. J., Henrissat, B. 2003. "An Evolving Hierarchical Family Classification for Glycosyltransferases." *Journal of Molecular Biology* 328 (2): 307–17.

Crowe, J. D., Zarger, R. A., Hodge, D. B. 2017. "Relating Nanoscale Accessibility within Plant Cell Walls to Improved Enzyme Hydrolysis Yields in Corn Stover Subjected to Diverse Pretreatments." *Journal of Agricultural and Food Chemistry* 65 (39): 8652–62.

Cui, W., Liu, W., Wu, G. 1995. "A Simple Method for the Transformation of *Agrobacterium tumefaciens* by Foreign DNA." *Chinese Journal of Biotechnology* 11 (4): 267–74.

Culbertson, A. T., Ehrlich, J. J., Choe, J-Y., Honzatko, R. B., Zabotina, O. A. 2018. "Structure of Xyloglucan Xylosyltransferase 1 Reveals Simple Steric Rules That Define Biological Patterns of Xyloglucan Polymers." *Proceedings of the National Academy of Sciences* 115 (23): 6064–69.

Currie, H. A., Perry, C. C. 2006. "Resolution of Complex Monosaccharide Mixtures from Plant Cell Wall Isolates by High pH Anion Exchange Chromatography." *Journal of Chromatography A* 1128 (1): 90–96.

Das, N. N., Das, S. K., Mukherjee, A. K. 1984. "On the Ester Linkage between Lignin and 4-O-Methyl-D-Glucurono-D-Xylan in Jute Fiber (Corchorus Capsularis)." *Carbohydrate Research* 127 (2): 345–48.

Daud, M. J., Jarvis, M. C. 1992. "Mannan of Oil Palm Kernel." *Phytochemistry* 31 (2): 463–64.

Davis, J., Brandizzi, F., Liepman, A. H., Keegstra, K. 2010. "Arabidopsis Mannan Synthase CSLA9 and Glucan Synthase CSLC4 Have Opposite Orientations in the Golgi Membrane." *The Plant Journal* 64 (6): 1028–37.

De Caroli, M., Lenucci, M. S., Di Sansebastiano, G-P., Tunno, M., Montefusco, A., Dalessandro, G., Piro, G. 2014. "Cellular Localization and Biochemical Characterization of a Chimeric Fluorescent Protein Fusion of Arabidopsis Cellulose Synthase-Like A2 Inserted into Golgi Membrane." *ScientificWorldJournal*. 2014: 792420.

Del Lungo, A., Ball, J., Carle, J.,. 2006. "Global Planted Forests Thematic Study: Results and Analysis, Planted Forests and Trees Working Paper 38." Food and Agriculture Organization of the United Nations. http://www.fao.org/forestry/12139-03441d093f070ea7d7c4e3ec3f306507.pdf.

Delgado, I. J., Wang, Z., de Rocher, A., Keegstra, K., Raikhel, N. V. 1998. "Cloning and Characterization of AtRGP1. A Reversibly Autoglycosylated Arabidopsis Protein Implicated in Cell Wall Biosynthesis." *Plant Physiology* 116 (4): 1339–50.

Delporte, M., Bernard, G., Legrand, G., Hielscher, B., Lanoue, A., Molinié, R., Rambaud, C., Mathiron, D., Besseau, S., Linka, N., Hilbert, J. L., Gagneul, D. 2018. "A BAHD Neofunctionalization Promotes Tetrahydroxycinnamoyl Spermine Accumulation in the

Pollen Coat of the Asteraceae Family." *Journal of Experimental Botany* 69 (22): 5355–71.

Devaraj, R. D., Reddy, C. K., Xu, B. 2019. "Health-Promoting Effects of Konjac Glucomannan and Its Practical Applications: A Critical Review." *International Journal of Biological Macromolecules* 126 (April): 273–81.

Dhugga, K. S., Tiwari, S. C., Ray, P. M. 1997. "A Reversibly Glycosylated Polypeptide (RGP1) Possibly Involved in Plant Cell Wall Synthesis: Purification, Gene Cloning, and Trans-Golgi Localization." *Proceedings of the National Academy of Sciences* 94 (14): 7679–84.

Dick-Pérez, M., Zhang, Y., Hayes, J., Salazar, A., Zabotina, O. A., Hong, M. 2011. "Structure and Interactions of Plant Cell-Wall Polysaccharides by Two- and Three-Dimensional Magic-Angle-Spinning Solid-State NMR." *Biochemistry* 50 (6): 989–1000.

Dilokpimol, A., Poulsen, C. P., Vereb, G., Kaneko, S., Schulz, A., Geshi, N. 2014. "Galactosyltransferases from *Arabidopsis thaliana* in the Biosynthesis of Type II Arabinogalactan: Molecular Interaction Enhances Enzyme Activity." *BMC Plant Biology* 14 (1): 90.

Dimmel, D. R., MacKay, J. J., Althen, E. M., Parks, C., Sederoff, R. R. 2001. "Pulping and Bleaching of cad-Deficient Wood." *Journal of Wood Chemistry and Technology* 21 (1): 1–17.

Ding, S-Y., Liu, Y-S., Zeng, Y., Himmel, M. E., Baker, J. O., Bayer, E. A. 2012. "How Does Plant Cell Wall Nanoscale Architecture Correlate with Enzymatic Digestibility?" *Science* 338 (6110): 1055–60.

Ding, S-Y, Zhao, S., Zeng, Y. 2013. "Size, Shape, and Arrangement of Native Cellulose Fibrils in Maize Cell Walls." *Cellulose* 21 (2): 863–71.

Doblin, M. S., Pettolino, F. A., Wilson, S. M., Campbell, R., Burton, R. A., Fincher, G. B., Newbigin, E., Bacic, A. 2009. "A Barley Cellulose Synthase-like CSLH Gene Mediates (1,3;1,4)-Beta-D-Glucan Synthesis in Transgenic Arabidopsis." *Proceedings of the National Academy of Sciences* 106 (14): 5996–6001.

Dou, J., Kim, H., Li, Y., Padmakshan, D., Yue, F., Ralph, J., Vuorinen, T. 2018. "Structural Characterization of Lignins from Willow Bark and Wood." *Journal of Agricultural and Food Chemistry* 66 (28): 7294–7300.

Douchkov, D., Lueck, S., Hensel, G., Kumlehn, J., Rajaraman, J., Johrde, A., Doblin, M. S., Beahan, C. T., Kopischke, M., Fuchs, R., Lipka, V., Niks, R. E., Bulone, V., Chowdhury, J., Little, A., Burton, R. A., Bacic, A., Fincher, G. B., Schweizer, P. 2016. "The Barley (*Hordeum vulgare*) Cellulose Synthase-like D2 Gene (*HvCslD2*) Mediates Penetration Resistance to Host-Adapted and Nonhost Isolates of the Powdery Mildew Fungus." *New Phytologist* 212 (2): 421–33.

Driemeier, C. E. personal communication.

Driemeier, C. E., Mendes, F. M., Oliveira, M. M. 2012. "Dynamic Vapor Sorption and Thermoporometry to Probe Water in Celluloses." *Cellulose* 19 (4): 1051–63.

Dupree, R., Simmons, T. J., Mortimer, J. C., Patel, D., Iuga, D., Brown, S. P., Dupree, P. 2015. "Probing the Molecular Architecture of *Arabidopsis thaliana* Secondary Cell Walls Using Two- and Three-Dimensional ^{13}C Solid State Nuclear Magnetic Resonance Spectroscopy." *Biochemistry* 54 (14): 2335-45.

Earl, W. L., VanderHart, D. L. 1981. "Observations by High-Resolution Carbon-13 Nuclear Magnetic Resonance of Cellulose I Related to Morphology and Crystal Structure." *Macromolecules* 14 (3): 570–74.

Edwards, M. E., Marshall, E., Gidley, M. J., Grant Reid, J. S. 2002. "Transfer Specificity of Detergent-Solubilized Fenugreek Galactomannan Galactosyltransferase." *Plant Physiology* 129 (3): 1391–97.

Egelund, J., Obel, N., Ulvskov, P., Geshi, N., Pauly, M., Bacic, A., Petersen, B. L. 2007. "Molecular Characterization of Two *Arabidopsis thaliana* Glycosyltransferase Mutants, Rra1 and Rra2, Which Have a Reduced Residual Arabinose Content in a Polymer Tightly Associated with the Cellulosic Wall Residue." *Plant Molecular Biology* 64 (4): 439–51.

Egelund, J., Petersen, B. L., Motawia, M. S., Damager, I., Faik, A., Olsen, C. E., Ishii, T., Clausen, H., Ulvskov, P., Geshi, N. 2006. "*Arabidopsis thaliana* RGXT1 and RGXT2 Encode Golgi-Localized (1,3)-α-D-Xylosyltransferases Involved in the Synthesis of Pectic Rhamnogalacturonan-II." *The Plant Cell* 18 (10): 2593–2607.

Eklöf, J. M., Brumer, H. 2010. "The XTH Gene Family: An Update on Enzyme Structure, Function, and Phylogeny in Xyloglucan Remodeling." *Plant Physiology* 153 (2): 456–66.

d'Errico, C., Börjesson, J., Ding, H., Krogh, K. B. R. M., Spodsberg, N., Madsen, R., Monrad, R. N. 2016. "Improved Biomass Degradation Using Fungal Glucuronoyl—Esterases—Hydrolysis of Natural Corn Fiber Substrate." *Journal of Biotechnology* 219 (February): 117–23.

d'Errico, C., Jørgensen, J. O., Krogh, K. B. R. M., Spodsberg, N., Madsen, R., Monrad, R. N. 2015. "Enzymatic Degradation of Lignin-Carbohydrate Complexes (LCCs): Model Studies Using a Fungal Glucuronoyl Esterase from *Cerrena unicolor*." *Biotechnology and Bioengineering* 112 (5): 914–22.

Etiégni, L., Campbell, A. G. 1991. "Physical and Chemical Characteristics of Wood Ash." *Bioresource Technology* 37 (2): 173–78.

Eudes, A., George, A., Mukerjee, P., Kim, J. S., Pollet, B., Benke, P. I., Yang, F., Mitra, P., Sun, L., Cetinkol, O. P., Chabout, S., Mouille, G., Soubigou-Taconnat, L., Balzergue, S., Singh, S., Holmes, B. M., Mukhopadhyay, A., Keasling, J. D., Simmons, B. A., Lapierre, C., Ralph, J., Loqué, D. 2012. "Biosynthesis and Incorporation of Side-Chain-Truncated Lignin Monomers to Reduce Lignin Polymerization and Enhance Saccharification." *Plant Biotechnology Journal* 10 (5): 609–20.

Fan, H., Dong, H., Xu, C., Liu, J., Hu, B., Ye, J., Mai, G., Li, H. 2017. "Pectin Methylesterases Contribute the Pathogenic Differences between Races 1 and 4 of *Fusarium oxysporum* f. Sp. Cubense." *Scientific Reports* 7 (1): 13140.

Fauré, R., Courtin, C. M., Delcour, J. A., Dumon, C., Faulds, C. B., Fincher, G. B., Fort, S., Fry, S. C., Halila, S., Kabel, M. A., Pouvreau, L., Quemener, B., Rivet, A., Saulnier, L., Schols, H. A., Driguez, H., O'Donohue, M. J. 2009. "A Brief and Informationally Rich Naming System for Oligosaccharide Motifs of Heteroxylans Found in Plant Cell Walls." *Australian Journal of Chemistry* 62 (6): 533–37.

Feijão, C. d. V. G. P. 2016. "Characterisation of Grass Xylan Structure and Biosynthesis." University of Cambridge.

Fernandes, A. N., Thomas, L. H., Altaner, C. M., Callow, P., Forsyth, V. T., Apperley, D. C., Kennedy, C. J., Jarvis, M. C. 2011. "Nanostructure of Cellulose Microfibrils in Spruce Wood." *Proceedings of the National Academy of Sciences* 108 (47): E1195–1203.

Fernández-Pérez, F., Vivar, T., Pomar, F., Pedreño, M. A., Novo-Uzal, E. 2015. "Peroxidase 4 Is Involved in Syringyl Lignin Formation in *Arabidopsis thaliana*." *Journal of Plant Physiology* 175 (March): 86–94.

Fornalé, S., Capellades, M., Encina, A., Wang, K., Irar, S., Lapierre, C., Ruel, K., Joseleau, J. P., Berenguer, J., Puigdomènech, P., Rigau, J., Caparrós-Ruiz, D. 2012. "Altered Lignin Biosynthesis Improves Cellulosic Bioethanol Production in Transgenic Maize Plants Down-Regulated for Cinnamyl Alcohol Dehydrogenase." *Molecular Plant* 5 (4): 817–30.

Franke, R., Humphreys, J. M., Hemm, M. R., Denault, J. W., Ruegger, M. O., Cusumano, J. C., Chapple, C. 2002. "The Arabidopsis REF8 Gene Encodes the 3-Hydroxylase of Phenylpropanoid Metabolism." *The Plant Journal: For Cell and Molecular Biology* 30 (1): 33–45.

Freiesleben, P. v., Spodsberg, N., Stenbæk, A., Stålbrand, H., Krogh, K. B. R. M., Meyer, A. S. 2018. "Boosting of Enzymatic Softwood Saccharification by Fungal GH5 and GH26 Endomannanases." *Biotechnology for Biofuels* 11 (1): 194.

