Cell Wall Chemotyping for Functional Genomics

Applications of Pyrolysis–Gas Chromatography / Mass Spectrometry

Lorenz Gerber
Faculty of Forest Sciences,
Department of Forest Genetics and Plant Physiology,
Umeå

Doctoral Thesis
Swedish University of Agricultural Sciences
Umeå 2012

Acta Universitatis Agriculturae Sueciae
2012:59

ISSN, 1652-6880
ISBN, 978-91-576-7706-8

Print: Arkitektkopia, Umeå 2012

Cell Wall Chemotyping for Functional Genomics

Applications of Pyrolysis–Gas Chromatography / Mass Spectrometry

Abstract

The interest in lignocellulose as a sustainable resource for energy and materials has fueled research on biotechnology applications in tree breeding to improve biomass production and wood properties. An important aspect of this research is the basic understanding of gene function in wood formation, where analysis of wood chemistry and wood structure is of utmost importance.

Current research strategies often involve large-scale screening of plant material, and therefore there is a need for rapid chemical characterisation and classification (chemotyping) of samples. Pyrolysis-Gas Chromatography/Mass Spectrometry (Py-GC/MS) provides a chemical fingerprint with more than hundred peaks representing each sample. Despite the fact that the technique is very informative, the wealth of data resulting from each sample has prohibited its use as a high–throughput method in large scale projects. For this reason, a novel application of Py-GC/MS was developed. An automated data processing pipeline was created and implemented, which together with a new simplified rule-set to analyse Py-GC/MS data overcame the bottleneck of time-consuming data handling. The added value of the method was demonstrated by fingerprinting and classifying the wood of 736 transgenic hybrid aspen (*Populus tremula x tremuloides*) trees representing 44 different genotypes. Among these genes were fructokinases (*FRK1* and *2*) and sucrose synthase (*SUS*) isoforms, both involved in primary carbon metabolism. A battery of chemotyping tools demonstrated that the downregulation of *FRK* resulted more specifically in a decrease in cellulose biosynthesis, whereas the repression of *SUS* affected the amount of all of the major wood polymers. It was further demonstrated that the wood of transgenic *SUS* trees was affected in structure and mechanical properties. In another study, it was shown that down-regulating a lignin biosynthetic gene, cinnamate 4-hydroxylase, resulted in a modified wood structure in hybrid aspen. Interestingly, the large reduction in lignin content resulted in only minor effects on the ultimate tensile strength of the wood. Finally, the battery of chemotyping tools was used in a study to show that the transcription factor MYB103 is required for ferrulate-5-hydroxylase (*F5H*) expression and S-lignin biosynthesis in *Arabidopsis thaliana*.

Keywords: aspen, *Populus*, cell wall analysis, analytical pyrolysis, wood formation, lignocellulose

Author's address: Lorenz Gerber, SLU, Department of Forest Genetics and Plant Physiology,
901 83 Umeå, Sweden.
E-mail: lorenz.gerber@slu.se

Contents

List of Publications **7**

Abbreviations **11**

1 Introduction **13**
1.1 Why Study Wood? 13
1.2 Wood Structure 13
1.2.1 Macrostructure of the Stem 14
1.2.2 Wood Anatomy 15
1.2.3 Wood Chemistry 16
1.2.4 Ultrastructure and Cell Wall Architecture 22
1.3 Bulk Properties 25
1.4 Chemical Characterisation of Wood 26
1.4.1 The Basic Tool-set 27
1.5 Wood Formation and Functional Genomics 31

2 Objectives **33**

3 Methodological Overview **33**
3.1 Plant Material and Sample Preparation 33
3.2 Pyrolysis–Gas Chromatography / Mass-spectrometry 34
3.3 Data Processing and Analysis 36

4 Results and Discussion **39**
4.1 Py-GC/MS for Wood Analysis 39
4.2 Data Integration in Screening Projects 41
4.3 Case Studies 42
4.3.1 Functional Studies of Genes in Primary Carbohydrate Metabolism 43
4.3.2 Functional Genomics of Genes involved in Lignin Biosynthesis 46

5 Conclusions and Future Perspectives **49**
5.1 Concluding Remarks 49
5.2 Future Perspectives 50

Acknowledgements **52**

References **53**

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I **Lorenz Gerber**, Mattias Eliasson, Johan Trygg, Thomas Moritz, Björn Sundberg (2012). Multivariate curve resolution provides a high-throughput data processing pipeline for Pyrolysis-Gas Chromatography/Mass Spectrometry. *Journal of Analytical and Applied Pyrolysis* 95, 95–100.

II Rui C. Pinto, **Lorenz Gerber**, Mattias Eliasson, Björn Sundberg, Johan Trygg (2012). Strategy for minimizing between-study variation of large scale phenotypic experiments using multivariate analysis. *submitted*

III **Lorenz Gerber**, Bo Zhang, Melissa Roach, András Gorzsás, Manoj Kumar, Totte Niittylä, Ingo Burgert, Björn Sundberg (2012). Reduced sucrose synthase activity in developing wood of hybrid aspen limits carbon allocation to all main cell wall polymers and modifies wood structure. *manuscript*

IV Melissa Roach, **Lorenz Gerber**, David Sandquist, András Gorzsás, Mattias Hedenström, Manoj Kumar, Marie-Caroline Steinhauser, Regina Feil, Geoffrey Daniel, Mark Stitt, Björn Sundberg, Totte Niittylä (2012). Fructokinase is required for carbon partitioning to cellulose in aspen wood. *The Plant Journal* 70(6), 967–977.

V Ingela Bjurhager, Anne-Mari Olsson, Bo Zhang, **Lorenz Gerber**, Manoj Kumar, Lars A. Berglund, Ingo Burgert, Björn Sundberg, Lennart Salmén (2010). Ultrastructure and mechanical properties of Populus wood with reduced lignin content caused by transgenic down-regulation of cinnamate 4-hydroxylase. *Biomacromolecules* 11(9), 2359–65.

VI David Öhman, Brecht Demendts, Manoj Kumar, **Lorenz Gerber**, András Gorzsás, Geert Goeminne, Mattias Hedenström, Brian Ellis, Wout Boerjan, Björn Sundberg (2012). MYB103 promotes FERULATE-5-HYDROXYLASE expression and syringyl lignin biosynthesis in Arabidopsis stems. *submitted*

The contribution of Lorenz Gerber to the papers included in this thesis was as follows:

I Design of experiments, practical work, data interpretation, writing article

II Establishing theoretical framework of data evaluation strategy, design of analytical experiments, practical work, data interpretation, writing article

III Designing experiments, Py-GC/MS analysis, density measurments, data interpretation, writing article

IV Py-GC/MS analysis, data interpretation, contributed to article writing

V Wet chemical cell wall analysis, Py-GC/MS analysis, data interpretation, contributed to article writing

VI Py-GC/MS analysis, data interpretation, contributed to article writing

Abbreviations

AFM	Atomic Force Microscopy.
AGP	Arabinogalactan Protein.
AIR	Alcohol Insoluble Residue.
C4H	Cinnamate-4-Hydroxylase.
CesA	Cellulose Synthase.
CMF	Cellulose Microfibril.
CML	Compound Middle Lamella.
DFRC	Derivatisation followed by reductive cleavage.
ESI-IT-MS	Electrospray-Ion Trap-Mass Spectrometry.
F5H	Ferulate 5-Hydroxylase.
FRK	Fructokinase.
FTIR	Fourier Transform Infrared.
G-Lignin	Guaiacyl type lignin.
H-Lignin	p-Hydoxyphenol lignin.
HPAEC-PAD	High Pressure Anion Exchange Chromatography-Pulsed Amperometric Detection.
HSQC	Heteronuclear Single Quantum Coherence.
HXK	Hexokinase.
INV	Invertase.
LC-MS/MS	Liquid Chromatography Tandem Mass Spectrometry.
LCC	Lignin–Carbohydrate Complex.
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time of Flight.
MCR-AR	Multivariate Curve Resolution-Alternating Regression.
MFA	Microfibril Angle.
ML	Middle Lamella.
MYB	family of transcription factors in higher plants, from *Myeloblastosis*.
NIR	Near Infrared.
NMR	Nuclear Magnetic Resonance.
OPLS-DA	Orthogonal Projection to Latent Structures-Discriminant Analysis.
P	Primary cell wall layer.
PCA	Principal Component Analysis.

Py-GC/MS	Pyrolysis-Gas Chromatography/Mass Spectrometry.
Py-MBMS	Pyrolysis-Molecular Beam Mass Spectrometry.
RG	Rhamnogalacturonan.
RNAi	Ribonucleic Acid Interference.
RSD	Relative Standard Deviation.
S-Lignin	Syringyl type lignin.
S/G	ratio of S to G type lignin.
S1,2,3	Secondary wall layer 1,2,3.
SND	Secondary wall-associated NAC domain.
SUS	Sucrose Synthase.
T-DNA	Transfer Deoxyribonucleic Acid.
TF	Transcription Factor.
TFA	Trifluoroacetic acid.
UDP-Glc	Uridine Diphosphate Glucose.

1 Introduction

1.1 Why Study Wood?

Wood, the characteristic tissue of trees, has been used throughout human history as a raw-material in many different applications: as construction material, for tools, or as an energy source, to name some of them. Still nowadays, more than two billion people use wood as a direct source of energy (http://www.fao.org/forestry/energy/en). Wood is also one of the most abundant biomasses on earth and vast forests account for an important share in the global carbon pools.

Reserves of easily available fossil carbon sources are limited. The increase of atmospheric CO_2 levels from burning fossil carbon and its impact on the global climate is today a fact. This sparked increasing efforts to develop and engineer industrial processes towards the use of renewable hydrocarbons (Gallezot, 2012). Lignocellulosic raw materials such as wood are a potential source of renewable biomaterial. However, there are legitimate fears for a surge in global food prices due to competing land use between food– and biomass production (http://www.fao.org/bioenergy/47280/en/). Further, increasing the area of cultivated land could have increasingly adverse effects on environment and biodiversity. An arguable strategy to address those issues is to promote methods to further increase the efficiency of agricultural and agro–forestry production systems. Recent techniques and practices used to increase the production of wood and lignocellulosic biomass include the use of biotechnology in tree breeding, clonal– and short rotation forestry (Neale and Savolainen, 2004; Tullus et al., 2012). Biotechnology can be used in breeding programs to select for offspring with favourable genes (marker assisted breeding), but also for genetic improvement of elite clones by transgenic technology (Harfouche et al., 2011). The potential of forest biotechnology is currently intensively explored and progress in this research will benefit from interdisciplinary collaborations across the fields of biology and chemistry as well as material– and computational sciences.

1.2 Wood Structure

Booker and Sell (1998) divided wood structure into four levels according to the length scale on which it can be observed. The macroscopic structure can be studied by bare eyes and includes tissues such as early and late wood, sap and heart wood. At the microscopic level, the distribution of single cell types can be observed by light microscopy. Together the macroscopic-

and microscopic features describe the anatomy of wood. At the ultrastructural level of organisation, best studied by electron microscopy, sub cellular details such as the distinct layered architecture of the cell wall become visible. Certain details of the nanostructure can still be resolved by electron microscopy while composition and properties of large molecular agglomerates in the range of nanometers are best studied by physico–chemical methods such as size-exclusion chromatography, viscosimetry, laser light scattering, thermal analysis, x–ray diffraction, NMR spectroscopy and similar approaches. Finally, the molecular level of organisation, in the order of tenths of nanometers can be visualised directly with AFM microscopy. However to explore the chemical composition of the various polymers, a wealth of indirect analytical methods and strategies are commonly employed. The subsequent paragraphs will describe in more detail all mentioned structural levels of wood and cell walls. As the description of ultrastructure and cell wall architecture readily refers to the chemical properties and structures of cell wall constituting molecular aggregates, chemical components will be described just after wood anatomy but before cell wall ultrastructure.

1.2.1 Macrostructure of the Stem

Anatomical features of a tree stem are shown in *Figure 1*. These are formed by cell divisions and differentiation in the cambial region (Mellerowicz et al., 2001). Briefly, to the inner side of the vascular cambium, cambial daughter cells differentiate to xylem tissue which is colloquially termed wood. To the outer side of the cambium phloem is formed. The bark subsequentially comprises functional phloem, non-functional living and dead phloem as well as cells formed in the cork cambium (Ray, 2006, p. 427). Seasonal growth rings in the xylem of trees from temperate regions are the result of radially wide cells formed during vigorous growth in the early growing season in contrast to more narrow cells formed in the late growing season. In the centre of the tree stem is the pith, a primary tissue that originates from the apical meristem. Trees, older than about 10–20 years, often have a dark coloration of the innermost xylem when exposed to air. This is called the heartwood, while the outer light colored part of the xylem is termed sapwood. The coloration of the heartwood is mostly due to phenolic extractives that become oxidised when exposed to air. The extractives make the heartwood hydrophobic and also provide antimicrobial properties (Scheffer, 1966). Physiologically, sapwood is the part of the wood which is active in water transport. Heartwood has structural function.

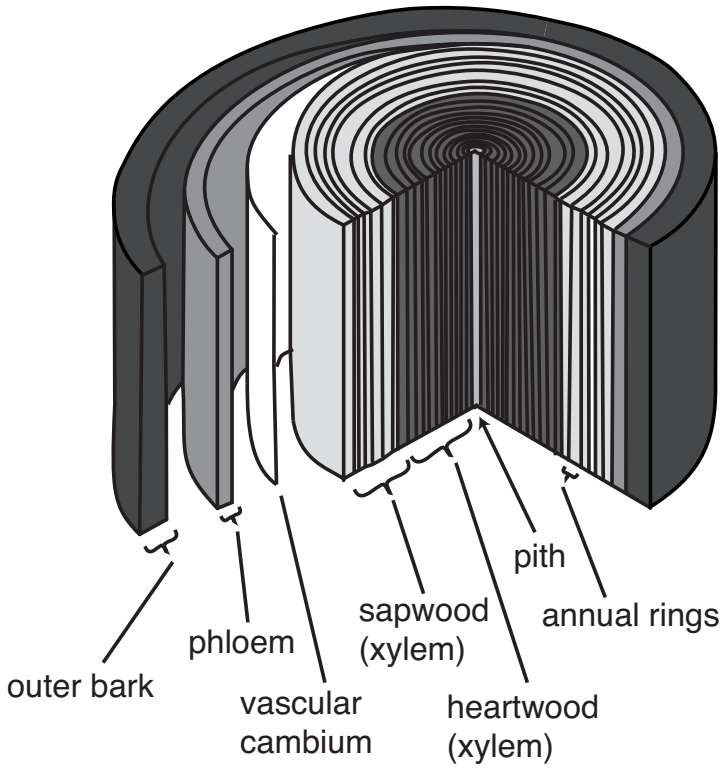

Figure 1: *Schematic picture describing the major tissues in a tree stem. Modified from Shmulsky and Jones (2011), with permission*

1.2.2 Wood Anatomy

The major cell types in hardwood xylem are fibres, vessel elements and ray cells (*Figure 2*). Fibres and vessel elements are axially elongated cell types. Fibres are the most abundant cell type and fulfill mostly a structural function, whereas vessel elements are transporting water and nutriments from the root. Both fibres and vessel elements undergo programmed cell death after deposition of the secondary cell wall. The radial dimension of the developing fibres and vessels zone is within the range of hundreds of micrometers to millimeter, so the major part of wood in trees is composed of dead cells. Ray cells are elongated in radial direction and function in radial transport of nutriments from the xylem and phloem streams. In contrast to fibres and vessel elements, ray cells stay alive for many years but finally undergo cell death at the border to the heartwood (Murakami et al., 1999). They are also involved in depositing the extractives of the heartwood.

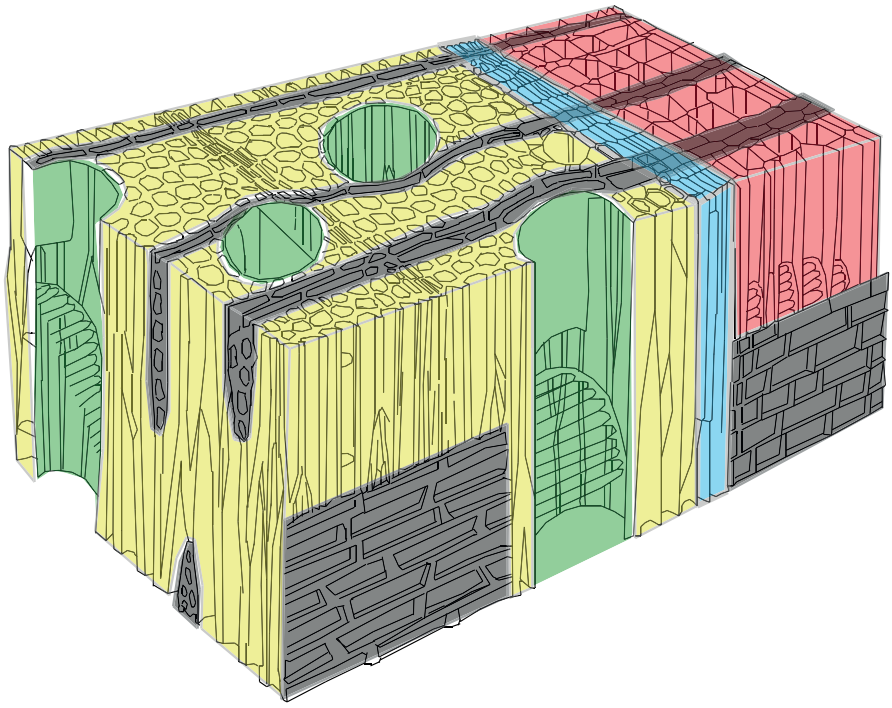

Figure 2: *Schematic picture showing the different cell types in hardwood: Fibres in yellow, ray cells in grey, cambium in blue, vessels in green and in red the phloem with sieve cells. Modified, from Mägdefrau (1951).*

1.2.3 Wood Chemistry

The underlying work of this thesis was mostly conducted on aspen wood. The following description of wood chemistry is therefore based on hardwood from angiosperms, mostly from *Populus* wood.

The chemical composition of wood can be indicated in many different ways. The elemental composition, for example, depicts the percentages of carbon, hydrogen, oxygen, nitrogen and sulphur in a sample and includes no information about structural aspects (*Table 1*). Summing up the yields from extracting several chemically distinct polymer fractions gives already some insight into the structure of wood. There are a large number of different protocols to extract sub-fractions from wood (Browning, 1967; Fengel and Wegener, 1989; Hon and Shiraishi, 2001). The origin and consequences of this variety will be discussed in more depth in *section 1.4*. A widely accepted nomenclature to define the distinct chemical fractions of wood are the terms 'cellulose', 'cross-linking glycans' (or 'matrix polysaccharides',

'hemicelluloses'), 'lignin', 'pectins', 'extractives', 'ashes' and 'structural wall proteins'.

Table 1: *Elemental composition of hybrid poplar wood according to Huang et al. (2009).*

Carbon	50.02 %
Hydrogen	6.28 %
Nitrogen	0.19 %
Oxygen	42.17 %
Sulphur	0.02 %

Table 2: *Chemical composition of* Populus tremuloides *composed from Panshin and de Zeeuw (1980, p. 92) and Fengel and Wegener (1989, p. 57)*

Cellulose	40 %
Cross–linking Glycans	31 %
Lignin	20 %
Pectin	3 %
Extractives	7 %
Ashes	>1 %

The distribution between the chemical fractions shown in *Table 2* are reference values with room for substantial variation. While the chemical composition at various heights of a tree stem is rather constant, substantial variation can occur in–between growth rings (Sykes et al., 2008). Further, there is also variation among different species and hybrids of aspen as shown in Sannigrahi et al. (2010). The abundance of 'structural wall proteins' in wood is very low, according to Lüdtke and Liefländer (1970) between 0.068 % and 0.21 %. They are therefore not accounted for in *Table 2*.