Froger, A., Hall, J. E. 2007. "Transformation of Plasmid DNA into E. Coli Using the Heat Shock Method." *Journal of Visualized Experiments : JoVE*, no. 6 (August).

Fu, C., Xiao, X., Xi, Y., Ge, Y., Chen, F., Bouton, J., Dixon, R. A., Wang, Z-Y. 2011. "Downregulation of Cinnamyl Alcohol Dehydrogenase (CAD) Leads to Improved Saccharification Efficiency in Switchgrass." *BioEnergy Research* 4 (3): 153–64.

Fukushima, R. S., Kerley, M. S., Ramos, M. H., Porter, J. H., Kallenbach, R. L. 2015. "Comparison of Acetyl Bromide Lignin with Acid Detergent Lignin and Klason Lignin and Correlation with in Vitro Forage Degradability." *Animal Feed Science and Technology* 201 (March): 25–37.

Geshi, N., Johansen, J. N., Dilokpimol, A., Rolland, A., Belcram, K., Verger, S., Kotake T, Tsumuraya, Y., Kaneko, S., Tryfona, T., Dupree, P., Scheller, H. V., Höfte, H., Mouille, G. 2013. "A Galactosyltransferase Acting on Arabinogalactan Protein Glycans Is Essential for Embryo Development in Arabidopsis." *The Plant Journal* 76 (1): 128–37.

Gille, S., Hänsel, U., Ziemann, M., Pauly, M. 2009. "Identification of Plant Cell Wall Mutants by Means of a Forward Chemical Genetic Approach Using Hydrolases." *Proceedings of the National Academy of Sciences* 106 (34): 14699–704.

Gille, S., de Souza, A., Xiong, G., Benz, M., Cheng, K., Schultink, A., Reca, I-B., Pauly, M. 2011. "O-Acetylation of Arabidopsis Hemicellulose Xyloglucan Requires AXY4 or AXY4L, Proteins with a TBL and DUF231 Domain." *The Plant Cell* 23 (11): 4041–53.

Giummarella, N., Pu, Y., Ragauskas, A. J., Lawoko, M. 2019. "A Critical Review on the Analysis of Lignin Carbohydrate Bonds." *Green Chemistry* 21 (7): 1573–95.

Gou, M., Ran, X., Martin, D. W., Liu, C-J. 2018. "The Scaffold Proteins of Lignin Biosynthetic Cytochrome P450 Enzymes." *Nature Plants* 4 (5): 299–310.

Goubet, F., Barton, C. J., Mortimer, J. C., Yu, X., Zhang, Z., Miles, G. P., Richens, J., Liepman, A. H., Seffen, K., Dupree, P. 2009. "Cell Wall Glucomannan in Arabidopsis Is Synthesised by CSLA Glycosyltransferases, and Influences the Progression of Embryogenesis." *The Plant Journal* 60 (3): 527–38.

Goubet, F., Misrahi, A., Park, S. K., Zhang, Z., Twell, D., Dupree, P. 2003. "AtCSLA7, a Cellulose Synthase-Like Putative Glycosyltransferase, Is Important for Pollen Tube Growth and Embryogenesis in Arabidopsis." *Plant Physiology* 131 (2): 547–57.

Grabber, J. H., Ralph, J., Hatfield, R. D. 1998. "Ferulate Cross-Links Limit the Enzymatic Degradation of Synthetically Lignified Primary Walls of Maize." *Journal of Agricultural and Food Chemistry* 46 (7): 2609–14.

Grantham, N. J., Wurman-Rodrich, J., Terrett, O. M., Lyczakowski, J. L., Stott, K., Iuga, D., Simmons, T. J., Durand-Tardif, M., Brown, S. P., Dupree, R., Busse-Wicher, M., Dupree, P. 2017. "An Even Pattern of Xylan Substitution Is Critical for Interaction with Cellulose in Plant Cell Walls." *Nature Plants* 3 (11): 859-65.

Grethlein, H. E. 1985. "The Effect of Pore Size Distribution on the Rate of Enzymatic Hydrolysis of Cellulosic Substrates." *Nature Biotechnology* 3 (2): 155–60.

Gu, Y., Somerville, C. R. 2010. "Cellulose Synthase Interacting Protein." *Plant Signaling & Behavior* 5 (12): 1571–74.

Guo, D., Chen, F., Inoue, K., Blount, J. W., Dixon, R. A. 2001. "Downregulation of Caffeic Acid 3-O-Methyltransferase and Caffeoyl CoA 3-O-Methyltransferase in Transgenic Alfalfa: Impacts on Lignin Structure and Implications for the Biosynthesis of G and S Lignin." *The Plant Cell* 13 (1): 73–88.

Hackney, J. M., Atalla, R. H., VanderHart, D. L. 1994. "Modification of Crystallinity and Crystalline Structure of *Acetobacter xylinum* Cellulose in the Presence of Water-Soluble β-1,4-Linked Polysaccharides: ^{13}C-NMR Evidence." *International Journal of Biological Macromolecules* 16 (4): 215–18.

Hannuksela, T., Hervé du Penhoat, C. 2004. "NMR Structural Determination of Dissolved O-Acetylated Galactoglucomannan Isolated from Spruce Thermomechanical Pulp." *Carbohydrate Research* 339 (2): 301–12.

Hannuksela, T., Tenkanen, M., Holmbom, B. 2002. "Sorption of Dissolved Galactoglucomannans and Galactomannans to Bleached Kraft Pulp." *Cellulose* 9 (3–4): 251–61.

Harholt, J., Jensen, J. K., Sørensen, S. O., Orfila, C., Pauly, M., Scheller, H. V. 2006. "ARABINAN DEFICIENT 1 Is a Putative Arabinosyltransferase Involved in Biosynthesis of Pectic Arabinan in Arabidopsis." *Plant Physiology* 140 (1): 49–58.

He, F. 2019. "Exploring the Function of Laccases from Miscanthus Sinensis in Lignin Biosynthesis." University of Heidelberg.

Herbaut, M., Zoghlami, A., Habrant, A., Falourd, X., Foucat, L., Chabbert, B., Paës, G. 2018. "Multimodal Analysis of Pretreated Biomass Species Highlights Generic Markers of Lignocellulose Recalcitrance." *Biotechnology for Biofuels* 11: 52.

Herbaut, M., Zoghlami, A., Paës, G. 2018. "Dynamical Assessment of Fluorescent Probes Mobility in Poplar Cell Walls Reveals Nanopores Govern Saccharification." *Biotechnology for Biofuels* 11: 271.

Hernández, F. L. 2018. "Identification of the Role of [Methyl]Glucuronic Acid on Arabinogalactan Polysaccharides in *Arabidopsis thaliana*." University of Cambridge.

Heux, L., Hägglund, P., Putaux, J-L., Chanzy, H. 2005. "Structural Aspects in Semicrystalline Samples of the Mannan II Family." *Biomacromolecules* 6 (1): 324–32.

195

Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, C., Meyer, D., Lapierre, C., Pollet, B., Legrand, M. 2004. "Silencing of Hydroxycinnamoyl-Coenzyme A Shikimate/Quinate Hydroxycinnamoyltransferase Affects Phenylpropanoid Biosynthesis." *The Plant Cell* 16 (6): 1446–65.

Holwerda, E. K., Worthen, R. S., Kothari, N., Lasky, R. C., Davison, B. H., Fu, C., Wang, Z. Y., Dixon, R. A., Biswal, A. K., Mohnen, D., Nelson, R. S., Baxter, H. L., Mazarei, M., Stewart, C. N. Jr., Muchero, W., Tuskan, G. A., Cai, C. M., Gjersing, E. E., Davis, M. F., Himmel, M. E., Wyman, C. E., Gilna, P., Lynd, L. R. 2019. "Multiple Levers for Overcoming the Recalcitrance of Lignocellulosic Biomass." *Biotechnology for Biofuels* 12: 15.

Houtman, C. J., Atalla, R. H. 1995. "Cellulose-Lignin Interactions (A Computational Study)." *Plant Physiology* 107 (3): 977–84.

Hrmova, M., Farkas, V., Lahnstein, J., Fincher, G. B. 2007. "A Barley Xyloglucan Xyloglucosyl Transferase Covalently Links Xyloglucan, Cellulosic Substrates, and (1,3;1,4)-β-D-Glucans." *Journal of Biological Chemistry* 282 (17): 12951–62.

Ishii, T. 1991. "Acetylation at O-2 of Arabinofuranose Residues in Feruloylated Arabinoxylan from Bamboo Shoot Cell-Walls." *Phytochemistry* 30 (7): 2317–20.

Ishizawa, C. I., Davis, M. F., Schell, D. F., Johnson, D. K. 2007. "Porosity and Its Effect on the Digestibility of Dilute Sulfuric Acid Pretreated Corn Stover." *Journal of Agricultural and Food Chemistry* 55 (7): 2575–81.

Iwai, H., Masaoka, N., Ishii, T., Satoh, S. 2002. "A Pectin Glucuronyltransferase Gene Is Essential for Intercellular Attachment in the Plant Meristem." *Proceedings of the National Academy of Sciences* 99 (25): 16319–24.

Jacquet, G., Pollet, B., Lapierre, C., Mhamdi, F., Rolando, C. 1995. "New Ether-Linked Ferulic Acid-Coniferyl Alcohol Dimers Identified in Grass Straws." *Journal of Agricultural and Food Chemistry* 43 (10): 2746–2751.

Jamet, E., Canut, H., Boudart, G., Pont-Lezica, R. F. 2006. "Cell Wall Proteins: A New Insight through Proteomics." *Trends in Plant Science* 11 (1): 33–39.

Jarvis, M. C. 2018. "Structure of Native Cellulose Microfibrils, the Starting Point for Nanocellulose Manufacture." *Phil. Trans. R. Soc. A* 376 (2112): 20170045.

Jarvis, M. C., Apperley, D. C. 1995. "Chain Conformation in Concentrated Pectic Gels: Evidence from ^{13}C NMR." *Carbohydrate Research* 275 (1): 131–45.

Jensen, J. K., Kim, H., Cocuron, J-C., Orler, R., Ralph, J., Wilkerson, C. G. 2011. "The DUF579 Domain Containing Proteins IRX15 and IRX15-L Affect Xylan Synthesis in Arabidopsis." *The Plant Journal* 66 (3): 387–400.

Jensen, J. K., Schultink, A., Keegstra, K., Wilkerson, C. G., Pauly, M. 2012. "RNA-Seq Analysis of Developing Nasturtium Seeds (*Tropaeolum majus*): Identification and Characterization of an Additional Galactosyltransferase Involved in Xyloglucan Biosynthesis." *Molecular Plant* 5 (5): 984–92.

Jensen, J. K., Johnson, N. R., Wilkerson, C. G. 2014. "*Arabidopsis thaliana* IRX10 and Two Related Proteins from Psyllium and *Physcomitrella patens* Are Xylan Xylosyltransferases." *The Plant Journal* 80 (2): 207–15.

Jensen, J. K., Sørensen, S. O., Harholt, J., Geshi, N., Sakuragi, Y., Møller, I., Zandleven, J., Bernal, A. J., Jensen, N. B., Sørensen, C., Pauly, M., Beldman, G., Willats, W. G., Scheller, H. V. 2008. "Identification of a Xylogalacturonan Xylosyltransferase Involved in Pectin Biosynthesis in Arabidopsis." *The Plant Cell* 20 (5): 1289–1302.

Jiang, N., Wiemels, R. E., Soya, A., Whitley, R., Held, M., Faik., A. 2016. "Composition, Assembly, and Trafficking of a Wheat Xylan Synthase Complex (XSC)." *Plant Physiology* 170 (4): 1999-2023.

Jing, B., Ishikawa, T., Soltis, N., Inada, N., Liang, Y., Murawska, G., Andeberhan, F., Pidatala, R., Yu, X., Baidoo, E., Kawai-Yamada, M., Loque, D., Kliebenstein, D. J., Dupree, P., Mortimer, J. C. 2018. "GONST2 Transports GDP-Mannose for Sphingolipid Glycosylation in the Golgi Apparatus of Arabidopsis." *BioRxiv*, June, 346775.

Jones, L. A., Ennos, R., Turner, S. R. 2001. "Cloning and Characterization of Irregular Xylem4 (irx4): A Severely Lignin-Deficient Mutant of Arabidopsis." *The Plant Journal* 26 (2): 205–16.

Jung, H. G., Phillips, R. L. 2010. "Putative Seedling Ferulate Ester (Sfe) Maize Mutant: Morphology, Biomass Yield, and Stover Cell Wall Composition and Rumen Degradability." *Crop Science; Madison* 50 (1): 403–18.

197

Kalmbach, L., Hématy, K., De Bellis, D., Barberon, M., Fujita, S., Ursache, R., Daraspe, J., Geldner, N. 2017. "Transient Cell-Specific EXO70A1 Activity in the CASP Domain and Casparian Strip Localization." *Nature Plants* 3 (5): 17058.

Kang, X., Kirui, A., Widanage, M. C. W., Mentink-Vigier, F., Cosgrove, D. J., Wang, T. 2019. "Lignin-Polysaccharide Interactions in Plant Secondary Cell Walls Revealed by Solid-State NMR." *Nature Communications* 10 (1): 347.

Kardošová, A., Malovíková, A., Rosík, J., Capek, P. 1990. "Distribution of 4-O-Methyl-D-Glucuronic Acid Units in 4-O-Methyl-D-Glucurono-D-Xylan Isolated from the Leaves of Marsh Mallow (*Althaea officinalis L., Var. Rhobusta*)." *Chemical Papers* 44 (1): 111–17.

Karlen, S. D., Free, H. C. A., Padmakshan, D., Smith, B. G., Ralph, J., Harris, P. J. 2018. "Commelinid Monocotyledon Lignins Are Acylated by P-Coumarate." *Plant Physiology* 177 (2): 513–21.

Keegstra, K. 2010. "Plant Cell Walls." *Plant Physiology* 154 (2): 483–86.

Kellett, L. E., Poole, D. M., Ferreira, L. M., Durrant, A. J., Hazlewood, G. P., Gilbert, H. J. 1990. "Xylanase B and an Arabinofuranosidase from *Pseudomonas fluorescens* Subsp. Cellulosa Contain Identical Cellulose-Binding Domains and Are Encoded by Adjacent Genes." *The Biochemical Journal* 272 (2): 369–76.

Keplinger, T., Konnerth, J., Aguié-Béghin, V., Rüggeberg, M., Gierlinger, N., Burgert, I. 2014. "A Zoom into the Nanoscale Texture of Secondary Cell Walls." *Plant Methods* 10 (1): 1.

Kim, H., Ralph, J. 2010. "Solution-State 2D NMR of Ball-Milled Plant Cell Wall Gels in DMSO-D6/Pyridine-D5." *Org. Biomol. Chem.* 8 (3): 576–91.

Kim, H., Ralph, J. 2014. "A Gel-State 2D-NMR Method for Plant Cell Wall Profiling and Analysis: A Model Study with the Amorphous Cellulose and Xylan from Ball-Milled Cotton Linters." *RSC Adv.* 4 (15): 7549–60.

Kim, H., Ralph, J., Lu, F., Ralph, S. A., Boudet, A. M., MacKay, J. J., Sederoff, R. R., Ito, T., Kawai, S., Ohashi, H., Higuchi, T. 2003. "NMR Analysis of Lignins in CAD-Deficient Plants. Part 1. Incorporation of Hydroxycinnamaldehydes and Hydroxybenzaldehydes into Lignins." *Organic & Biomolecular Chemistry* 1 (2): 268–81.

Kim, S-J, Held, M. A., Zemelis, S., Wilkerson, C. G., Brandizzi, F. 2015. "CGR2 and CGR3 Have Critical Overlapping Roles in Pectin Methylesterification and Plant Growth in *Arabidopsis thaliana*." *The Plant Journal* 82 (2): 208–20.

Kimura, S., Laosinchai, W., Itoh, T., Cui, X., Linder, C. R., Brown, R. M. 1999. "Immunogold Labeling of Rosette Terminal Cellulose-Synthesizing Complexes in the Vascular Plant *Vigna angularis*." *The Plant Cell* 11 (11): 2075–85.

Knoch, E., Dilokpimol, A., Tryfona, T., Poulsen, C. P., Xiong, G., Harholt, J., Petersen, B. L., Ulvskov, P., Hadi, M. Z., Kotake, T., Tsumuraya, Y., Pauly, M., Dupree, P., Geshi, N. 2013. "A β–Glucuronosyltransferase from *Arabidopsis thaliana* Involved in Biosynthesis of Type II Arabinogalactan Has a Role in Cell Elongation during Seedling Growth." *The Plant Journal* 76 (6): 1016–29.

Kolenová, K., Vrsanská, M., Biely, P. 2006. "Mode of Action of Endo-Beta-1,4-Xylanases of Families 10 and 11 on Acidic Xylooligosaccharides." *Journal of Biotechnology* 121 (3): 338–45.

Kong, Y., Zhou, G., Abdeen, A. A., Schafhauser, J., Richardson, B., Atmodjo, M. A., Jung, J., Wicker, L., Mohnen, D., Western, T., Hahn, M. G. 2013. "GALACTURONOSYLTRANSFERASE-LIKE5 Is Involved in the Production of Arabidopsis Seed Coat Mucilage." *Plant Physiology* 163 (3): 1203–17.

Konishi, T., Takeda, T., Miyazaki, Y., Ohnishi-Kameyama, M., Hayashi, T., O'Neill, M. A., Ishii, T. 2007. "A Plant Mutase That Interconverts UDP-Arabinofuranose and UDP-Arabinopyranose." *Glycobiology* 17 (3): 345–54.

Konishi, T., Miyazaki, Y., Yamakawa, S., Iwai, H., Satoh, S., Ishii, T. 2010. "Purification and Biochemical Characterization of Recombinant Rice UDP-Arabinopyranose Mutase Generated in Insect Cells." *Bioscience, Biotechnology, and Biochemistry* 74 (1): 191–94.

Kont, R., Kurašin, M., Teugjas, H., Väljamäe, P. 2013. "Strong Cellulase Inhibitors from the Hydrothermal Pretreatment of Wheat Straw." *Biotechnology for Biofuels* 6 (1): 135.

Kormelink, F.J.M., Voragen, A. G. J. 1993. "Degradation of Different [(Glucurono)Arabino]Xylans by a Combination of Purified Xylan-Degrading Enzymes." *Applied Microbiology and Biotechnology* 38 (5).

Kubicki, J. D., Yang, H., Sawada, D., O'Neill, H., Oehme, D., Cosgrove, D. J. 2018. "The Shape of Native Plant Cellulose Microfibrils." *Scientific Reports* 8 (1): 13983.

Kumar, M., Mishra, L., Carr, P., Pilling, M., Gardner, P., Mansfield, S. D., Turner, S. R. 2018. "Exploiting CELLULOSE SYNTHASE (CESA) Class Specificity to Probe Cellulose Microfibril Biosynthesis." *Plant Physiology* 177 (1): 151–67.

Kumar, M., Wightman, R., Atanassov, I., Gupta, A., Hurst, C. H., Hemsley, P. A., Turner, S. R. 2016. "S-Acylation of the Cellulose Synthase Complex Is Essential for Its Plasma Membrane Localization." *Science* 353 (6295): 166–69.

Lairson, L. L., Henrissat, B., Davies, G. J., Withers, S. G. 2008. "Glycosyltransferases: Structures, Functions, and Mechanisms." *Annual Review of Biochemistry* 77: 521–55.

Laitinen, T., Morreel, K., Delhomme, N., Gauthier, A., Schiffthaler, B., Nickolov, K., Brader, G., Lim, K. J., Teeri, T. H., Street, N. R., Boerjan, W., Kärkönen, A. 2017. "A Key Role for Apoplastic H2O2 in Norway Spruce Phenolic Metabolism." *Plant Physiology* 174 (3): 1449–75.

Lamport, D. T. A., Kieliszewski, M. J., Chen, Y., Cannon, M. C. 2011. "Role of the Extensin Superfamily in Primary Cell Wall Architecture." *Plant Physiology* 156 (1): 11–19.

Lamport, D. T. A., Várnai, P. 2013. "Periplasmic Arabinogalactan Glycoproteins Act as a Calcium Capacitor That Regulates Plant Growth and Development." *New Phytologist* 197 (1): 58–64.

Lan, W., Lu, F., Regner, M., Zhu, Y., Rencoret, J., Ralph, S. A., Zakai, U. I., Morreel, K., Boerjan, W., Ralph, J. 2015. "Tricin, a Flavonoid Monomer in Monocot Lignification." *Plant Physiology* 167 (4): 1284–95.

Lee, C., Teng, Q., Zhong, R., Ye, Z-H. 2011. "The Four Arabidopsis Reduced Wall Acetylation Genes Are Expressed in Secondary Wall-Containing Cells and Required for the Acetylation of Xylan." *Plant & Cell Physiology* 52 (8): 1289–1301.

Lee, Y., Rubio, M. C., Alassimone, J., Geldner, N. 2013. "A Mechanism for Localized Lignin Deposition in the Endodermis." *Cell* 153 (2): 402–12.

Lefebvre, V., Fortabat, M-N., Ducamp, A., North, H. M., Maia-Grondard, A., Trouverie, J., Boursiac, Y., Mouille, G., Durand-Tardif, M. 2011. "ESKIMO1 Disruption in Arabidopsis Alters Vascular Tissue and Impairs Water Transport." *PLoS ONE* 6 (2).

Lei, L., Li, S., Du, J., Bashline, L., Gu, Y. 2013. "CELLULOSE SYNTHASE INTERACTIVE3 Regulates Cellulose Biosynthesis in Both a Microtubule-Dependent and Microtubule-Independent Manner in Arabidopsis." *The Plant Cell* 25 (12): 4912–23.

Lei, L., Zhang, T., Strasser, R., Lee, C. M., Gonneau, M., Mach, L., Vernhettes, S., Kim, S. H., Cosgrove, D. J., Li, S., Gu, Y. 2014. "The Jiaoyao1 Mutant Is an Allele of Korrigan1 That Abolishes Endoglucanase Activity and Affects the Organization of Both Cellulose Microfibrils and Microtubules in Arabidopsis." *The Plant Cell* 26 (6): 2601–16.

Levesque-Tremblay, G., Müller, K., Mansfield, S. D., Haughn, G. W. 2015. "HIGHLY METHYL ESTERIFIED SEEDS Is a Pectin Methyl Esterase Involved in Embryo Development." *Plant Physiology* 167 (3): 725–37.

Levesque-Tremblay, G., Pelloux, J., Braybrook, S. A., Müller, K. 2015. "Tuning of Pectin Methylesterification: Consequences for Cell Wall Biomechanics and Development." *Planta* 242 (4): 791–811.

Lewis, N. S. 2016. "Research Opportunities to Advance Solar Energy Utilization." *Science* 351 (6271)

Li, X., Jackson, P., Rubtsov, D. V., Faria-Blanc, N., Mortimer, J. C., Turner, S. R., Krogh, K. B. R. M., Johansen, K. S., Dupree, P. 2013. "Development and Application of a High Throughput Carbohydrate Profiling Technique for Analyzing Plant Cell Wall Polysaccharides and Carbohydrate Active Enzymes." *Biotechnology for Biofuels* 6 (1): 94.

Li, Y., Kim, J. I., Pysh, L., Chapple, C. 2015. "Four Isoforms of Arabidopsis 4-Coumarate:CoA Ligase Have Overlapping yet Distinct Roles in Phenylpropanoid Metabolism." *Plant Physiology* 169 (4): 2409–21.

Liang, Y., Basu, D., Pattathil, S., Xu, W-l., Venetos, A., Martin, S. L., Faik, A., Hahn, M. G., Showalter, A. M. 2013. "Biochemical and Physiological Characterization of Fut4 and Fut6 Mutants Defective in Arabinogalactan-Protein Fucosylation in Arabidopsis." *Journal of Experimental Botany* 64 (18): 5537–51.

Liepman, A. H., Wilkerson, C. G., Keegstra, K. 2005. "Expression of Cellulose Synthase-like (Csl) Genes in Insect Cells Reveals That CslA Family Members Encode Mannan Synthases." *Proceedings of the National Academy of Sciences* 102 (6): 2221–26.

Linder, H. P., Lehmann, C. E. R., Archibald, S., Osborne, C. P., Richardson, D. M. 2018. "Global Grass (Poaceae) Success Underpinned by Traits Facilitating Colonization, Persistence and Habitat Transformation." *Biological Reviews* 93 (2): 1125–44.

Little, A., Lahnstein, J., Jeffery, D. W., Khor, S. F., Schwerdt, J. G., Shirley, N. J., Hooi, M., Xing, X., Burton, R. A., Bulone, V. 2019. "A Novel (1,4)-β-Linked Glucoxylan Is Synthesized by Members of the Cellulose Synthase-Like F Gene Family in Land Plants." *ACS Central Science* 5 (1): 73–84.

Little, A., Schwerdt, J. G., Shirley, N. J., Khor, S. F., Neumann, K., O'Donovan, L. A., Jelle Lahnstein, Collins, H. M., Henderson, M., Fincher, G. B., Burton, R. A. 2018. "Revised Phylogeny of the Cellulose Synthase Gene Superfamily: Insights into Cell Wall Evolution." *Plant Physiology* 177 (3): 1124–41.

Liu, X.-L., Liu, L., Niu, Q-K., Xia, C., Yang, K-Z., Li, R., Chen, L-Q., Zhang, X-Q., Zhou, Y., Ye, D. 2011. "Male Gametophyte Defective 4 Encodes a Rhamnogalacturonan II Xylosyltransferase and Is Important for Growth of Pollen Tubes and Roots in Arabidopsis." *The Plant Journal: For Cell and Molecular Biology* 65 (4): 647–60.

Liwanag, A. J. M., Ebert, B., Verhertbruggen, Y., Rennie, E. A., Rautengarten, C., Oikawa, A., Andersen, M. C. F., Clausen, M. H., Scheller, H, V. 2012. "Pectin Biosynthesis: GALS1 in *Arabidopsis thaliana* Is a β-1,4-Galactan β-1,4-Galactosyltransferase." *The Plant Cell* 24 (12): 5024–36.

Lucas, W J., Groover, A., Lichtenberger, R., Furuta, K., Yadav, S-R., Helariutta, Y., He, X. Q., Fukuda, H., Kang, J., Brady, S. M., Patrick, J. W., Sperry, J., Yoshida, A., López-Millán, A. F., Grusak, M. A., Kachroo, P. 2013. "The Plant Vascular System: Evolution, Development and Functions." *Journal of Integrative Plant Biology* 55 (4): 294–388.