Below follows a short description of the various cell wall constituents, with the aim to give an overview about molecular composition, intra– as well as extra–molecular structure, physico–chemical properties and a short account for their biosynthesis.

Cellulose, the largest and most abundant macromolecular assembly of wood, is made of long β-(1 $\rightarrow$ 4)-*D*- glucose chains. The orientation of each glucose molecule is 180° rotated in respect to both its neighbouring

molecules. As such, the actual repeating unit of cellulose is the disaccharide cellobiose. The degree of polymerisation of each such glucose chain has been determined to an average of 10'000 for *Populus* wood (Fengel and Wegener, 1989, p. 75).

During biosynthesis a distinct number of elementary glucan chains are bundled together by hydrogen bonding into larger aggregates called cellulose microfibrils (CMF) (Ding and Himmel, 2006; Fernandes et al., 2011). Such microfibrils as well as further aggregations of multiple microfibrils comprise both amorph and crystalline zones while all other cell wall polymers are exclusively amorph. Two important properties resulting from this fact are the sizes of individual crystallites and the degree of crystalline regions expressed in percentage of the total cellulose fraction (Zugenmaier, 2008, p. 207–212). Cellulose is hydrophilic and can swell substantially when it absorbs water (Zugenmaier, 2008, p. 213). The enzymatic machinery for cellulose synthesis is thought to be a multi-subunit complex located in the plasma membrane, including a number of cellulose synthases (CesA) as the catalytic subunits (Mutwil et al., 2008; Kumar et al., 2009).

The major types of **cross–linking glycans** in *Populus* wood are glucuronoxylan, glucomannan and xyloglucan. Glucuronoxylan, the most abundant cross–linking glycan in *Populus* consists of a β-(1 $\rightarrow$ 4)-*D*- xylose backbone chain. About 70 % of the backbone xylosyl units are acetylated in the C–2 or C–3 position. About 10 % of the backbone xylosyl units are decorated through a (1 $\rightarrow$ 2)-linkage with 4–O–methyl–α–D–glucuronosyl acid residues (Sjöström, 1993, p.67–68). The reducing end has a well conserved pattern including rhamnnose and galacturonic acid (Fengel and Wegener, 1989, p. 110) (Peña et al., 2007).

Glucomannan is the second most abundant cross–linking glycan in *Populus* wood with a mannose:glucose ratio of about 1:1 – 1:2 (Sjöström, 1993, p.68–69) (Fengel and Wegener, 1989, p.115)(Kim and Daniel, 2012). It is largely unbranched. Xyloglucan (Simson and Timell, 1978) consists of a β-(1 $\rightarrow$ 4)-*D*- glucosyl backbone where about 70% of the C–6 positions are decorated with α-*D*- xylosyl units. Those side-chains can be several sugar units long incorporating also galactosyl- and fucosyl units (Simson and Timell, 1978).

The physico–chemical properties of cross–linking glycans vary according to their molecular composition. However, they are all of hydrophilic character and can therefore form hydrogen bonds. The glucuronic–acid branchings of the glucuronoxylan contribute due to their high abundance to a large extent to the chemical environment in *Populus* cell walls. While a large number of enzymes involved in biosynthesis of cross–linking gly-

cans have been identified and located to the golgi apparatus, it is today still largely unknown how the enzymes work together to build the well defined and complex polysaccharide structures (Scheller and Ulvskov, 2010). After initial assembly, the polymers are packed in vesicles and transported to the growing cell wall.

Lignin is a three dimensional–linked hetero-polymer consisting of phenylpropanoic building blocks, the monolignols. In angiosperm trees, three types of monolignols are incorporated into the lignin polymer: p-coumaryl–, coniferyl–, and sinapylalcohol (Boerjan et al., 2003). The commonly used acronyms H, G and S type lignin refer to the first letters of p-**H**ydroxyphenyl, **G**uaiacyl and **S**yringyl lignin units (Lapierre, 2010, p. 12). The three monolignols differ in the degree of methoxylation of their phenyl group.

Polymerisation of the macromolecular structure of lignin takes place *in muro*. The actual deposition process of lignin is called 'end–wise radical coupling' and opposed to the other cell wall polymers, it is not under direct enzymatic control (Ralph et al., 1999; Hatfield and Vermerris, 2001). It complies to chemical equilibrium rules where the abundance of monolignols and catalytic radicals determine the rate of reaction. As such, the abundance of different types of bonds that are formed in a lignin polymer can be explained with the reactivities of the respective atoms according to the rules of organic chemistry (Syrjänen and Brunow, 2001). Small deviations from such theoretical values can occur due to a variety of conditions in the plant cell wall such as pH, oxidation potential, relative concentration of the reactants, steric constraints and also chemical bonds to other cell wall constituents (Brunow and Lundqvist, 2010, p. 270). A commonly cited characteristic of the lignin polymer is the ratio between the different monolignol building blocks incorporated into the polymer ('S/G ratio'). This ratio is of interest because different types of monolignols preferentially form different intramolecular bonds which have different susceptibilities towards degradation (Lapierre et al., 1999; Huntley et al., 2003; Studer et al., 2011). As there is no conserved sequence or larger consistently conserved three–dimensional polymeric structure in lignin, it is best characterised chemically by the abundance of the different ether and carbon–carbon bonds (Dimmel, 2010). The main bonding patterns are shown in *Figure 3*.

4–O–β ether bonds are the most abundant lignin bonds in hardwoods and make up about 60% of all bonds. From the naturally occurring carbon-carbon bonds, those involving the aryl C–5 position are of special interest as they are significantly less reactive than for example ether bonds and therefore more resistant to degradation. This specific structure is also called condensed lignin (Dimmel, 2010, p. 3). The aryl C–5 carbon is only avail-

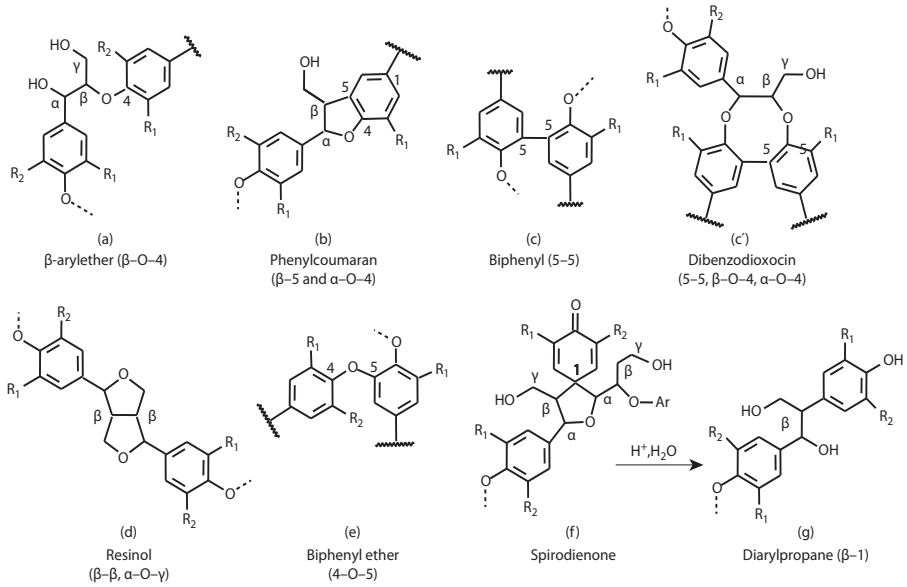

Figure 3: *Main bonding patterns of native lignins ($R_1 = R_2 = H$ in p-hydroxyphenyl units; $R_1 = OMe$, $R_2 = H$ in guaiacyl units, $R_1 = R_2 = OMe$ in syringyl units). With permission, adapted from Ralph et al. (2004).*

able for bonding in guaiacyl and p–hydroxyphenyl monolignols. The latter are of rather low abundance in hardwoods, whereas guaiacyl monolignols constitute about 30–40% of the total lignin. Except a few percent of p–hydroxyphenyl type, the rest is syringyl type lignin. The phenyl group makes the lignin hydrophobic, which is in contrast to the other cell wall polymers. The degree of polymerisation of native lignin polymer is hard to determine and widely different values have been published depending on the method used. According Sjöström (1993), hardwood lignins have usually a degree of polymerisation below 20'000 and rather high polydispersities of 2.5 to 3.

Pectin is a collective name for a number of polymers where some have very complex structures (Caffall and Mohnen, 2009). Generally, pectins are of rather low abundance in wood. The main pectic polysaccharide in *Populus* wood is rhamnogalacturonan, a (1 → 4)-β-*D*- galacturonosyl -(1 → 2)-*L*-rhamnosyl main-chain with (1 → 4)-β-*D*- galactan and (1 → 5)-α-*L*- arabinan branchings at the rhamnosyl C4 position. Some of them carry an arabinose or fucose terminal group (Simson and Timell, 1978). This polymer is often also referred to as Rhamnogalacturonan–I. Rhamnogalacturonan–II (RG–II), a pectic polysaccharide with a similar backbone but much more

complex branching structures, is found and described in many other model species including *Arabidopsis thaliana*. There is today little or no literature about RG-II in *Populus*. Due to the rather complex structure and expected low abundance, it has probably been left out so far from systematic investigations in this species.

Biosynthesis of pectins occurs in the golgi apparatus. Upon initial assembly, for vesicle transport to the wall, most galacturonosyl groups of the backbone are methyl esterified. After deposition in the wall, pectin methylesterases actively modify the chemical properties of the pectic polymers (Cosgrove, 2005). The galacturonic residues are of importance not just to regulate the pH in the cell wall but also for building calcium bridges to other pectic polymers.

A large variety of **extractives** exists in the wood of angiosperm trees. Functionally, extractives are not part of the load bearing cell wall structure. Many of them are involved in pathogen resistance, active defense or plant–insect interactions (Scheffer, 1966; Ralph, 2009). Chemically, extractives are almost exclusively of hydrophobic character. Inclusion of extractives, therefore modifies the cell wall chemistry towards a more aseptic environment. The major extractives of *Populus* wood are fatty acids such as linoleic acid and oleic acid, wax acids, glycerol, wax alcohols, paraffin waxes and sterols such as sitosterol (Rowe and Connor, 1979). A large range of small phenolic and poly-phenolic molecules are another major class of extractives.