Lyczakowski, J. J., Wicher, K. B., Terrett, O. M., Faria-Blanc, N., Yu, X., Brown, D., Krogh, K. B. R. M., Dupree, P., 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: 224.

Łyczakowski, J. J. 2018. "Biosynthesis and Function of Glucuronic Acid Substitution Patterns on Softwood Xylan." University of Cambridge.

Madson, M., Dunand, C., Li, X., Verma, R., Vanzin, G. F., Caplan, J., Shoue, D. A., Carpita, N. C., Reiter, W-D. 2003. "The MUR3 Gene of Arabidopsis Encodes a Xyloglucan Galactosyltransferase That Is Evolutionarily Related to Animal Exostosins." *The Plant Cell* 15 (7): 1662–70.

Mahon, E. L., Mansfield, S. D. 2019. "Tailor-Made Trees: Engineering Lignin for Ease of Processing and Tomorrow's Bioeconomy." *Current Opinion in Biotechnology* 56 (April): 147–55.

Manabe, Y., Verhertbruggen, Y., Gille, S., Harholt, J., Chong, S-L., Pawar, P. M-A., Mellerowicz, E. J., Tenkanen, M., Cheng, K., Pauly, M., Scheller, H. V. 2013. "Reduced Wall Acetylation Proteins Play Vital and Distinct Roles in Cell Wall O-Acetylation in Arabidopsis." *Plant Physiology* 163 (3): 1107–17.

Mansfield, S. D., Kang, K-Y., Chapple, C. 2012. "Designed for Deconstruction – Poplar Trees Altered in Cell Wall Lignification Improve the Efficacy of Bioethanol Production." *New Phytologist* 194 (1): 91–101.

Mares, D.J., Stone, B. A. 1973. "Studies on Wheat Endosperm I. Chemical Composition and Ultrastructure of the Cell Walls." *Australian Journal of Biological Sciences* 26 (4): 793.

Marriott, P. E., Sibout, R., Lapierre, C., Fangel, J. U., Willats, W. G. T., Hofte, H., Gómez, L. D., McQueen-Mason, S. J. 2014. "Range of Cell-Wall Alterations Enhance Saccharification in *Brachypodium distachyon* Mutants." *Proceedings of the National Academy of Sciences* 111 (40): 14601–6.

Martínez, P. M., Appeldoorn, M. M., Gruppen, H., Kabel, M. A. 2016. "The Two *Rasamsonia emersonii* α-Glucuronidases, ReGH67 and ReGH115, Show a Different Mode-of-Action towards Glucuronoxylan and Glucuronoxylo-Oligosaccharides." *Biotechnology for Biofuels* 9: 105.

Martínez-Abad, A., Berglund, J., Toriz, G., Gatenholm, P., Henriksson, G., Lindström, M., Wohlert, J., Vilaplana, F. 2017. "Regular Motifs in Xylan Modulate Molecular Flexibility and Interactions with Cellulose Surfaces." *Plant Physiology* 175 (4): 1579–92.

Martínez-Abad, A., Giummarella, N., Lawoko, M., Vilaplana, F. 2018. "Differences in Extractability under Subcritical Water Reveal Interconnected Hemicellulose and Lignin Recalcitrance in Birch Hardwoods." *Green Chemistry* 20 (11): 2534–46.

McCann, M. C., Carpita, N. C. 2015. "Biomass Recalcitrance: A Multi-Scale, Multi-Factor and Conversion-Specific Property." *Journal of Experimental Botany* 66 (14): 4109-18.

McKee, L. S., Sunner, H., Anasontzis, G. E., Toriz, G., Gatenholm, P., Bulone, V., Vilaplana, F., Olsson, L. 2016. "A GH115 α-Glucuronidase from *Schizophyllum commune* Contributes to the Synergistic Enzymatic Deconstruction of Softwood Glucuronoarabinoxylan." *Biotechnology for Biofuels* 9: 2.

Meester, B. De., de Vries, L., Özparpucu, M., Gierlinger, N., Corneillie, S., Pallidis, A., Goeminne, G., Morreel, K., De Bruyne, M., De Rycke, R., Vanholme, R., Boerjan, W. 2018. "Vessel-Specific Reintroduction of CINNAMOYL-COA REDUCTASE1 (CCR1) in Dwarfed Ccr1 Mutants Restores Vessel and Xylary Fiber Integrity and Increases Biomass." *Plant Physiology* 176 (1): 611–33.

Meng, X., Ragauskas, A. J. 2014. "Recent Advances in Understanding the Role of Cellulose Accessibility in Enzymatic Hydrolysis of Lignocellulosic Substrates." *Current Opinion in Biotechnology* 27: 150–58.

Meyer, K., Shirley, A. M., Cusumano, J. C., Bell-Lelong, D. A., Chapple, C. 1998. "Lignin Monomer Composition Is Determined by the Expression of a Cytochrome P450-Dependent Monooxygenase in Arabidopsis." *Proceedings of the National Academy of Sciences* 95 (12): 6619–23.

Miao, Y-C., Liu, C-J. 2010. "ATP-Binding Cassette-like Transporters Are Involved in the Transport of Lignin Precursors across Plasma and Vacuolar Membranes." *Proceedings of the National Academy of Sciences* 107 (52): 22728–33.

Micheli, F. 2001. "Pectin Methylesterases: Cell Wall Enzymes with Important Roles in Plant Physiology." *Trends in Plant Science* 6 (9): 414–19.

Mitchell, R. A. C., Dupree, P., Shewry, P. R. 2007. "A Novel Bioinformatics Approach Identifies Candidate Genes for the Synthesis and Feruloylation of Arabinoxylan." *PLANT PHYSIOLOGY* 144 (1): 43–53.

Moglia, A., Acquadro, A., Eljounaidi, K., Milani, A. M., Cagliero, C., Rubiolo, P., Genre, A., Cankar, K., Beekwilder, J., Comino, C. 2016. "Genome-Wide Identification of BAHD Acyltransferases and In Vivo Characterization of HQT-like Enzymes Involved in Caffeoylquinic Acid Synthesis in Globe Artichoke." *Frontiers in Plant Science* 7: 1424.

Mohnen, D. 2008. "Pectin Structure and Biosynthesis." *Current Opinion in Plant Biology*, 11 (3): 266–77.

Møller, S. R., Yi, X., Velásquez, S. M., Gille, S., Hansen, P. L. M., Poulsen, C. P., Olsen, C. E., Rejzek, M., Parsons, H., Yang, Z., Wandall, H. H., Clausen, H., Field, R. A., Pauly, M., Estevez, J. M., Harholt, J., Ulvskov, P., Petersen, B. L. 2017. "Identification and Evolution of a Plant Cell Wall Specific Glycoprotein Glycosyl Transferase, ExAD." *Scientific Reports* 7 (March): 45341.

Moreira-Vilar, F. C., Siqueira-Soares, R. d. C., Finger-Teixeira, A., Oliveira, D. M. d., Ferro, A. P., da Rocha, G. J., Ferrarese, M. d. L. L., dos Santos, W. D., Ferrarese-Filho, O. 2014. "The Acetyl Bromide Method Is Faster, Simpler and Presents Best Recovery of Lignin in Different Herbaceous Tissues than Klason and Thioglycolic Acid Methods." *PLOS ONE* 9 (10): e110000.

Morgan, J. L. W., McNamara, J. T., Fischer, M., Rich, J., Chen, H-M., Withers, S. G., Zimmer, J. 2016. "Observing Cellulose Biosynthesis and Membrane Translocation in Crystallo." *Nature* 531 (7594): 329–34.

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., 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): 17409–14.

Mortimer, J. C., Faria-Blanc, N., Yu, X., Tryfona, T., Sorieul, M., Ng, Y. Z., Zhang, Z., Stott, K., Anders, N., Dupree, P. 2015. "An Unusual Xylan in Arabidopsis Primary Cell Walls Is Synthesised by GUX3, IRX9L, IRX10L and IRX14." *The Plant Journal* 83 (3): 413–26.

Mortimer, J. C., Yu, X., Albrecht, S., Sicilia, F., Huichalaf, M., Ampuero, D., Michaelson, L. V., Murphy, A. M., Matsunaga, T., Kurz, S., Stephens, E., Baldwin, T. C., Ishii, T., Napier, J. A., Weber, A. P., Handford, M. G., Dupree, P. 2013. "Abnormal Glycosphingolipid Mannosylation Triggers Salicylic Acid–Mediated Responses in Arabidopsis." *The Plant Cell* 25 (5): 1881–94.

Mosbech, C., Holck, J., Meyer, A. S., Agger, J. W. 2018. "The Natural Catalytic Function of CuGE Glucuronoyl Esterase in Hydrolysis of Genuine Lignin–Carbohydrate Complexes from Birch." *Biotechnology for Biofuels* 11: 1.

Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J., Mansfield, S. D. 2016. "Designer Lignins: Harnessing the Plasticity of Lignification." *Current Opinion in Biotechnology* 37: 190–200.

Mueller, S. C. 1980. "Evidence for an Intramembrane Component Associated with a Cellulose Microfibril-Synthesizing Complex in Higher Plants." *The Journal of Cell Biology* 84 (2): 315–26.

Ndeh, D., Gilbert, H. J. 2018. "Biochemistry of Complex Glycan Depolymerisation by the Human Gut Microbiota." *FEMS Microbiology Reviews* 42 (2): 146–64.

Ndeh, D., Rogowski, A., Cartmell, A., Luis, A. S., Baslé, A., Gray, J., Venditto, I., Briggs, J., Zhang, X., Labourel, A., Terrapon, N., Buffetto, F., Nepogodiev, S., Xiao, Y., Field, R. A., Zhu, Y., O'Neil, M. A., Urbanowicz, B. R., York, W. S., Davies, G. J., Abbott, D. W., Ralet, M. C., Martens, E. C., Henrissat, B., Gilbert, H. J. 2017. "Complex Pectin Metabolism by Gut Bacteria Reveals Novel Catalytic Functions." *Nature* 544 (7648): 65–70.

Neumüller, K. G., de Souza, A. C., van Rijn, J. H. J., Streekstra, H., Gruppen, H., Schols, H. A. 2015. "Positional Preferences of Acetyl Esterases from Different CE Families towards Acetylated 4-O-Methyl Glucuronic Acid-Substituted Xylo-Oligosaccharides." *Biotechnology for Biofuels* 8: 7.

Nishikubo, N., Awano, T., Banasiak, A., Bourquin, V., Ibatullin, F., Funada, R., Brumer, H., Teeri, T. T., Hayashi, T., Sundberg, B., Mellerowicz, E. J. 2007. "Xyloglucan Endo-Transglycosylase (XET) Functions in Gelatinous Layers of Tension Wood Fibers in Poplar—A Glimpse into the Mechanism of the Balancing Act of Trees." *Plant and Cell Physiology* 48 (6): 843–55.

Nishimura, H., Kamiya, A., Nagata, T., Katahira, M., Watanabe, T. 2018. "Direct Evidence for α Ether Linkage between Lignin and Carbohydrates in Wood Cell Walls." *Scientific Reports* 8 (1).

Nishiyama, Y., Sugiyama, J., Chanzy, H., Langan, P. 2003. "Crystal Structure and Hydrogen Bonding System in Cellulose Iα from Synchrotron X-Ray and Neutron Fiber Diffraction." *Journal of the American Chemical Society* 125 (47): 14300–306.

Novozymes. 2017. "Cellic CTec3 HS Application Sheet." Novozymes. https://www.novozymes.com › Cellic-CTec3-HS-application-sheet-NA.

O. Buanafina, M. M. de. 2009. "Feruloylation in Grasses: Current and Future Perspectives." *Molecular Plant* 2 (5): 861–72.

Obayashi, T., Aoki, Y., Tadaka, S., Kagaya, Y., Kinoshita, K. 2018. "ATTED-II in 2018: A Plant Coexpression Database Based on Investigation of the Statistical Property of the Mutual Rank Index." *Plant & Cell Physiology* 59 (1): e3.

Obel, N., Porchia, A. C., Scheller, H. V. 2003. "Intracellular Feruloylation of Arabinoxylan in Wheat: Evidence for Feruloyl-Glucose as Precursor." *Planta* 216 (4): 620–29.

Ogawa-Ohnishi, M., Matsushita, W., Matsubayashi, Y. 2013. "Identification of Three Hydroxyproline O-Arabinosyltransferases in *Arabidopsis thaliana*." *Nature Chemical Biology* 9 (11): 726–30.

O'Neill, M. A., Eberhard, S., Albersheim, P., Darvill, A. G. 2001. "Requirement of Borate Cross-Linking of Cell Wall Rhamnogalacturonan II for Arabidopsis Growth." *Science* 294 (5543): 846–49.

Onkokesung, N., Gaquerel, E., Kotkar, H., Kaur, H., Baldwin, I. T., Galis, I. 2012. "MYB8 Controls Inducible Phenolamide Levels by Activating Three Novel Hydroxycinnamoyl-Coenzyme A:Polyamine Transferases in *Nicotiana attenuata*." *Plant Physiology* 158 (1): 389–407.