Ashes is the inorganic fraction that remains after complete combustion of the cell wall. *Populus* wood contains about 0.4 % of ashes (Fengel and Wegener, 1989, p. 57). Calcium is the by far most abundant inorganic element in wood. In contrast to most other inorganic elements in plants, it is structurally involved in cell walls by connecting two chains of the pectin polysaccharide homogalacturonan through ionic bonds (McNeil et al., 1984). A similar function has also the much less abundant element boron, which builds borate diester cross-links between the apiose residues of two RG–II chains (O'Neill et al., 2004). The other highly abundant inorganic elements, potassium and magnesium, are not know to have a structural function in the cell wall.

The most prominent **Structural wall proteins** in wood are the APG's which are built of a protein part and a carbohydrate part with substantial complexity (Seifert and Roberts, 2007). Not much data on chemical quantitation and characterisation of Arabinogalactan proteins (AGP) in *Populus* is available. However, studies elucidating the genes involved in cell wall biosynthesis suggest that AGPs are present in *Populus* (Andersson-Gunneråsa et al., 2006; Geisler-Lee et al., 2006).

1.2.4 Ultrastructure and Cell Wall Architecture

The ultrastructure of the fibre cell wall is to a large degree determined by the cellulose polymer because the cellulose microfibrils extend over a much longer distance than any other cell wall polymer. Further, they are the first polymer that is laid down in the growing cell wall, they are usually well ordered and function as a scaffold for the other polymers. The angle between the longitudinal axis of a fibre and the direction of the cellulose microfibril is called micro fibril angle (MFA) and is an important parameter to describe the ultrastructure of a cell wall (Donaldson, 2008)(*Figure 4*). The main ultrastructural feature of a fibre wall that is visible by electron microscopy on transverse sections are concentric layers with different radial dimensions. These layers reflect different orientations of the cellulose microfibrils. The outermost layer is the primary cell wall (P), the three inner layers consist of the secondary wall and are denoted S1, S2 and S3 (*Figure 4*).

The primary cell wall of a fully developed fibre has no uniform orientation of the cellulose microfibrils. The criss-cross structure is a result of a high MFA and 'spring' like reorientation of the microfibrils during cell expansion (Burgert and Fratzl, 2009). The MFA of the S1 and S3 layers is rather high, while S2 has a much lower MFA (Donaldson, 2008). The intercellular layer, or middle lamella (ML), is the volume between two adjacent primary cell walls. There is usually no clear border between the middle lamella and the primary walls, therefore it is often described as compound middle lamella (CML).

While the ultrastructural features of cell wall layers are on a micro-scale, many cell wall architectural details are on a nano-scale. The cell wall architecture is determined by the abundance of the cell wall polymers in different morphological regions (cellwall–layers, –corner, –tip, tangential– and radial walls), their orientation and linkage to neighbouring polymers. To describe the structural architecture of cell walls, Booker and Sell (1998) used the analogy of reinforced concrete. The partly crystalline cellulose fibrils make up a scaffold similar to iron rods, cross–linking glycans make the connections between the cellulose fibrils, similar to cement and lignin compares to the sand as a filling material in the afore mentioned analogy. In primary cell walls of *Populus*, xyloglucan is the most abundant cross–linking glycan, whereas glucuronoxylan, the most abundant of the secondary cell wall. Pectin polymers are highly abundant in the middle lamella, but also in the primary walls. Investigation on the orientation and interlinkage of cell wall polymers is an active field of research and many questions are still unanswered. Generally, more is known about the primary cell wall, often from studies in *Arabidopsis*.

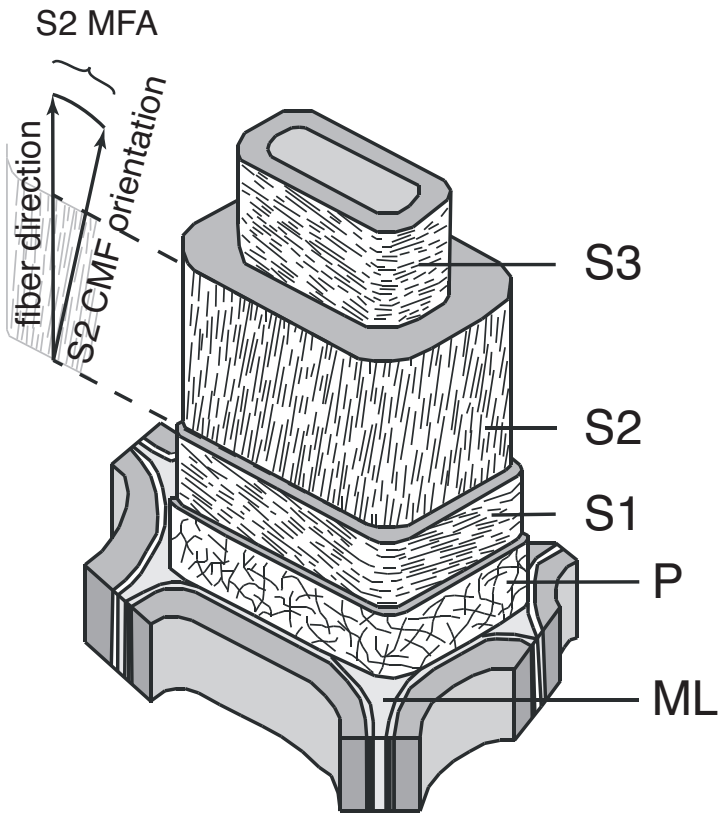

Figure 4: *Ultrastructural features of a fibre: Four concentric layers with distinct cellulose microfibril (CMF) orientations. The microfibrils of the primary (P) wall are criss–crossed, while the secondary (S) wall layers show distinct microfibril angles (MFA). The middle lamella (ML) is the space between primary walls. Modified from Shmulsky and Jones (2011), with permission.*

A schematic model of the primary cell wall before lignification, referred to as 'sticky network model' is shown in *Figure 5*. In this model, stretches of xyloglucan chains are hydrogen bonded on the surface of cellulose microfibrils. The majority of xyloglucan chains are bonded to more than one cellulose microfibril, hence they connect them loosely to each other. Some xyloglucan chain ends are also thought to be attached to cellulose microfibrils by being buried on the inside of them during biosynthesis (Cosgrove, 2000). In another model, the 'multicoat model', cross–linking glycans are suggested to coat individual cellulose microfibrils without linking them together (Talbott and Ray, 1992). Based on recent structural analysis by Fer-

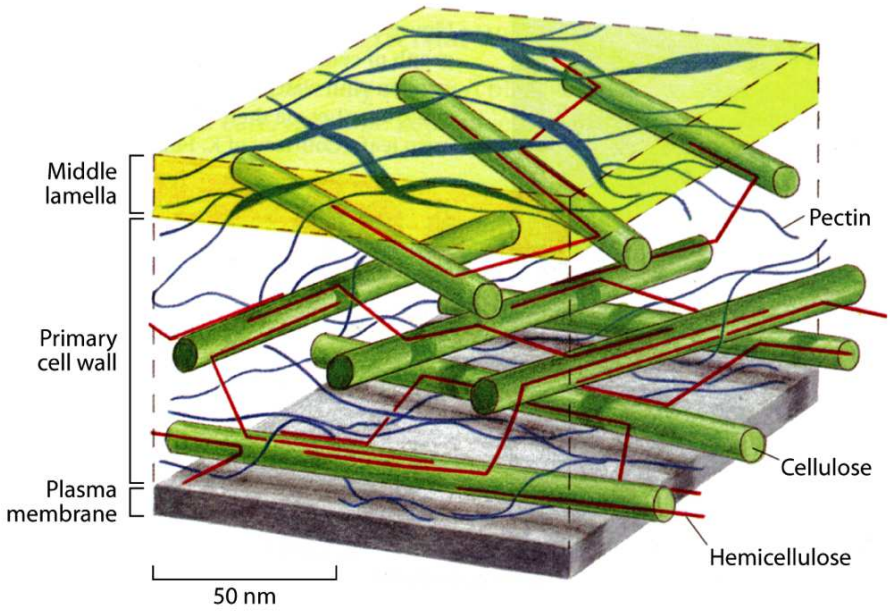

Figure 5: *The schematic drawing is a model of the primary cell walls' architecture. This model is referred to as the 'sticky network model', which is one of several. The main distinguishing feature to other models is the direct connection of several microfibrils by the cross–linking glycan xyloglucan. From Scheller and Ulvskov (2010), with permission.*

nandes et al. (2011) of the secondary cell wall in spruce fibers a 'multicoat model' is favoured.

The lignin polymer forms within the matrix of the existing wall structure. When interpreting literature data regarding the spatial distribution of different cell wall polymers to specific morphological regions it is important to note that values also vary depending on the developmental stage (Terashima, 1990). For example, during formation, the primary cell wall does not contain notable amounts of lignin. However, at a much later developmental stage, the primary cell wall is not excluded from the lignification process. Indeed the lignification has been even found to start in the middle lamella of the cell corners where also the lignin concentration was reported to be the highest (Ma et al., 2011). From there, lignification is thought to spread by end wise radical polymerisation throughout the whole cell wall. It has been described that the aromatic rings of the lignin polymer align parallel to the cell wall (Atalla and Agarwal, 1985; Åkerholm and

Salmén, 2005). Terashima (1990) proposed that the anisotropic shrinkage, when the swollen hydrophilic cell wall becomes hydrophobic due to deposition of lignin, causes the phenyl rings to align parallel to the cell wall. Other research groups proposed that non–covalent interactions with the cellulose network cause this alignment (Siegel, 1957; Houtman and Atalla, 1995; Roussel and Lim, 1995; Li and Eriksson, 2005). Since a long time there has been a discussion about the presence of covalent bonds between lignin entities and several other cell wall polymers, so called LCC's (Merewether, 1957; Lai and Sarkanen, 1971; Koshijima and Watanabe, 2003). Such bonds seem to have a large influence on the degradability of cell walls for example in pulping processes (Chakar and Ragauskas, 2004; Henriksson et al., 2007; Dimmel and Gellerstedt, 2010). But due to their low abundance, not much is known about their formation, chemical structure and distribution throughout the cell wall layers.

The primary cell wall is enzymatically modified *in muro* after its formation to facilitate cell expansion (Cosgrove, 2005). The 'pectin methyl esterase' for example, removes the methyl esterification of homogalacturonan pectins when freshly deposited into the cell wall (Wolf et al., 2009). The formed weak acid is prone to constitute ionic calcium bridges with other acidic pectin groups. Other enzymes that actively modify deposited cell wall polymers are 'Xyloglucan endotransglucosyslase/hydrolase' and a number of endo-glucanases such as for example 'Korrigan' (Molhoj et al., 2002). It has also been proposed that the secondary wall is enzymatically modified after its formation. A recent idea proposes that endoglucanases chop cross–linking glycans that are trapped in the so called gelatinous layer of tension wood fibres in angiosperm trees. This would make the gelatinous layer swell and fulfill its function as the force generator in the tension wood response (Mellerowicz and Gorshkova, 2012).

1.3 Bulk Properties

Bulk properties of wood refer to parameters that can be measured on a macro, and sometimes on a micro scale. They are therefore also dependent on the nano– and ultrastructure of the fibre cell wall. **Density**, or 'mass divided by volume' can be assessed at different levels of organisation in wood. Usually, bulk density refers to the density of a piece of wood and is based on dry mass and wet volume. It is mainly dependent on volume and density of the cell wall material. The volume of cell wall material per volume wood is defined by the wall thickness and the volume of the lumen. In hardwood, fibers are the most important determinant for density, but vessels and ray cells will also have an influence.

The density of the cell wall material *per se*, however, is for the sake of simplicity usually assumed to be constant (1.056 $g\,cm^{-3}$) (Kollmann, 1951). When working with transgenic plant material, where the range of chemical composition and structure can vary beyond the natural variation, this assumption may not hold true. It is therefore in certain cases relevant to determine the density of the actual cell wall material in order to fully understand the effect of the transgene. The density of cell wall material depends on its chemical composition, structure and the resulting porosity. When the density of actual cell wall material shall be determined, precise anatomical measures of the sample material are needed. Porosity, can be measured for example by cryoporometry (Petrov and Furo, 2009).

Mechanical properties of solid wood, wood sections or even single fibres are bulk properties that describe the mechanical behaviour of the tested specimen. Because wood and fibres are anisotropic, mechanical properties are dependent on the measuring direction (Salmén and Burgert, 2008). Commonly determined parameters for wood sections or single fibers are 'stiffness', and 'ultimate stress'. In a typical experiment, an increasing force is applied to a specimen in a defined direction and axis until it breaks. Applied force and deformation are constantly measured and represented in a 'stress–strain' diagram. A typical such diagram for wood is presented in *Figure 6*.

Wood has generally a very high 'tensile strength' in fibre direction and can withstand substantial forces until the point of 'ultimate stress' is reached. Even the 'compressive strength' in fiber direction is rather high. These properties are mostly a consequence of the cellulose ultrastructure. The strength properties perpendicular to the fiber direction are in contrast rather poor. Wood withstands in this direction just low tensile or compressive stresses (Bodig and Jayne, 1993).

1.4 Chemical Characterisation of Wood

A large number of different methods and techniques to chemically characterise wood and plant cell walls are described in literature. Among them, there is a basic tool-set of more readily used methods in plant cell wall research that mostly aims at characterising a range of different sample materials in a comparable way. There are also a number of advanced and specific methods that can be used to further explore details about the cell wall. Often those methods require tedious sample preparation and specifically trained researchers.

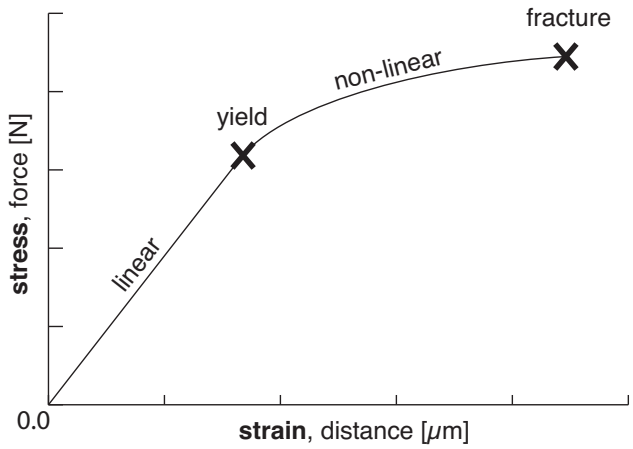

Figure 6: *Typical stress–strain curve for micro-mechanical tensile strength testing of thin wood sections as presented in Paper III. The slope of the linear portion determines the property 'stiffness' or 'elasticity' of the tested specimen. The sample deformation in this region fully recovers after stress release. In contrast, the deformation after the yield point, in the non-linear part of the curve, is permanent. This property is called sample 'plasticity'. The force at which the specimen fractures is the point of 'ultimate stress'.*

1.4.1 The Basic Tool-set

Before describing in more detail a number of analytical methods, some important terms from analytical chemistry, taken from Harris (2007), are introduced: *Accuracy*, a measure of how close a measured value is to the 'true' value. *Precision*, a measure of the reproducibility of a measurement. *Robustness*, ability of an analytical method not to be affected by small deliberate changes in operating conditions.

The accuracy for most of the basic methods, depends on specificity of chemical reactions: The different bonds within and between wood polymers are of rather similar character, it is therefore difficult to find reactions that exclusively isolate a specific wall component. The type of reactions and reaction conditions used are usually optimised for a specific type of sample. If the robustness of a method has been assessed, it should include how much the accuracy of the method varies when applied to a large array of different sample types. However, this information is usually not available and therefore the comparison of analytical results between species and tissue types should be met with caution.

Chemical characterisation of wood can be done at three different lev-

els. First the main functional fractions of wood (lignin, cross–linking glycans, cellulose and extractives) are quantified. In respect to wood analysis, pectins are usually included without mentioning in the cross–linking glycan fraction as they are low abundant and the strategy for analysis is the same. The second level is the compositional analysis of individual wood polymers. And finally, the structural determination of wood polymers can be done.

Quantification of wood polymers Initially, suitable portions of woody tissue have to be sampled, dried, ground and sieved. Pulp and paper industry's standardised procedures (http://www.tappi.org, http://www.iso.org) for those tasks are a good choice when dealing with bulk material. If just milligrams of sample material are available, substantial adaption of mills, sieves and general glassware is usually required. Determination of the first fraction, the extractives, is done gravimetric by weighing prior and after extraction. A generic extraction procedure uses the soxhlet extractor with three steps of extraction, 'ethanol–toluene', 'ethanol' and 'water' (Ona et al., 1995). The resulting sample material is called 'extractive free wood'. Determination and quantitation of the various extractives is usually not a standard experiment in laboratories that work mostly on structural features of plant cell walls.

In a next step, the lignin content is determined. This is done on a small subfraction of the 'extractive free wood'. Two commonly used procedures are the 'Klason lignin' determination (http://www.tappi.org, T222 om-88) and the 'Acetyl-Bromide' method (Fukushima and Hatfield, 2004). The 'Klason lignin' determination is done gravimetric after all carbohydrate polymers are hydrolyzed and solubilised. In the 'acetyl bromide' method, the lignin polymer is dissolved using acetyl-bromide in glacial acetic acid. Then quantitation is done by measuring the absorption at 280 nm.

For cellulose determination, the 'extractive free wood' has to be delignified. The most common method is by sodium chlorite where a number of reactive chlor species are formed in solution that specifically degrade the lignin polymer (Ona et al., 1995). The sample after lignin extraction is called 'holocellulose'. In the next step, the cross–linking glycans are degraded by 'alkali dissolution', while the much more resistant cellulose is dissolved as intact macromolecules and determined gravimetric. Cellulose determined with this method is usually termed 'α-cellulose'. Cross–linking glycans can then be calculated as mass closure. Another standard protocol to determine cellulose content is the 'Updegraff' method (Updegraff, 1969). Here, all polymers except cellulose are degraded by a mixture of acetic– and concentrated nitric acid. Finally, the total sugar con-

tent of the non-dissolved portion is measured photospectrometrically by the phenol-sulphuric acid method from 'Dubois' (Dubois et al., 1956; Masuko et al., 2005). Direct quantitative methods to determine the content of cross–linking glycans are scarce. Sometimes, the values from molecular composition determination, described in the following sections, can be used.

Molecular composition of individual wood polymers The second step in basic chemical characterisation of wood is determination of the molecular composition of lignin and the cross–linking glycans. Whereas there are two readily used standard procedures for lignin composition, analysis of cross–linking glycans can be taken to many levels of sophistication. This is due to the large variety of different cross–linking glycans and pectins as well as their well conserved complex structures.

The most popular procedures for lignin composition analysis are 'thioacidolysis' (Rolando et al., 1992) and 'derivatisation followed by reductive cleavage' (DFRC) (Lu and Ralph, 1997). In both methods, wood powder is dissolved (in the case of DFRC after bromination derivatisation to improve solubility), before the lignin polymer is degraded by specific chemical reactions, that will free lignin monomers without degrading them. In the case of 'thioacidolysis' this is done with an acid–catalyzed reaction specific for the highly abundant $\beta - O - 4$ bonds. In the 'DFRC' method, cleavage of the same bond type is obtained by reductive cleavage using zinc dust. Finally, the dissolved monomers are derivatised for separation and analysis by GC/MS. The analytical result will give a good estimation on the abundance of the three different monolignols in the lignin polymer. With slight modifications, the same methods can also be used to detect and quantify a number of oligo–lignols (Peng et al., 1998).