Østergaard, L., Teilum, K., Mirza, O., Mattsson, O., Petersen, M., Welinder, K. G., Mundy, J., Gajhede, M., Henriksen, A. 2000. "Arabidopsis ATP A2 Peroxidase. Expression and High-Resolution Structure of a Plant Peroxidase with Implications for Lignification." *Plant Molecular Biology* 44 (2): 231–43.

Park, S., Szumlanski, A. L., Gu, F., Guo, F., Nielsen, E. 2011. "A Role for CSLD3 during Cell-Wall Synthesis in Apical Plasma Membranes of Tip-Growing Root-Hair Cells." *Nature Cell Biology* 13 (8): 973–80.

Park, Y. B., Cosgrove, D. J. 2012. "Changes in Cell Wall Biomechanical Properties in the Xyloglucan-Deficient Xxt1/Xxt2 Mutant of Arabidopsis." *Plant Physiology* 158 (1): 465–75.

Patron, N. J., Orzaez, D., Marillonnet, S., Warzecha, H., Matthewman, C., Youles, M., Raitskin, O., Leveau, A., Farré, G., Rogers, C., Smith, A., Hibberd, J., Webb, A. A., Locke, J., Schornack, S., Ajioka, J., Baulcombe, D. C., Zipfel, C., Kamoun, S., Jones, J. D., Kuhn, H., Robatzek, S., Van Esse, H. P., Sanders, D., Oldroyd, G., Martin, C., Field, R., O'Connor, S., Fox, S., Wulff, B., Miller, B., Breakspear, A., Radhakrishnan, G., Delaux, P. M., Loqué, D., Granell, A., Tissier, A., Shih, P., Brutnell, T. P., Quick, W. P., Rischer, H., Fraser, P. D., Aharoni, A., Raines, C., South, P. F., Ané, J. M., Hamberger, B. R., Langdale, J., Stougaard, J., Bouwmeester, H., Udvardi, M., Murray, J. A., Ntoukakis, V., Schäfer, P., Denby, K., Edwards, K. J., Osbourn, A., Haseloff, J. 2015. "Standards for Plant Synthetic Biology: A Common Syntax for Exchange of DNA Parts." *New Phytologist* 208 (1): 13–19.

Pauly, M., Keegstra, K. 2008. "Cell-Wall Carbohydrates and Their Modification as a Resource for Biofuels." *The Plant Journal* 54 (4): 559–68.

Pauly, M., Keegstra, K. 2016. "Biosynthesis of the Plant Cell Wall Matrix Polysaccharide Xyloglucan." *Annual Review of Plant Biology* 67 (1): 235–59.

Peña, M. J., Kulkarni, A. R., Backe, J., Boyd, M., O'Neill, M. A., York, W. S. 2016. "Structural Diversity of Xylans in the Cell Walls of Monocots." *Planta* 244 (3): 589–606.

Peña, M. J., Zhong, R., Zhou, G-K., Richardson, E. A., O'Neill, M. A., Darvill, A. G., York, W. S., Ye, Z-H. 2007. "Arabidopsis Irregular Xylem8 and Irregular Xylem9: Implications for the Complexity of Glucuronoxylan Biosynthesis." *The Plant Cell* 19 (2): 549–63.

Perkins, M., Smith, R. A., Samuels, L. 2019. "The Transport of Monomers during Lignification in Plants: Anything Goes but How?" *Current Opinion in Biotechnology* 56: 69–74.

Perrin, R. M., DeRocher, A. E., Bar-Peled, M., Zeng, W., Norambuena, L., Orellana, A., Raikhel, N. V., Keegstra, K. 1999. "Xyloglucan Fucosyltransferase, an Enzyme Involved in Plant Cell Wall Biosynthesis." *Science* 284 (5422): 1976–79.

Phyo, P., Wang, T., Kiemle, S. N., O'Neill, H., Pingali, S. V., Hong, M., Cosgrove, D. J. 2017. "Gradients in Wall Mechanics and Polysaccharides along Growing Inflorescence Stems." *Plant Physiology* 175 (4): 1593–1607.

Phyo, P., Wang, T., Yang, Y., O'Neill, H., Hong, M. 2018. "Direct Determination of Hydroxymethyl Conformations of Plant Cell Wall Cellulose Using [1]H Polarization Transfer Solid-State NMR." *Biomacromolecules* 19 (5): 1485–97.

Pihlajaniemi, V., Sipponen, M. H., Kallioinen, A., Nyyssölä, A., Laakso, S. 2016. "Rate-Constraining Changes in Surface Properties, Porosity and Hydrolysis Kinetics of Lignocellulose in the Course of Enzymatic Saccharification." *Biotechnology for Biofuels* 9: 18.

Pihlajaniemi, V., Sipponen, M, H., Liimatainen, H., Sirviö, J. A., Nyyssölä, A., Laakso, S. 2016. "Weighing the Factors behind Enzymatic Hydrolyzability of Pretreated Lignocellulose." *Green Chemistry* 18 (5): 1295–1305.

Piston, F., Uauy, C., Fu, L., Langston, J., Labavitch, J., Dubcovsky, J. 2010. "Down-Regulation of Four Putative Arabinoxylan Feruloyl Transferase Genes from Family PF02458 Reduces Ester-Linked Ferulate Content in Rice Cell Walls." *Planta* 231 (3): 677–91.

Qin, L., Li, W-C., Liu, L., Zhu, J-Q., Li, X., Li, B-Z., Yuan, Y-J. 2016. "Inhibition of Lignin-Derived Phenolic Compounds to Cellulase." *Biotechnology for Biofuels* 9 (1): 70.

Qing, Q., Wyman, C. E. 2011. "Supplementation with Xylanase and β-Xylosidase to Reduce Xylo-Oligomer and Xylan Inhibition of Enzymatic Hydrolysis of Cellulose and Pretreated Corn Stover." *Biotechnology for Biofuels* 4 (1): 18.

Qu, Y., Egelund, J., Gilson, P. R., Houghton, F., Gleeson, P. A., Schultz, C. J., Bacic, A. 2008. "Identification of a Novel Group of Putative *Arabidopsis thaliana* β-(1,3)-Galactosyltransferases." *Plant Molecular Biology* 68 (1): 43–59.

Quinlan, R. J., Sweeney, M. D., Leggio, L. L., Otten, H., Poulsen, J-C. N., Johansen, K. S., Krogh, K. B. R. M., Jørgensen, C. I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C. P., Dupree, P., Xu, F., Davies, G. J., Walton, P. H. 2011. "Insights into the Oxidative Degradation of Cellulose by a Copper Metalloenzyme That Exploits Biomass Components." *Proceedings of the National Academy of Sciences* 108 (37): 15079–84.

Ralph, J., Akiyama, T., Kim, H., Lu, F., Schatz, P. F., Marita, J. M., Ralph, S. A., Reddy, M. S., Chen, F., Dixon, R. A. 2006. "Effects of Coumarate 3-Hydroxylase Down-Regulation on Lignin Structure." *Journal of Biological Chemistry* 281 (13): 8843–53.

Ralph, J., Grabber, J. H., Hatfield, R. D. 1995. "Lignin-Ferulate Cross-Links in Grasses: Active Incorporation of Ferulate Polysaccharide Esters into Ryegrass Lignins." *Carbohydrate Research* 275 (1): 167–78.

Ralph, J., Lapierre, C., Boerjan, W. 2019. "Lignin Structure and Its Engineering." *Current Opinion in Biotechnology* 56: 240–49.

Ralph, J., Lapierre, C., Marita, J. M., Kim, H., Lu, F., Hatfield, R. D., Ralph, S., Chapple, C., Franke, R., Hemm, M. R., Van Doorsselaere, J., Sederoff, R. R., O'Malley, D. M., Scott, J. T., MacKay, J. J., Yahiaoui, N., Boudet, A., Pean, M., Pilate, G., Jouanin, L., Boerjan, W. 2001. "Elucidation of New Structures in Lignins of CAD-and COMT-Deficient Plants by NMR." *Phytochemistry* 57 (6): 993–1003.

Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P. F., Marita, J. M., Hatfield, R. D., Ralph, S. A., Christensen, J. H., Boerjan, W. 2004. "Lignins: Natural Polymers from Oxidative Coupling of 4-Hydroxyphenyl- Propanoids." *Phytochemistry Reviews* 3 (1–2): 29–60.

Ralph, J., Quideau, S., Grabber, J. H., Hatfield, R. D. 1994. "Identification and Synthesis of New Ferulic Acid Dehydrodimers Present in Grass Cell Walls." *Journal of the Chemical Society, Perkin Transactions 1*, no. 23: 3485–3498.

Ramírez, V., Xiong, G., Mashiguchi, K., Yamaguchi, S., Pauly, M. 2018. "Growth- and Stress-Related Defects Associated with Wall Hypoacetylation Are Strigolactone-Dependent." *Plant Direct* 2 (6): e00062.

Rantanen, H., Virkki, L., Tuomainen, P., Kabel, M., Schols, H. A., Tenkanen, M. 2007. "Preparation of Arabinoxylobiose from Rye Xylan Using Family 10 Aspergillus Aculeatus Endo-1,4-β-d-Xylanase." *Carbohydrate Polymers* 68 (2): 350–59.

Rautengarten, C., Birdseye, D., Pattathil, S., McFarlane, H. E., Saez-Aguayo, S., Orellana, A., Persson, S., Hahn, M. G., Scheller, H. V., Heazlewood, J. L., Ebert, B. 2017. "The Elaborate Route for UDP-Arabinose Delivery into the Golgi of Plants." *Proceedings of the National Academy of Sciences* 114 (16): 4261–66.

Reiter, W. D. 2002. "Biosynthesis and Properties of the Plant Cell Wall." *Current Opinion in Plant Biology* 5 (6): 536–42.

Ren, Y., Hansen, S. F., Ebert, B., Lau, J., Scheller, H. V. 2014. "Site-Directed Mutagenesis of IRX9, IRX9L and IRX14 Proteins Involved in Xylan Biosynthesis: Glycosyltransferase Activity Is Not Required for IRX9 Function in Arabidopsis." *PLOS ONE* 9 (8): e105014.

Río, J. C. del., Rencoret, J., Gutiérrez, A., Kim, H., Ralph, J. 2018. "Structural Characterization of Lignin from Maize (*Zea mays L.*) Fibers: Evidence for Diferuloylputrescine Incorporated into the Lignin Polymer in Maize Kernels." *Journal of Agricultural and Food Chemistry* 66 (17): 4402–13.

Río, J. C. del., Rencoret, J., Gutiérrez, A., Kim, H., Ralph, J. 2017. "Hydroxystilbenes Are Monomers in Palm Fruit Endocarp Lignins." *Plant Physiology* 174 (4): 2072–82.

Roberts, A. W., Lahnstein, J., Hsieh, Y. S. Y., Xing, X., Yap, K., Chaves, A. M., Scavuzzo-Duggan, T. R., Dimitroff, G., Lonsdale, A., Roberts, E., Bulone, V., Fincher, G. B., Doblin, M. S., Bacic, A., Burton, R. A. 2018. "Functional Characterization of a Glycosyltransferase from the Moss *Physcomitrella patens* Involved in the Biosynthesis of a Novel Cell Wall Arabinoglucan." *The Plant Cell* 30 (6): 1293–1308.

Rogowski, A., Briggs, J. A., Mortimer, J. C., Tryfona, T., Terrapon, N., Lowe, E. C., Baslé, A., Morland, C., Day, A. M., Zheng, H., Rogers, T. E., Thompson, P., Hawkins, A. R., Yadav, M. P., Henrissat, B., Martens, E. C., Dupree, P., Gilbert, H. J., Bolam, D. N. 2015. "Glycan Complexity Dictates Microbial Resource Allocation in the Large Intestine." *Nature Communications* 6: 7481.

Rondeau-Mouro, C., Defer, D., Leboeuf, E., Lahaye, M. 2008. "Assessment of Cell Wall Porosity in *Arabidopsis thaliana* by NMR Spectroscopy." *International Journal of Biological Macromolecules* 42 (2): 83–92.

Roudier, F., Fernandez, A. G., Fujita, M., Himmelspach, R., Borner, G. H., Schindelman, G., Song, S., Baskin, T. I., Dupree, P., Wasteneys, G. O., Benfey, P. N. 2005. "COBRA, an Arabidopsis Extracellular Glycosyl-Phosphatidyl Inositol-Anchored Protein, Specifically Controls Highly Anisotropic Expansion through Its Involvement in Cellulose Microfibril Orientation." *The Plant Cell* 17 (6): 1749–63.

Saez-Aguayo, S., Rautengarten, C., Temple, H., Sanhueza, D., Ejsmentewicz, T., Sandoval-Ibañez, O., Doñas, D., Parra-Rojas, J. P., Ebert, B., Lehner, A., Mollet, J. C., Dupree, P., Scheller, H. V., Heazlewood, J. L., Reyes, F. C., Orellana, A. 2017. "UUAT1 Is a Golgi-Localized UDP-Uronic Acid Transporter That Modulates the Polysaccharide Composition of Arabidopsis Seed Mucilage." *The Plant Cell* 29 (1): 129–43.

Saito, F., Suyama, A., Oka, T., Yoko-O, T., Matsuoka, K., Jigami, Y., Shimma, Y. I. 2014. "Identification of Novel Peptidyl Serine α-Galactosyltransferase Gene Family in Plants." *The Journal of Biological Chemistry* 289 (30): 20405–20.