For the molecular composition determination of cross–linking glycans a whole range of methods are readily applied. In fact, it is more like a modular system of different experimental steps that are put together depending on the problem to be solved. All analyses consist however of three steps where the first step is to obtain the starting material. In the simplest case, this is either extractive free wood or holocellulose and the major fraction of abundant cross–linking glycans and pectins can be analysed without further fractionation. It is also possible to fractionate the sample to obtain for example samples enriched in pectins or acidic cross–linking glycans etc. Such procedures are tedious and vary substantial depending on the used sample material. The second step is the depolymerisation of the polysaccharides. A number of different protocols, using a range of different acids are described

in literature. The most common are in aqueous solvent TFA, sulphuric acid (Saeman hydrolysis), or in dry methanol, by hydrochloric acid (methanolysis) (Biermann and McGinnis, 1989). While the first two are classic hydrolytic cleavage reactions, the third is based on solvolysis. After depolymerisation, the samples consist of free monosaccharides, which in solution freely 'mutarotate' between the $\alpha-$ and $\beta-$ and the 'furanose' and 'pyranose' conformations. The last step is the analytical determination. The two most common techniques here are GC/MS and HPAEC-PAD. For GC/MS two popular protocols exist. In the determination as 'alditol acetates', the monosaccharides are first reduced with sodium borohydride to alditols, then derivatised by acetylation and separated on a rather polar column. Here, the reduction step will yield one peak for each sugar. This method is compatible with TFA and sulphuric acid hydrolysis (Biermann and McGinnis, 1989). Methanolysis, when analysed by gas–chromatography, is usually combined with tri-methyl silylation derivatisation and separation on a non-polar column (Biermann and McGinnis, 1989). Here, each monosaccharide can result in up to four peaks due to the above mentioned 'mutarotation' of the molecules in solution. An advantage of the 'methanolysis' procedure is the ability to analyse also uronic acids, while the acidic group becomes reduced and the information lost in the reduction step of the 'alditol acetates' procedure. A general strength of analysis by GC/MS is the possibility for analyte identification by the mass spectra and retention time comparison with standards. For analysis by HPAEC–PAD, no derivatisation is needed. Here, each monosaccharide will result in one peak. Analyte identification is done purely by retention time. Chromatographic separation of the monosaccharides common to wood polymer by HPAEC–PAD is a bit slower than with GC. It has also been reported, that the separation of certain monosaccharides can be problematic (Perry and Currie, 2006). Results from carbohydrate analysis can be reported as weight or mol per weight, respectively normalised to weight or mol percentage. The former can give information on the depolymerisation efficiency, which for small wood samples is rather common to deviate substantially from 100%. Due to this problem, the mol % visualisation will give more reliable values on the molecular composition of the hydrolyzed polymers.

Sequence and structure of wood polymers As a third step in basic chemical characterisation, the sequence and structure of cell wall polymers can be determined. For lignin a detailed quantification of all abundant bonding structures in the lignin polymer is readily done by solution state 2D NMR experiments (Lu and Ralph, 2011). However, data processing and analysis

is tedious and needs NMR spectroscopists' expert knowledge.

For complete structural elucidation of carbohydrate polymers a sequence or linkage analysis of the polymer is done. A commonly used method is to methylate all free $-OH$ groups prior to TFA hydrolysis (Hakamori, 1964; Ciucanu and Kerek, 1984). After methylation, the reduction step is conducted with deuterated sodium borohydride to conserve the linkage information. Then, the samples are derivatised by acetylation and analysed by GC/MS. A detailed list with abundances of different linkage types is the analytical result obtained by this method.

The actual structural information can only be determined from oligo saccharides obtained by specific enzymes. Two different types of mass–spectrometers are popular for the structural determination of oligo–saccharides: MALDI-TOF or ESI-IT-MS (Zaia, 2004). Both methods result in mass spectra that allow interpretation of the oligo–saccharides actual structure.

1.5 Wood Formation and Functional Genomics

Gene expression analysis and functional genomics research have revealed many of the genes and enzymes that are involved in cell wall biosynthesis (e.g. Mouille et al., 2003; Bhalerao et al., 2003; Aspeborg et al., 2005; Boerjan, 2005; Brown et al., 2005; Andersson-Gunneråas et al., 2006; Mitchell et al., 2007; Yang et al., 2011). Establishing of the model organism *Arabidopsis* was an important premise towards standardising methods and resources for functional genomics among plant researchers around the world. For research on woody species, the early availability of the *Populus* genome sequence and secondary cell wall transcriptome databases were the departure for successful research (Tuskan et al., 2006; Sterky et al., 2004; Hertzberg et al., 2001).

High–throughput methodologies for chemotyping lignocellulose are becoming increasingly critical for large–scale approaches in functional genomics involving screening of natural or mutant populations. Moreover, there is often a need to couple high–throughput analysis with the capability to handle small sample amounts. For these purposes wet chemistry approaches are less suitable because they are in general tedious, require large sample amounts, and are difficult to automatise. Some recent developments have, however, been presented for high–throughput analysis of lignin and non–cellulosic carbohydrates (Tuskan et al., 1999; Lerouxel et al., 2002; Mouille et al., 2003; Bauer, 2012).

As an alternative to isolate and analyse individual wall components, the whole cell wall fraction can be analysed without prior extraction and separation. This approach provides a chemical fingerprint (chemotype) of the

sample, and is particularly advantageous in studies where an overall picture of cell wall composition is wanted. There are several techniques that can provide cell wall chemotypes. One possibility is to dissolve the whole finely ground sample with appropriate solvents, followed by 2D solution state NMR spectroscopy (Lu and Ralph, 2003, 2011). This approach provides as output complex spectral maps where individual signals can be assigned to their origin cell wall polymers. Due to the complexity of the data, it is preferentially analysed and classified by multivariate data analysis prior to tedious manual spectral assignments (Hedenström et al., 2009). However, required sample amount is in the range of 50 to 100 mg, the sample preparation and analysis is not truly high-throughput and quantitative analysis is difficult due to many spectral overlaps in a complex two-dimensional dataset.

Another possibility to obtain cell wall chemotypes is by transmittance FTIR spectroscopy. Here, analysis of powdered samples does not require laborious sample preparation beyond the milling procedure, and the rather quick acquisition time allows throughputs of several hundred samples per day. Also, FTIR fingerprints are rich in information, but due to the overlap of characteristic bands from different polymers they do not translate so easily into chemical structure and composition. Therefore quantitative estimates will be inaccurate (Gorzsás et al., 2011). However, chemical fingerprints obtained by FTIR are well suited for sample classification (Brown et al., 2005; Mouille et al., 2003).

A third technique that has been demonstrated to be useful for high-throughput analysis of complete cell wall samples is Py-MBMS. Here, pyrolysis of powderised samples is used for reproducible sample fractionation prior to mass spectrometric detection. The degradation products were shown to be specific for the different types of lignin as well as pentose and hexose sugars (Tuskan et al., 1999; Sykes et al., 2008; Studer et al., 2011). Thus, this approach provides quantitative estimates of the main cell wall polymers, but it is rather insensitive for sample classification due to the lack of a chromatographic step.

A chemical fingerprinting method that has been demonstrated mostly for primary cell wall tissues with a focus on cross–linking glycans and pectins employs matrix assisted laser desorption ionisation as sample introduction to mass spectrometry (Obel et al., 2009). This specific method has sometimes been called oligomer mass profiling (OLIMP) (Lerouxel et al., 2002). The method employs an enzymatic pre–treatment directly on the tissue of interest, which is then mounted on the sample stage of a MALDI–TOF instrument. Prior to analysis the samples are coated with MALDI matrix.

The resulting data is dependent on the efficiency of employed enzymes.

2 Objectives

The objectives of the work described in this thesis were to develop and evaluate Py-GC/MS, together with other cell wall analysis approaches, as a technique to be used in a wide range of analytical scenarios in functional genomics research. Such applications could include compositional and/or structural cell wall analysis, high-throughput chemotyping and micro-analysis of woody tissues for different sample types. After several initial experiments it became obvious that an automated data processing pipeline would be needed to achieve the main objective. A toolbox of data processing and evaluation functions were therefore also included in the list of objectives. The tools developed was used in interdisciplinary projects together with plant biologists and material scientists to explore the function of several genes important in wood formation.

3 Methodological Overview

3.1 Plant Material and Sample Preparation

Samples from *Arabidopsis thaliana* and aspen (*Populus tremula x tremuloides*) were used in a majority of the experiments. For the study of secondary cell walls in *Arabidopsis* plants, either the hypocotyl or the inflorescence stem was used. Samples were grown under standard conditions in the growth chamber. Aspen trees were grown in the greenhouse according to an experimental design that minimises environmental effects (Pinto et al., 2011).

Depending on the specific type of sample, slightly adapted sample preparation protocols were employed. The generic sample preparation for Py-GC/MS analysis consists of freeze drying the frozen material and then either grinding or excising the tissue to be investigated. In most cases, the samples were not extracted prior to analysis. If the application demanded extraction, standard procedures such as preparation of AIR (Coimbra et al., 1996) in *Arabidopsis* or toluene:ethanol soxhlet extraction (Ona et al., 1995) for aspen wood were employed. To grind *Arabidopsis* hypocotyls and stems, 1.5 ml cups with one 7 mm ball were used in a vibratory ball mill (MM 400, Retsch, Germany) at 30 Hz for 2 min. Small sample amounts of aspen wood were ground in the same way. Larger wood samples were ground in 10 ml containers with either one 12 mm ball for 2 min at 30 Hz or with two 12 mm balls at 30 Hz for 70 s. If wet chemical procedures were applied to

a wood sample, it was initially disintegrated by a centrifugal mill (ZM 200, Retsch, Germany) with a 0.25 mm cutting sieve. Powder was then produced as described above for small sample amounts. All ball–milling procedures used stainless steel containers and grinding balls.

3.2 Pyrolysis–Gas Chromatography / Mass-spectrometry

Development and application of an analytical pyrolysis method for plant research was a major effort of this thesis. Therefore a short account for pyrolysis as an analytical technique in general is presented here. Pyrolysis is the disintegration of molecules by thermal energy only, while analytical pyrolysis is the technique where the resulting fragments of pyrolytic degradation are studied to gain information about the initial intact molecules (Wampler, 2007, pp. 1–2). The most common applied conditions for analytical pyrolysis are an inert atmosphere and temperatures in the range between 300 °C–800 °C. The pyrolytic process is complex and consists of a whole chain of reactions, especially when applied to biological matrices. Commonly occurring types of reactions include eliminations, rearrangements, substitutions and additions (Moldoveanu, 1998, pp. 9–31).