Sakamoto, T., Ogura, A., Inui, M., Tokuda, S., Hosokawa, S., Ihara, H., Kasai, N. 2011. "Identification of a GH62 α-L-Arabinofuranosidase Specific for Arabinoxylan Produced by Penicillium Chrysogenum." *Applied Microbiology and Biotechnology* 90 (1): 137–46.

Salmén, L. 2015. "Wood Morphology and Properties from Molecular Perspectives." *Annals of Forest Science* 72 (6): 679–84.

Santos, C. A., Morais, M. A. B., Terrett, O. M., Lyczakowski, J. J., Zanphorlin, L. M., Ferreira-Filho, J. A., Tonoli, C. C. C., Murakami, M. T., Dupree, P., Souza, A. P. 2019. "An Engineered GH1 β-Glucosidase Displays Enhanced Glucose Tolerance and Increased Sugar Release from Lignocellulosic Materials." *Scientific Reports* 9 (1): 4903.

Sato, Y., Namiki, N., Takehisa, H., Kamatsuki, K., Minami, H., Ikawa, H., Ohyanagi, H., Sugimoto, K., Itoh, J., Antonio, B. A., Nagamura, Y. 2013. "RiceFREND: A Platform for Retrieving Coexpressed Gene Networks in Rice." *Nucleic Acids Research* 41: 8.

Scheller, H. V., Ulvskov, P. 2010. "Hemicelluloses." *Annual Review of Plant Biology* 61 (1): 263–89.

Schilmiller, A. L., Stout, J., Weng, J. K., Humphreys, J., Ruegger, M. O., Chapple, C. 2009. "Mutations in the Cinnamate 4-Hydroxylase Gene Impact Metabolism, Growth and Development in Arabidopsis." *The Plant Journal* 60 (5): 771–82.

Schultink, A., Cheng, K., Park, Y. B., Cosgrove, D. J., Pauly, M. 2013. "The Identification of Two Arabinosyltransferases from Tomato Reveals Functional Equivalency of Xyloglucan Side Chain Substituents." *Plant Physiology* 163 (1): 86–94..

Schultink, A., Naylor, D., Dama, M., Pauly, M. 2015. "The Role of the Plant-Specific ALTERED XYLOGLUCAN9 Protein in Arabidopsis Cell Wall Polysaccharide O-Acetylation." *Plant Physiology* 167 (4): 1271–83.

"Sequoia Sempervirens (Coast Redwood) Description." n.d. Accessed August 20, 2019. https://www.conifers.org/cu/Sequoia.php.

Sewalt, V., Ni, W., Blount, J. W., Jung, H. G., Masoud, S. A., Howles, P. A., Lamb, C., Dixon, R. A. 1997. "Reduced Lignin Content and Altered Lignin Composition in Transgenic Tobacco Down-Regulated in Expression of L-Phenylalanine Ammonia-Lyase or Cinnamate 4-Hydroxylase." *Plant Physiology* 115 (1): 41–50.

Shadle, G., Chen, F., Srinivasa Reddy, M. S., Jackson, L., Nakashima, J., Dixon, R. A. 2007. "Down-Regulation of Hydroxycinnamoyl CoA: Shikimate Hydroxycinnamoyl Transferase in Transgenic Alfalfa Affects Lignification, Development and Forage Quality." *Phytochemistry* 68 (11): 1521–29..

Shi, J., Pattathil, S., Parthasarathi, R., Anderson, N. A., Kim, J. I., Venketachalam, S., Hahn, M. G., Chapple, C., Simmons, B. A., Singh, S. 2016. "Impact of Engineered Lignin Composition on Biomass Recalcitrance and Ionic Liquid Pretreatment Efficiency." *Green Chemistry* 18: 4884-4895.

Shigeto, J., Kiyonaga, Y., Fujita, K., Kondo, R., Tsutsumi, Y. 2013. "Putative Cationic Cell-Wall-Bound Peroxidase Homologues in Arabidopsis, AtPrx2, AtPrx25, and AtPrx71, Are Involved in Lignification." *Journal of Agricultural and Food Chemistry* 61 (16): 3781–88.

Shimada, T. L., Shimada, T., Hara-Nishimura, I. 2010. "A Rapid and Non-Destructive Screenable Marker, FAST, for Identifying Transformed Seeds of *Arabidopsis thaliana*." *The Plant Journal* 61 (3): 519–28.

Showalter, A. M., Basu, D. 2016. "Extensin and Arabinogalactan-Protein Biosynthesis: Glycosyltransferases, Research Challenges, and Biosensors." *Frontiers in Plant Science* 7 (June).

Sibout, R., Le Bris, P., Legée, F., Cézard, L., Renault, H., Lapierre, C. 2016. "Structural Redesigning Arabidopsis Lignins into Alkali-Soluble Lignins through the Expression of p-Coumaroyl-CoA:Monolignol Transferase PMT." *Plant Physiology* 170 (3): 1358–66.

Sibout, R., Eudes, A., Mouille, G., Pollet, B., Lapierre, C., Jouanin, L.,, Séguin A. 2005. "CINNAMYL ALCOHOL DEHYDROGENASE-C and -D Are the Primary Genes Involved in Lignin Biosynthesis in the Floral Stem of Arabidopsis." *The Plant Cell* 17 (7): 2059–76.

Sibout, R., Proost, S., Hansen, B. O., Vaid, N., Giorgi, F. M., Ho-Yue-Kuang, S., Legée, F., Cézart, L., Bouchabké-Coussa, O., Soulhat, C., Provart, N., Pasha, A., Le Bris, P., Roujol, D., Hofte, H., Jamet, E., Lapierre, C., Persson, S., Mutwil, M. 2017. "Expression Atlas and Comparative Coexpression Network Analyses Reveal Important Genes Involved in the Formation of Lignified Cell Wall in *Brachypodium distachyon*." *New Phytologist* 215 (3): 1009–25.

Simmons, T. J., Mortimer, J. C., Bernardinelli, O. D., Pöppler, A. C., Brown, S. P., deAzevedo, E. R., Dupree, R., Dupree, P. 2016. "Folding of Xylan onto Cellulose Fibrils in Plant Cell Walls Revealed by Solid-State NMR." *Nature Communications* 7: 13902.

Simonović, J., Stevanic, J., Djikanović, D., Salmén, L., Radotić, K. 2011. "Anisotropy of Cell Wall Polymers in Branches of Hardwood and Softwood: A Polarized FTIR Study." *Cellulose* 18 (6): 1433–40.

Sims, I. M., Craik, D. J., Bacic, A. 1997. "Structural Characterisation of Galactoglucomannan Secreted by Suspension-Cultured Cells of *Nicotiana plumbaginifolia*." *Carbohydrate Research* 303 (1): 79–92.

Smith, R. A., Schuetz, M., Roach, M., Mansfield, S. D., Ellis, B., Samuels, L. 2013. "Neighboring Parenchyma Cells Contribute to Arabidopsis Xylem Lignification, While Lignification of Interfascicular Fibers Is Cell Autonomous." *The Plant Cell* 25 (10): 3988–99.

Smith, R. A., Gonzales-Vigil, E., Karlen, S. D., Park, J. Y., Lu, F., Wilkerson, C. G., Samuels, L., Ralph, J., Mansfield, S. D. 2015. "Engineering Monolignol P-Coumarate Conjugates into Poplar and Arabidopsis Lignins." *Plant Physiology* 169 (4): 2992–3001.

Somerville, C. R. 2006. "Cellulose Synthesis in Higher Plants." *Annual Review of Cell and Developmental Biology* 22: 53–78.

Somssich, M., Vandenbussche, F., Ivakov, A., Funke, B., Ruprecht, C., Vissenberg, K., Van Der Straeten, D., Persson, S., Suslov, D. 2019. "Brassinosteroids Control Gravitropism by

Changing Cellulose Orientation and Mannan Content in Arabidopsis Hypocotyls." *BioRxiv*, February, 557777.

Sørensen, I., Pettolino, F. A., Wilson, S. M., Doblin, M. S., Johansen, B., Bacic, A., Willats, W. G. 2008. "Mixed-Linkage (1→3),(1→4)-β-d-Glucan Is Not Unique to the Poales and Is an Abundant Component of *Equisetum arvense* Cell Walls." *The Plant Journal* 54 (3): 510–21.

de Souza, W. R., Martins, P. K., Freeman, J., Pellny, T. K., Michaelson, L. V., Sampaio, B. L., Vinecky, F., Ribeiro, A. P., da Cunha, B. A. D. B., Kobayashi, A. K., de Oliveira, P. A., Campanha, R. B., Pacheco, T. F., Martarello, D. C. I., Marchiosi, R., Ferrarese-Filho, O., Dos Santos, W. D., Tramontina, R., Squina, F. M., Centeno, D. C., Gaspar, M., Braga, M. R., Tiné, M. A. S., Ralph, J., Mitchell, R. A. C., Molinari, H. B. C. 2018. "Suppression of a Single BAHD Gene in *Setaria viridis* Causes Large, Stable Decreases in Cell Wall Feruloylation and Increases Biomass Digestibility." *New Phytologist* 218 (1): 81–93.

de Souza, W. R., Pacheco, T. F., Duarte, K. E., Sampaio, B. L., de Oliveira Molinari, P. A., Martins, P. K., Santiago, T. R., Formighieri, E. F., Vinecky, F., Ribeiro, A. P., da Cunha, B. A. D. B., Kobayashi, A. K., Mitchell, R. A. C., de Sousa Rodrigues Gambetta, D., Molinari, H. B. C. 2019. "Silencing of a BAHD Acyltransferase in Sugarcane Increases Biomass Digestibility." *Biotechnology for Biofuels* 12 (1): 111.

Sterling, J. D., Atmodjo, M. A., Inwood, S. E., Kumar Kolli, V. S., Quigley, H. F., Hahn, M. G., Mohnen, D. 2006. "Functional Identification of an Arabidopsis Pectin Biosynthetic Homogalacturonan Galacturonosyltransferase." *Proceedings of the National Academy of Sciences* 103 (13): 5236–41.

Stewart, J. J., Akiyama, T., Chapple, C., Ralph, J., Mansfield, S. D. 2009. "The Effects on Lignin Structure of Overexpression of Ferulate 5-Hydroxylase in Hybrid Poplar." *Plant Physiology* 150 (2): 621–35.

Stranne, M., Ren, Y., Fimognari, L., Birdseye, D., Yan, J., Bardor, M., Mollet, J. C., Komatsu, T., Kikuchi, J., Scheller, H. V., Sakuragi, Y. 2018. "TBL10 Is Required for O-Acetylation of Pectic Rhamnogalacturonan-I in *Arabidopsis thaliana*." *The Plant Journal* 96 (4): 772–85.

Šuchová, K., Kozmon, S., Puchart, V., Malovíková, A., Hoff, T., Krogh, K. B. R. M., Biely, P. 2018. "Glucuronoxylan Recognition by GH 30 Xylanases: A Study with Enzyme and Substrate Variants." *Archives of Biochemistry and Biophysics* 643: 42–49.

Sundin, L., Vanholme, R., Geerinck, J., Goeminne, G., Höfer, R., Kim, H., Ralph, J., Boerjan, W. 2014. "Mutation of the Inducible ARABIDOPSIS THALIANA CYTOCHROME P450 REDUCTASE2 Alters Lignin Composition and Improves Saccharification." *Plant Physiology* 166 (4): 1956–71.

Szybalski, W., Kim, S. C., Hasan, N., Podhajska, A. J. 1991. "Class-IIS Restriction Enzymes — a Review." *Gene* 100 (April): 13–26.

Szyjanowicz, P. M., McKinnon, I., Taylor, N. G., Gardiner, J., Jarvis, M. C., Turner, S. R. 2004. "The Irregular Xylem 2 Mutant Is an Allele of Korrigan That Affects the Secondary Cell Wall of *Arabidopsis thaliana*." *The Plant Journal: For Cell and Molecular Biology* 37 (5): 730–40.

Takenaka, Y., Kato, K., Ogawa-Ohnishi, M., Tsuruhama, K., Kajiura, H., Yagyu, K., Takeda, A., Takeda, Y., Kunieda, T., Hara-Nishimura, I., Kuroha, T., Nishitani, K., Matsubayashi, Y., Ishimizu, T. 2018. "Pectin RG-I Rhamnosyltransferases Represent a Novel Plant-Specific Glycosyltransferase Family." *Nature Plants* 4 (9): 669–76.

Takenaka, Y., Watanabe, Y., Schuetz, M., Unda, F., Hill, J. L. Jr., Phookaew, P., Yoneda, A., Mansfield, S. D., Samuels, L., Ohtani, M., Demura, T. 2018. "Patterned Deposition of Xylan and Lignin Is Independent from That of the Secondary Wall Cellulose of Arabidopsis Xylem Vessels." *The Plant Cell* 30 (11): 2663–76.

Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., Hao, Z., Zhu, X., Avci, U., Miller, J. S., Baldwin, D., Pham, C., Orlando, R., Darvill, A., Hahn, M. G., Kieliszewski, M. J., Mohnen, D. 2013. "An Arabidopsis Cell Wall Proteoglycan Consists of Pectin and Arabinoxylan Covalently Linked to an Arabinogalactan Protein." *The Plant Cell* 25 (1): 270–87.

Tavares, E. Q., De Souza, A. P., Buckeridge, M. S. 2015. "How Endogenous Plant Cell-Wall Degradation Mechanisms Can Help Achieve Higher Efficiency in Saccharification of Biomass." *Journal of Experimental Botany* 66 (14): 4133–43.