A pyrolyser is a device that allows heating of a sample to the desired pyrolysis temperature under a controlled atmosphere. This can be achieved in many different ways and accordingly a number of different types of pyrolysers are available. A more detailed description of these can be found elsewhere (Wampler, 2007, pp. 27–46). When used as an analytical technique, pyrolysis is often hyphenated online with gas–chromatography for separation and mass-spectrometry (Py–GC/MS) for characterisation and quantitation of the pyrolysate. Alternatively, online direct analysis by mass–spectrometry (Py–MS) without prior separation is possible (Wampler, 2007, pp. 47–64).

In the past, pyrolysis has often been used for analysis of lignocellulosic materials (Moldoveanu, 1998, pp. 342–345). It is particularly suited for analysis of the lignin fraction, as the pyrolytic degradation products of this polymer are very informative to its composition. Notably, quick determination of the S/G ratio is a major application of analytical pyrolysis (Rodrigues et al., 1999; Barbosa et al., 2008; Fagundes Lopes et al., 2011). Compared to spectroscopic techniques such as FTIR or NIR, analytical pyrolysis yields reliable S/G values even without multivariate calibrations (Nunes et al., 2010; Alves et al., 2011; Sun et al., 2012).

In contrast, the carbohydrate fraction of the lignocellulosic material does not result in a high yield of characteristic degradation products. This is somehow surprising as pyrolysis of almost every single molecule that com-

poses the lignocellulosic polymer matrix results in a well distinguishable fingerprint (Moldoveanu, 1998, pp. 217–237). However, the data obtained when pyrolysing the whole matrix at once is much more than just the sum of those fingerprints. A whole range of additonal paramters contributes to the resulting pyrolysate. An obvious example are linkages between different molecules that are not accounted for when the single molecules or homogenous polymers are pyrolysed. But also other more subtle differences, such as sample particle size, rest water content or inorganic species can influence the pyrolysis process. The degradation of the carbohydrate matrix seems to be more affected by such effects than the lignin polymer. Therefore, thorough structural and compositional characterisation of the whole lignocellulosic matrix by analytical pyrolysis is in the opinion of the author not easy, or maybe even impossible, to achieve.

Initially a Pyrola 2000 (PyrolAB, Lund, Sweden) mounted on a GC/MS (Agilent 5975C/7890, Agilent Sweden, Kista, Sweden) was employed in some of the initial method development experiments. This pyrolyser has a resistive heated platinum foil that serves as sample stage. To allow unattended analysis of large numbers of samples, the setup was upgraded to a Py-2020 pyrolyser with an AS-1020 auto-sampler (both Frontier Lab, Fukushima, Japan). This pyrolyser has an oven with a pneumatic controlled sample introduction system that allows automatic operation.

The GC injector port was operated in split mode to adapt for various sample amounts. Sample amounts were usually in the range between 30 μg to 100 μg. The GC oven was ramped from 40 °C to 300 °C. To achieve best possible analyte separation in the rather short GC runs of about 20 minutes, different heating gradients were chosen to adapt for varied peak–dense parts of the chromatogram. When lower and upper temperatures of the oven program were chosen, the length of the chromatographic run was adapted so that the heating gradient always stayed below the maximum oven heating rate. This is important in order to achieve stable retention times and overall good reproducibility of the chromatograms as otherwise small changes in environmental conditions can influence the oven heating rate. The quadrupole mass spectrometer was operated in scan mode. To achieve the highest possible number of data points over one peak, the scan range was reduced to 35–250 m/z, which includes most of the known pyrolytic lignocellulose degradation products. The above mentioned conditions were initially published in Paper I but subsequently also used in most of the other papers that include Py-GC/MS analysis.

3.3 Data Processing and Analysis

Rawdata from Py-GC/MS has the same basic properties as GC/MS data, i.e. an x–y matrix where x is time and y is an individual mass channel. For each pair of x–y coordinates, there is a value for the signal abundance (the actual measured value). However GC/MS is a technique with a very large application envelope that will require different strategies and steps of data processing, depending on the application. Even within the application of Py-GC/MS, different scenarios can be thought of, that have different requirements to the data processing strategy. For high-throughput application of Py-GC/MS, the obtained data compares to data from GC/MS metabolomics applications. The distinguishing properties of GC/MS data from high–throughput methods will be discussed in further detail below, when describing the commonly implemented data processing steps such as smoothing, chromatogram alignment, background correction, peak recognition, spectral deconvolution and peak integration.

- **Smoothing** is used to remove signal noise of different types. The trade–off in signal smoothing filters is between removing noise artifacts and loss of real data features. Common methods include moving averages, savitzki–golay, or smoothing splines. Mass–spectrometers have a large number of parameters that can be optimised and that will affect the noise level of the basic signal. Therefore optimising the data smoothing is usually a two way process that involves adjusting both processing and instrumental settings.

- **Aligning** of multiple chromatograms can help to identify corresponding peaks in each chromatogram. Generally it can be said that gas–chromatography is one of the most robust chromatographic methods and recent instruments exhibit excellent oven temperature stability, which is the most important factor that affects retention time shifts. This is at least true when other general chromatographic conditions are optimal. However, column overload and similar methodological failures will affect retention time stability. Column overload is sometimes unavoidable when analysing complex samples with hundreds of analytes typical in Py-GC/MS applications. Therefore alignment steps are usually integrated in Py-GC/MS workflows. A number of different approaches are used readily, including both linear and non–linear methods (Nielsen et al., 1998; Fraga et al., 2001; Johnson et al., 2003).

- **Background Correction** in Py-GC/MS rawdata is required due to

several reasons. Especially online pyrolysis setups that use furnaces at constant temperature can exhibit sample bleed if the remaining char is not removed from the reactor during the chromatographic separation. This bleed can give rise to an unwanted constant background signal. During gas chromatography, when elevated temperatures are needed to elute the analytes, column bleed from the stationary phase of the GC column can cause a background signal. Background correction algorithms usually detect points of local signal minima, and subtract the signal under the linear interpolation between two such local minima (Liang et al., 1993).

- **Peak recognition and integration** is the process of identifying start and end point of the chromatographic peak and determining the peak area. For a small number of well separated and ideally shaped peaks, it is rather straightforward to automatise this process. However if these conditions are not met, applying an automated process is not trivial. Chosen Py-GC/MS run conditions are, from a chromatographic point of view, non-ideal, as they should fit for all kind of different peaks (early– or late eluting, small or large). It is simply impossible to obtain ideal peak shapes throughout a whole run. A further major difficulty is the deconvolution of overlapping peaks. Deconvolution should result in an integral of the deconvoluted peak and the underlying mass–spectra.

  Many different approaches have been presented of which none became the only prevailing solution (Tauler, 1995; Stein, 1999; Smith et al., 2006; Dixon et al., 2006; Aberg et al., 2008; van Stokkum et al., 2009; Amigo et al., 2010; Hantao et al., 2012). During the work of this thesis several approaches were tested. The method finally implemented as standard data processing pipeline is based on an algorithm called multivariate curve resolution by alternating regression (MCR-AR) (Tauler, 1995; Jonsson et al., 2005). Its strength is the parallel processing of all chromatograms at once. This guarantees a consistent global peak list opposed to methods where chromatograms are processed in sequence.

While the raw data consisted of a single data matrix, the resulting data of a processing pipeline that includes the above mentioned steps, consists of two matrices. The first contains samples and integrated peak areas per peak and the second contains mass spectra per peak. Now the data are ready for analysis. The strategy or method chosen to analyse processed Py-GC/MS data depends on the aim of the analysis. Either the complete information

from the chemical fingerprint is used to identify and classify variation in the data set (for example in a mutant screen), or annotated peaks are used for quantitative estimates of specific compounds (i.e. lignin type and carbohydrates). Multivariate data analysis is chosen when the complete dataset is used, whereas univariate data analysis is used to make quantitative estimates.

Multivariate Data Analysis and classification methods can be applied directly to the peak table of a sample set. A brief, non-mathematical description of the methods and some common terms follows here. Principal component analysis (PCA) and orthogonal projection to latent structures by partial least squares–discriminant analysis (OPLS-DA) are the two methods mostly used in the presented work (Eriksson et al., 2001; Bylesjö et al., 2006). The main applications of PCA is data simplification and overview, while OPLS-DA is used for sample classification. The PCA algorithm finds iterative a model that describes in a single 'linear combination', or 'principal component' (PC) as much variation of the data as possible. This PC is then subtracted from the initial data, the process is repeated and the next PC can be determined. The number of significant PCs has to be determined with an additional algorithm. The benefit of this method is that a dataset with hundreds of data variables (in this case integrated peaks from the Py-GC/MS chromatogram) can be represented by a few numbers. Graphically, the most common representations is to show two principal components as an x–y plot. In the 'scores plot', every sample is represented by a spot. Samples that cluster in the x–y plot are similar to each other, according to the measured variables. In the 'loadings plot', each measured variable is represented by one spot. The x–y positions of variables in the 'loading plot' explain the contribution of this variable (chromatographic peak) to the variation in the dataset. No prior information about the samples is required for PCA analysis, and the method is therefore 'unsupervised'. PCA is usefull to have an initial overview of a large dataset, to identify potential outliers and to find groups of samples that clearly differ from others.

The graphical output of OPLS-DA analysis is also in the form as 'scores' and 'loadings plots', however with some important differences for the interpretation. OPLS-DA is in contrast to PCA a supervised method. Prior to analysis, the class membership of each sample has to be defined. The aim of OPLS-DA is to find the variation that separates classes. Often, there are just two classes, for example 'mutant' and 'wildtype' plants. This is done by finding iterative a component that maximises the co–variation between samples of the defined classes. OPLS-DA constructs a first model term (termed 'predictive component') that contains all class separating variation, i.e. the difference between mutant and wildtype plants. Eventual further compo-

nents (termed 'orthogonal component') contain only in–class variation.