Taylor, N. G., Howells, R. M., Huttly, A. K., Vickers, K., Turner, S. R. 2003. "Interactions among Three Distinct CesA Proteins Essential for Cellulose Synthesis." *Proceedings of the National Academy of Sciences* 100 (3): 1450–55.

Temple, H., Mortimer, J.C., Tryfona, T., Yu, X., Lopez-Hernandez, F., Sorieul, M., Anders, N., Dupree, P. 2019. "Two Members of the DUF579 Family Are Responsible for Arabinogalactan Methylation in Arabidopsis." *Plant Direct* 3 (2): e00117.

Temple, H., Saez-Aguayo, S., Reyes, F. C., Orellana, A. 2016. "The inside and Outside: Topological Issues in Plant Cell Wall Biosynthesis and the Roles of Nucleotide Sugar Transporters." *Glycobiology* 26 (9): 913–25.

Temple, H. personal communication.

Terashima, N., Awano, T., Takabe, K., Yoshida, M. 2004. "Formation of Macromolecular Lignin in Ginkgo Xylem Cell Walls as Observed by Field Emission Scanning Electron Microscopy." *Comptes Rendus Biologies* 327 (9–10): 903–10.

Terashima, N., Kitano, K., Kojima, M., Yoshida, M., Yamamoto, H., Westermark, U. 2009. "Nanostructural Assembly of Cellulose, Hemicellulose, and Lignin in the Middle Layer of Secondary Wall of Ginkgo Tracheid." *Journal of Wood Science* 55 (6): 409–16.

Terrett, O. M., Dupree, P. 2019. "Covalent Interactions between Lignin and Hemicelluloses in Plant Secondary Cell Walls." *Current Opinion in Biotechnology* 56: 97–104.

Tetreault, H. M., Scully, E. D., Gries, T., Palmer, N. A., Funnell-Harris, D. L., Baird, L., Seravalli, J., Dien, B. S., Sarath, G., Clemente, T. E., Sattler, S. E. 2018. "Overexpression of the *Sorghum bicolor* SbCCoAOMT Alters Cell Wall Associated Hydroxycinnamoyl Groups." *PLOS ONE* 13 (10): e0204153.

Teutsch, H. G., Hasenfratz, M. P., Lesot, A., Stoltz, C., Garnier, J. M., Jeltsch, J. M., Durst, F., Werck-Reichhart, D. 1993. "Isolation and Sequence of a CDNA Encoding the Jerusalem Artichoke Cinnamate 4-Hydroxylase, a Major Plant Cytochrome P450 Involved in the General Phenylpropanoid Pathway." *Proceedings of the National Academy of Sciences* 90 (9): 4102–6.

Thakur, B. R., Singh, R. K., Handa, A. K. 1997. "Chemistry and Uses of Pectin — A Review." *Critical Reviews in Food Science and Nutrition* 37 (1): 47–73.

Thomas, L. H., Forsyth, V. T., Martel, A., Grillo, I., Altaner, C. M., Jarvis, M. C. 2014. "Structure and Spacing of Cellulose Microfibrils in Woody Cell Walls of Dicots." *Cellulose* 21 (6): 3887–95.

217

Tobimatsu, Y., Schuetz, M. 2019. "Lignin Polymerization: How Do Plants Manage the Chemistry so Well?" *Current Opinion in Biotechnology* 56: 75–81.

Tokoh, C., Takabe, K., Sugiyama, J., Fujita, M. 2002. "CP MAS ^{13}C NMR and Electron Diffraction Study of Bacterial Cellulose Structure Affected by Cell Wall Polysaccharides." *Cellulose* 9 (3–4): 351–60.

Tolbert, A., Akinosho, H., Khunsupat, R., Naskar, A. K., Ragauskas, A. J. 2014. "Characterization and Analysis of the Molecular Weight of Lignin for Biorefining Studies." *Biofuels, Bioproducts and Biorefining* 8 (6): 836–56.

Tryfona, T., Liang, H. C., Kotake, T., Tsumuraya, Y., Stephens, E., Dupree, P. 2012. "Structural Characterization of Arabidopsis Leaf Arabinogalactan Polysaccharides." *Plant Physiology* 160 (2): 653–66.

Tryfona, T., Sorieul, M., Feijao, C., Stott, K., Rubtsov, D. V., Anders, N., Dupree, P. 2019. "Development of an Oligosaccharide Library to Characterise the Structural Variation in Glucuronoarabinoxylan in the Cell Walls of Vegetative Tissues in Grasses." *Biotechnology for Biofuels* 12: 109.

Tryfona, T., Theys, T. E., Wagner, T., Stott, K., Keegstra, K., Dupree, P. 2014. "Characterisation of FUT4 and FUT6 α-(1→2)-Fucosyltransferases Reveals That Absence of Root Arabinogalactan Fucosylation Increases Arabidopsis Root Growth Salt Sensitivity." *PLOS ONE* 9 (3): e93291.

Turlapati, P. V., Kim, K. W., Davin, L. B., Lewis, N. G. 2011. "The Laccase Multigene Family in *Arabidopsis thaliana*: Towards Addressing the Mystery of Their Gene Function(s)." *Planta* 233 (3): 439–70.

Turner, S., Kumar, M. 2018. "Cellulose Synthase Complex Organization and Cellulose Microfibril Structure." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376 (2112): 20170048.

Umemura, M., Yuguchi, Y. 2005. "Conformational Folding of Xyloglucan Side Chains in Aqueous Solution from Molecular Dynamics Simulation." *Carbohydrate Research* 340 (16): 2520–32.

Urbániková, L., Vršanská, M., Krogh K. B. R. M., Hoff, T., Biely, P. 2011. "Structural Basis for Substrate Recognition by Erwinia Chrysanthemi GH30 Glucuronoxylanase." *FEBS Journal* 278 (12): 2105–16.

Urbanowicz, B. R., Peña, M. J., Moniz, H. A., Moremen, K. W., York, W. S. 2014. "Two Arabidopsis Proteins Synthesize Acetylated Xylan in Vitro." *The Plant Journal : For Cell and Molecular Biology* 80 (2): 197–206.

Urbanowicz, B. R., Peña, M. J., Ratnaparkhe, S., Avci, U., Backe, J., Steet, H. F., Foston, M., Li, H., O'Neill, M. A., Ragauskas, A. J., Darvill, A. G., Wyman, C., Gilbert, H. J., York, W. S. 2012. "4-O-Methylation of Glucuronic Acid in Arabidopsis Glucuronoxylan Is Catalyzed by a Domain of Unknown Function Family 579 Protein." *Proceedings of the National Academy of Sciences* 109 (35): 14253–58.

Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sørlie, M., Eijsink, V. G. 2010. "An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides." *Science* 330 (6001): 219–22.

Vain, T., Crowell, E. F., Timpano, H., Biot, E., Desprez, T., Mansoori, N., Trindade, L. M., Pagant, S., Robert, S., Höfte, H., Gonneau, M., Vernhettes, S. 2014. "The Cellulase KORRIGAN Is Part of the Cellulose Synthase Complex." *Plant Physiology* 165 (4): 1521–32.

Van Acker, R., Déjardin, A., Desmet, S., Hoengenaert, L., Vanholme, R., Morreel, K., Laurans, F., Kim, H., Santoro, N., Foster, C., Goeminne, G., Légée, F., Lapierre, C., Pilate, G., Ralph, J., Boerjan, W. 2017. "Different Routes for Conifer- and Sinapaldehyde and Higher Saccharification upon Deficiency in the Dehydrogenase CAD1." *Plant Physiology* 175 (3): 1018–39.

Van Acker, R., Vanholme, R., Storme, V., Mortimer, J. C., Dupree, P., Boerjan, W. 2013. "Lignin Biosynthesis Perturbations Affect Secondary Cell Wall Composition and Saccharification Yield in *Arabidopsis thaliana*." *Biotechnology for Biofuels* 6: 46.

Van Sandt, V. S., Suslov, D., Verbelen, J. P., Vissenberg, K. 2007. "Xyloglucan Endotransglucosylase Activity Loosens a Plant Cell Wall." *Annals of Botany* 100 (7): 1467–73.

219

Vandavasi, V. G., Putnam, D. K., Zhang, Q., Petridis, L., Heller, W. T., Nixon, B. T., Haigler, C. H., Kalluri, U., Coates, L., Langan, P., Smith, J. C., Meiler, J., O'Neill, H. 2016. "A Structural Study of CESA1 Catalytic Domain of Arabidopsis Cellulose Synthesis Complex: Evidence for CESA Trimers." *Plant Physiology* 170 (1): 123–35.

Vanholme, Ruben. personal communication.

Vanholme, R., Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., Kim, H., Cross, J., Morreel, K., Araujo, P., Welsh, L., Haustraete, J., McClellan, C., Vanholme, B., Ralph, J., Simpson, G. G., Halpin, C., Boerjan, W. 2013. "Caffeoyl Shikimate Esterase (CSE) Is an Enzyme in the Lignin Biosynthetic Pathway in Arabidopsis." *Science* 341 (6150): 1103–6.

Vanholme, R., De Meester, B., Ralph, J., Boerjan, W. 2019. "Lignin Biosynthesis and Its Integration into Metabolism." *Current Opinion in Biotechnology* 56 (April): 230–39.

Vanholme, R., Ralph, J., Akiyama, T., Lu, F., Pazo, J. R., Kim, H., Christensen, J. H., Van Reusel, B., Storme, V., De Rycke, R., Rohde, A., Morreel, K., Boerjan, W. 2010. "Engineering Traditional Monolignols out of Lignin by Concomitant Up-Regulation of F5H1 and down-Regulation of COMT in Arabidopsis." *The Plant Journal* 64 (6): 885–97.

Vardakou, M., Dumon, C., Murray, J. W., Christakopoulos, P., Weiner, D. P., Juge, N., Lewis, R. J., Gilbert, H. J., 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.

Vega-Sánchez, M. E., Verhertbruggen, Y., Scheller, H. V., Ronald, P. C. 2013. "Abundance of Mixed Linkage Glucan in Mature Tissues and Secondary Cell Walls of Grasses." *Plant Signaling & Behavior* 8 (2): e23143

Verhertbruggen, Y., Yin, L., Oikawa, A., Scheller, H. V. 2011. "Mannan Synthase Activity in the CSLD Family." *Plant Signaling & Behavior* 6 (10): 1620–23.

Vermaas, J. V., Petridis, L., Qi, X., Schulz, R., Lindner, B., Smith, J. C. 2015. "Mechanism of Lignin Inhibition of Enzymatic Biomass Deconstruction." *Biotechnology for Biofuels* 8: 217.

Voiniciuc, C., Dama, M., Gawenda, N., Stritt, F., Pauly, M. 2019. "Mechanistic Insights from Plant Heteromannan Synthesis in Yeast." *Proceedings of the National Academy of Sciences* 116 (2): 522–27.

Voiniciuc, C., Engle, K. A., Günl, M., Dieluweit, S., Schmidt, M. H., Yang, J. Y., Moremen, K. W., Mohnen, D., Usadel, B. 2018. "Identification of Key Enzymes for Pectin Synthesis in Seed Mucilage." *Plant Physiology* 178 (3): 1045–64.

Voiniciuc, C., Günl, M., Schmidt, M. H., Usadel, B. 2015. "Highly Branched Xylan Made by IRREGULAR XYLEM14 and MUCILAGE-RELATED21 Links Mucilage to Arabidopsis Seeds." *Plant Physiology* 169 (4): 2481–95.

Voiniciuc, C., Schmidt, M. H., Berger, A., Yang, B., Ebert, B., Scheller, H. V., North, H. M., Usadel, B., Günl, M. 2015. "MUCILAGE-RELATED10 Produces Galactoglucomannan That Maintains Pectin and Cellulose Architecture in Arabidopsis Seed Mucilage." *Plant Physiology* 169 (1): 403–20.

Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., Laue, E. D. 2005. "The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline." *Proteins* 59 (4): 687–96.

Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Geddes, B., Lu, F., Ralph, J. 2015. "Syringyl Lignin Production in Conifers: Proof of Concept in a Pine Tracheary Element System." *Proceedings of the National Academy of Sciences* 112 (19): 6218–23.

Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Torr, K., Donaldson, L., Pears, L., Ralph, J. 2011. "CCoAOMT Suppression Modifies Lignin Composition in *Pinus radiata*." *The Plant Journal* 67 (1): 119–29.

Wang, H. M., Loganathan, D., Linhardt, R. J. 1991. "Determination of the pK_a of Glucuronic Acid and the Carboxy Groups of Heparin by ^{13}C-Nuclear-Magnetic-Resonance Spectroscopy." *Biochemical Journal* 278 (Pt 3): 689–95.

Wang, L., Wang, W., Wang, Y. Q., Liu, Y. Y., Wang, J. X., Zhang, X. Q., Ye, D., Chen, L. Q. 2013. "Arabidopsis Galacturonosyltransferase (GAUT) 13 and GAUT14 Have Redundant Functions in Pollen Tube Growth." *Molecular Plant* 6 (4): 1131–48.

221

Wang, T., Chen, Y., Tabuchi, A., Cosgrove, D. J., Hong, M. 2016a. "The Target of β-Expansin EXPB1 in Maize Cell Walls from Binding and Solid-State NMR Studies." *Plant Physiology* 172 (4): 2107-2119.

Wang, T., Park, Y. B., Caporini, M. A., Rosay, M., Zhong, L., Cosgrove, D. J., Hong, M. 2013. "Sensitivity-Enhanced Solid-State NMR Detection of Expansin's Target in Plant Cell Walls." *Proceedings of the National Academy of Sciences* 110 (41): 16444–49.

Wang, T., Park, Y. B., Cosgrove, D. J., Hong, M. 2015. "Cellulose-Pectin Spatial Contacts Are Inherent to Never-Dried Arabidopsis Primary Cell Walls: Evidence from Solid-State Nuclear Magnetic Resonance." *Plant Physiology* 168 (3): 871–84.

Wang, T., Salazar, A., Zabotina, O. A., Hong, M. 2014. "Structure and Dynamics of Brachypodium Primary Cell Wall Polysaccharides from Two-Dimensional 13C Solid-State Nuclear Magnetic Resonance Spectroscopy." *Biochemistry* 53 (17): 2840–54.

Wang, T., Yang, H., Kubicki, J. D., Hong, M. 2016b. "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): 2210-22.

Wang, Y., Mortimer, J. C., Davis, J., Dupree, P., Keegstra, K. 2013. "Identification of an Additional Protein Involved in Mannan Biosynthesis." *The Plant Journal* 73 (1): 105–17.

Watanabe, T., Koshijima, T. 1988. "Evidence for an Ester Linkage between Lignin and Glucuronic Acid in Lignin–Carbohydrate Complexes by DDQ-Oxidation." *Agricultural and Biological Chemistry* 52 (11): 2953–55.

White, P. B., Wang, T., Park, Y. B., Cosgrove, D. J., Hong, M. 2014. "Water–Polysaccharide Interactions in the Primary Cell Wall of *Arabidopsis thaliana* from Polarization Transfer Solid-State NMR." *Journal of the American Chemical Society* 136 (29): 10399–409.

Whitney, S. E. C., Brigham, J. E., Darke, A. H., Reid, J. S. G., Gidley, M. J. 1998. "Structural Aspects of the Interaction of Mannan-Based Polysaccharides with Bacterial Cellulose." *Carbohydrate Research* 307 (3): 299–309.

Wilkerson, C. G., Mansfield, S. D., Lu, F., Withers, S., Park, J. Y., Karlen, S. D., Gonzales-Vigil, E., Padmakshan, D., Unda, F., Rencoret, J., Ralph, J. 2014. "Monolignol Ferulate Transferase Introduces Chemically Labile Linkages into the Lignin Backbone." *Science* 344 (6179): 90–93.

Wormit, A., Usadel, B. 2018. "The Multifaceted Role of Pectin Methylesterase Inhibitors (PMEIs)." *International Journal of Molecular Sciences* 19 (10).

Xiao, C., Zhang, T., Zheng, Y., Cosgrove, D. J., Anderson, C. T. 2016. "Xyloglucan Deficiency Disrupts Microtubule Stability and Cellulose Biosynthesis in Arabidopsis, Altering Cell Growth and Morphogenesis." *Plant Physiology* 170 (1): 234–49.

Xiong, G., Cheng, K., Pauly, M. 2013. "Xylan O-Acetylation Impacts Xylem Development and Enzymatic Recalcitrance as Indicated by the Arabidopsis Mutant *tbl29*." *Molecular Plant* 6 (4): 1373–75.

Xiong, G., Dama, M., Pauly, M. 2015. "Glucuronic Acid Moieties on Xylan Are Functionally Equivalent to O-Acetyl-Substituents." *Molecular Plant* 8 (7): 1119–21.

Xue, S., Uppugundla, N., Bowman, M. J., Cavalier, D., Da Costa Sousa, L., E Dale, B., Balan, V. 2015. "Sugar Loss and Enzyme Inhibition Due to Oligosaccharide Accumulation during High Solids-Loading Enzymatic Hydrolysis." *Biotechnology for Biofuels* 8: 195.

Yan, X., Liu, J., Kim, H., Liu, B., Huang, X., Yang, Z., Lin, Y. J., Chen, H., Yang, C., Wang, J. P., Muddiman, D. C., Ralph, J., Sederoff, R. R., Li, Q., Chiang, V. L. 2019. "CAD1 and CCR2 Protein Complex Formation in Monolignol Biosynthesis in *Populus trichocarpa*." *The New Phytologist* 222 (1): 244–60.

Yang, H., Wang, T., Oehme, D., Petridis, L., Hong, M., Kubicki, J. D. 2018. "Structural Factors Affecting ^{13}C NMR Chemical Shifts of Cellulose: A Computational Study." *Cellulose* 25 (1): 23–36.

Yang, W., Schuster, C., Beahan, C. T., Charoensawan, V., Peaucelle, A., Bacic, A., Doblin, M. S., Wightman, R., Meyerowitz, E. M. 2016. "Regulation of Meristem Morphogenesis by Cell Wall Synthases in Arabidopsis." *Current Biology* 26 (11): 1404–15.

Yarbrough, J. M., Mittal, A., Mansfield, E., Taylor, L. E., Hobdey, S. E., Sammond, D. W., Bomble, Y. J., Crowley, M. F., Decker, S. R., Himmel, M. E., Vinzant, T. B. 2015. "New Perspective on Glycoside Hydrolase Binding to Lignin from Pretreated Corn Stover." *Biotechnology for Biofuels* 8: 214.

Yi Chou, E., Schuetz, M., Hoffmann, N., Watanabe, Y., Sibout, R., Samuels, A. L. 2018. "Distribution, Mobility, and Anchoring of Lignin-Related Oxidative Enzymes in Arabidopsis Secondary Cell Walls." *Journal of Experimental Botany* 69 (8): 1849–59.

Yin, Y., Chen, H., Hahn, M. G., Mohnen, D., Xu, Y. 2010. "Evolution and Function of the Plant Cell Wall Synthesis-Related Glycosyltransferase Family 8." *Plant Physiology* 153 (4): 1729–46.

Yin, Y., Mohnen, D., Gelineo-Albersheim, I., Xu, Y., Hahn, M. G. 2010. "Glycosyltransferases of the GT8 Family." *Annual Plant Reviews*, 167–211.

Yu, L., Lyczakowski, J. J., Pereira, C. S., Kotake, T., Yu, X., Li, A., Mogelsvang, S., Skaf, M. S., Dupree, P. 2018. "The Patterned Structure of Galactoglucomannan Suggests It May Bind to Cellulose in Seed Mucilage." *Plant Physiology* 178 (3): 1011–1026.

Yu, L., Shi, D., Li, J., Kong, Y., Yu, Y., Chai, G., Hu, R., Wang, J., Hahn, M. G., Zhou, G. 2014. "CELLULOSE SYNTHASE-LIKE A2, a Glucomannan Synthase, Is Involved in Maintaining Adherent Mucilage Structure in Arabidopsis Seed." *Plant Physiology* 164 (4): 1842–56.

Yuan, Y., Teng, Q., Zhong, R., Haghighat, M., Richardson, E. A., Ye, Z. H. 2016. "Mutations of Arabidopsis TBL32 and TBL33 Affect Xylan Acetylation and Secondary Wall Deposition." *PLoS ONE* 11 (1): e0146460.

Yuan, Y., Teng, Q., Zhong, R., Ye, Z. H. 2016. "TBL3 and TBL31, Two Arabidopsis DUF231 Domain Proteins, Are Required for 3-O-Monoacetylation of Xylan." *Plant and Cell Physiology* 57 (1): 35–45.

Yue, F., Lu, F., Ralph, S., Ralph, J. 2016. "Identification of 4–O–5-Units in Softwood Lignins via Definitive Lignin Models and NMR." *Biomacromolecules* 17 (6): 1909–20.

Zabotina, O. A. 2012. "Xyloglucan and Its Biosynthesis." *Frontiers in Plant Science* 3: 134.

Zabotina, O. A., Avci, U., Cavalier, D., Pattathil, S., Chou, Y. H., Eberhard, S., Danhof, L., Keegstra, K., Hahn, M. G. 2012. "Mutations in Multiple XXT Genes of Arabidopsis Reveal the Complexity of Xyloglucan Biosynthesis." *Plant Physiology* 159 (4): 1367–84.

Zeng, W., Jiang, N., Nadella, R., Killen, T. L., Nadella, V., Faik, A. 2010. "A Glucurono(Arabino)Xylan Synthase Complex from Wheat Contains Members of the GT43, GT47, and GT75 Families and Functions Cooperatively." *Plant Physiology* 154 (1): 78–97.

Zeng, W., Lampugnani, E. R., Picard, K. L., Song, L., Wu, A. M., Farion, I. M., Zhao, J., Ford, K., Doblin, M. S., Bacic, A. 2016. "Asparagus IRX9, IRX10, and IRX14A Are Components

of an Active Xylan Backbone Synthase Complex That Forms in the Golgi Apparatus." *Plant Physiology* 171 (1): 93–109.

Zhang, B., Zhang, L., Li, F., Zhang, D., Liu, X., Wang, H., Xu, Z., Chu, C., Zhou, Y. 2017. "Control of Secondary Cell Wall Patterning Involves Xylan Deacetylation by a GDSL Esterase." *Nature Plants* 3 (3): 17017.

Zhang, L., Gao, C., Mentink-Vigier, F., Tang, L., Zhang, D., Wang, S., Cao, S., Xu, Z., Liu, X., Wang, T., Zhou, Y., Zhang, B. 2019. "Arabinosyl Deacetylase Modulates the Arabinoxylan Acetylation Profile and Secondary Wall Formation." *The Plant Cell* 31 (5): 1113-1126

Zhang, X., Dominguez, P. G., Kumar, M., Bygdell, J., Miroshnichenko, S., Sundberg, B., Wingsle, G., Niittylä, T. 2018. "Cellulose Synthase Stoichiometry in Aspen Differs from Arabidopsis and Norway Spruce." *Plant Physiology* 177 (3): 1096–1107.

Zhang, Y., Nikolovski, N., Sorieul, M., Vellosillo, T., McFarlane, H. E., Dupree, R., Kesten, C., Schneider, R., Driemeier, C., Lathe, R., Lampugnani, E., Yu, X., Ivakov, A., Doblin, M. S., Mortimer, J. C., Brown, S. P., Persson, S., Dupree, P. 2016. "Golgi-Localized STELLO Proteins Regulate the Assembly and Trafficking of Cellulose Synthase Complexes in Arabidopsis." *Nature Communications* 7 (1): 11656.

Zhao, Q., Tobimatsu, Y., Zhou, R., Pattathil, S., Gallego-Giraldo, L., Fu, C., Jackson, L. A., Hahn, M. G., Kim, H., Chen, F., Ralph, J., Dixon, R. A. 2013. "Loss of Function of Cinnamyl Alcohol Dehydrogenase 1 Leads to Unconventional Lignin and a Temperature-Sensitive Growth Defect in *Medicago truncatula*." *Proceedings of the National Academy of Sciences* 110 (33): 13660–65.

Zhao, Zhen, Vincent H. Crespi, James D. Kubicki, Daniel J. Cosgrove, and Linghao Zhong. 2014. "Molecular Dynamics Simulation Study of Xyloglucan Adsorption on Cellulose Surfaces: Effects of Surface Hydrophobicity and Side-Chain Variation." *Cellulose* 21 (2): 1025–39. https://doi.org/10.1007/s10570-013-0041-1.

Zheng, Y., Wang, X., Chen, Y., Wagner, E., Cosgrove, D. J. 2018. "Xyloglucan in the Primary Cell Wall: Assessment by FESEM, Selective Enzyme Digestions and Nanogold Affinity Tags." *The Plant Journal* 93 (2): 211–26.

Zhong, R., Cui, D., Phillips, D. R., Ye, Z. H. 2018. "A Novel Rice Xylosyltransferase Catalyzes the Addition of 2-O-Xylosyl Side Chains onto the Xylan Backbone." *Plant and Cell Physiology* 59 (3): 554–65.

Zhong, R., Cui, D., Ye, Z. H. 2017. "Regiospecific Acetylation of Xylan Is Mediated by a Group of DUF231-Containing O-Acetyltransferases." *Plant and Cell Physiology* 58 (12): 2126–38.

Zhong, R., Cui, D., Ye, Z. H. 2018. "Members of the DUF231 Family Are O-Acetyltransferases Catalyzing 2-O- and 3-O-Acetylation of Mannan." *Plant and Cell Physiology* 59 (11): 2339–49.

Zhu, L., Dama, M., Pauly, M. 2017. "Identification of an Arabinopyranosyltransferase from Physcomitrella Patens Involved in the Synthesis of the Hemicellulose Xyloglucan." *BioRxiv* 241521.

Zubieta, C., Kota, P., Ferrer, J. L., Dixon, R. A., Noel, J. P. 2002. "Structural Basis for the Modulation of Lignin Monomer Methylation by Caffeic Acid/5-Hydroxyferulic Acid 3/5-O-Methyltransferase." *The Plant Cell* 14 (6): 1265–77.

Zuo, J., Niu, Q. W., Nishizawa, N., Wu, Y., Kost, B., Chua, N. H. 2000. "KORRIGAN, an Arabidopsis Endo-1,4-β-Glucanase, Localizes to the Cell Plate by Polarized Targeting and Is Essential for Cytokinesis." *The Plant Cell* 12 (7): 1137–53.