Similar to regression analysis, certain metrics are calculated for multivariate models to describe their qualities, such as percent of described variation and predictive power of the model. Many of those metrics are calculated in a process called cross–validation. In a widely used cross–validation method (leave–one–out), the model is re–calculated from a subgroup of the available samples and the samples not used in model building are predicted. This process is repeated until every sample has been predicted once. The variability of the model and the accuracy of the predictions during this process are a good measure for the predictive power of the full model.

Univariate data analysis of Py-GC/MS data allows quantitation of the different lignin types, total lignin and total carbohydrate fractions. For this purpose, the mass spectra of each peak in the chromatogram are consulted and potentially identified. A number of freely available libraries of pyrolytic degradation products from lignocellulosic biomass simplify this process significantly (Faix et al., 1991, 1990; Ralph and Hatfield, 1991). Then the identified peaks are grouped and evaluated with univariate statistics such as t–tests. There are several such examples included in the presented work.

4 Results and Discussion

4.1 Py-GC/MS for Wood Analysis

We have developed an analytical protocol for high-throughput applications of Py-GC/MS instruments. We were specifically interested in obtaining a method for high–throughput chemotyping, suitable for minute samples because *Arabidopsis* is commonly used as an experimental plant in functional genomics research on cell walls. Initially, we tried to optimise the instrumental conditions to achieve as high precision as possible. However, we realised that it was not easy to broadly asses analytical reproducibility of chromatograms that typically contained more than 100 peaks. Therefore we concentrated on implementing an automated data processing pipeline that would initially facilitate method development, and eventually become instrumental for high–throughput applications.

In Paper I three main achievements are presented. First, an automated data processing pipeline is implemented that made it possible to quickly obtain integrated peak tables from large Py-GC/MS sample sets. For parallel data processing of large Py-GC/MS data-sets, we adapted multivariate curve resolution by alternating regression (MCR-AR) (Jonsson et al., 2005). The presented pipeline includes data smoothing, background correction, chro-

matogram alignment and the actual MCR–AR algorithm (*Paper I, Figure 1*). The latter is used for peak recognition and deconvolution, peak integration and extraction of deconvoluted mass spectra. MCR–AR has the advantage over many other automated methods, in that it processes chromatograms simultaneously. This results in consistent master peak tables with corresponding mass spectra, valid for all chromatograms.

Second, we employed the data processing pipeline to optimise the instrumental conditions for chemotyping by Py-GC/MS. We used the average relative standard deviation (RSD) of ten technical replicates as a metric for the quality of the whole method. For all optimisation experiments, the sample set consisted of each ten aspen and spruce wood powder samples that were analysed in random sequence. Several parameters were systematically varied to find optimal settings for both instrumental analysis and for the automated processing pipeline. Further, we evaluated the contributions of genuine signal, background, and noise to each total ion count chromatogram. It was found that 37 % of the raw signal contributed to the integrated peaks, 58 % was unavoidable background signal and 5 % instrument noise. For the two different sample types, we achieved an average RSD's below 9 % for a selection of around 100 peaks that represented 99 % of the total integrated peak area.

Third, we defined a new simplified rule-set to analyse Py-GC/MS data from lignocellulosic samples. For many applications the only needed information is the composition according to the different types of lignin (S–, G–, H–lignin) and fraction of carbohydrates. We found that the largest peak in the mass spectra was for 90 % of the peaks unequivocally assignable to one of the above mentioned compositional groups (*Paper VI, Table 1*). We wrote a script that used this feature to determine the composition of hundreds of samples within seconds. The result of this script compared well with manual peak identification with just minor deviations for lignin (below 1 %) and still acceptable for carbohydrates (below 5%) (*Paper VI, Table 3*).

In summary, we developed and validated an analytical method that, to our knowledge, for the first time combines Py-GC/MS with a highly automated multivariate curve resolution based data processing pipeline. Method validation using authentic test samples demonstrated an excellent overall reproducibility, indicated by average RSD values around 10 % for close to 100 peaks analyzed in ten technical replicates. The Py-GC/MS method is applied frequently both in research (forensics, art object analysis) and industry (pulp and paper, paint and coating, rubber tires, plastic recycling). To the authors knowledge, tedious manual data processing is still commonly used. Our novel approach to quickly classify peaks of pyrolytic degradation prod-

ucts according to their highest abundant mass channel in the mass spectra is particularly of interest for studies on lignocellulose. This method significantly reduces time from analysis to interpretable results with an acceptable trade-off in accuracy. It is especially of interest in laboratories where Py-GC/MS is used by non–chemists as a standard procedure to obtain quantitative values on the samples' lignin composition.

4.2 Data Integration in Screening Projects

In Paper II we demonstrate a novel application of the OPLS-DA method (Bylesjö et al., 2006) for data mining of combined data-sets with large between–study variation. Between–study variation is a common but often unavoidable problem when complex datasets from large projects that involve parallel experiments are to be combined for global data analysis. In our case, we evaluated a large–scale functional genomics project where 736 transgenic trees representing 44 genotypes and 174 independent lines were grown together with 165 wildtype trees in the greenhouse (*Paper VI, Tables S1 and S2*). Because of the size of the experiments the trees were grown in four batches at different times of the year. The Py-GC/MS method developed in Paper I was used for chemical fingerprinting of wood samples for all trees. Datasets consisted of tables with one row per analyzed sample and one column for each integrated peak from the chromatogram. PCA analysis of all obtained datasets revealed substantial between experiment variation (*Paper II, Figure 2A*). This is not surprising since environmental conditions will unavoidable vary in the greenhouse during different times of the year.

When only few mutant lines are investigated it is common practice to establish OPLS-DA models for each mutant line and compare them to corresponding wildtypes (Bylesjö et al., 2006). But in large screening projects this approach becomes cumbersome as each model yields several graphs and metrics to be considered. Here we developed a method where a matrix was based on consensus profiles for all chemical fingerprints. As consensus profiles we chose the predictive loading vectors from OPLS-DA analysis. Because each transgenic line is compared to a group of wildtype samples, the predictive loading vector becomes independent of between–study variation, which is the precondition for global data analysis.

The experiment was designed to contain a large number of control wildtype trees in each growth batch, whereas the number of replicates per transgenic line varied between three to six. Initially, the number of OPLS-DA components was determined by leave–one–out cross–validation. To obtain class–balanced models, we implemented a second important achievement: a boot–strapping approach where we calculated the average from 300

randomly re-sampled class-balanced models, each built using five wildtype trees. This process also yielded cross validation metrics by using the number of misclassifications from the 300 re-sampled models. At this stage the method was validated by comparing the initial PCA analysis of the raw-data with the PCA analysis of the consensus profile matrix that confirmed the successful exclusion of between–study variations (*Paper II, Figures 2A and 2B*).

A third achievement in Paper II is the visualisation of Py-GC/MS data by hierarchical clustered heatmaps for quick and intuitive interpretation of the dataset (*Paper II, Figure 3*). We used 'Ward's clustering' method to reorder the heatmap according to similarities of the consensus profiles from individual transgenic lines. The clustering was validated by PCA analysis (*Paper II, Figures 4A and 4B*). It was found that transgenic lines with similar chemotypes indeed were grouped into the same cluster. Along with the transgenic lines, the dataset also contained positive controls, i.e. transgenics with already well described chemotypes. These are used to predict the effect of unknown genes in transgenic lines that clustered with the positive controls.

In summary, we have demonstrated a novel data normalisation strategy for large–scale screening studies based on the well–established multivariate classification method OPLS-DA. Despite an increasing use of –omics platforms in modern research, there are to our knowledge just a few data processing strategies that allow the global analysis of such complex datasets. Also the use of boot–strapping to obtain better comparable OPLS-DA models is novel in this field, and solves a problem that is inherent in many life science experiments. Often the number of bio-replicates at the end of an experiment can not be known in advance, and it is often not feasible to have the same number of samples for every sample type. In all such cases, it is a good solution to create class–balanced models by applying the bootstrapping strategy presented here. To make optimal use of the obtained global data set, an appropriate method to access the data was established. Using hierarchical clustered heatmaps for a matrix of Py-GC/MS effect profiles is a large simplification for intuitive and quick evaluation of large Py-GC/MS chemotyping datasets.

4.3 Case Studies

In Paper I and II we developed a Py-GC/MS method for high–throughput analysis and dealt with issues arising when analyzing large chemotyping datasets. In the following Papers III–VI, Py-GC/MS was applied along with many other techniques and experiments to explore gene function in cell wall

biosynthesis both in aspen and in *Arabidopsis*. Py-GC/MS served either for initial quick chemotyping or for quantitation of the various lignin fractions.

4.3.1 Functional Studies of Genes in Primary Carbohydrate Metabolism

Function of Sucrose Synthase in Wood Formation of Aspen In many tree species, including *Populus sp.*, photosynthate is transported as sucrose from photosynthetically active leaves through the phloem to sink tissues (Turgeon, 1996). Most of the biomass in a tree is laid down as wood in the stem and the wood forming tissues are therefore the trees' most important carbon sink (Pregitzer and Euskirchen, 2004; Bonan, 2008). Two classes of enzymes that catalyze sucrose hydrolysis are invertases (INV) and sucrose synthases (SUS) (Sturm and Tang, 1999). Invertases are a group of β–fructosidases that hydrolyze sucrose irreversibly. Sucrose hydrolysis catalyzed by SUS is reversible and yields UDP-Glc and fructose. Despite a large number of studies in many different species, the relative importance for these two enzymes in sucrose hydrolysis is unclear (Zrenner et al., 1995; Nguyen et al., 1997; Ruan et al., 2003; Chourey et al., 1998; Barratt et al., 2009; Baroja-Fernández et al., 2011).

About 40 % of wood is composed of cellulose. A concept initially based on immuno co–localisation of CesA and SUS by Amor et al. (1995), proposed that SUS associated with the cellulose biosynthetic machinery at the plasmamembrane for efficient provision of UDP-Glc (Haigler et al., 2001). While conclusive experimental evidence for the direct interaction of SUS and CesAs is still missing, a recent study showed that an *Arabidopsis sus1 sus2sus3sus4* quadruple mutant grew normally and had indifferent proportion of cellulose in the stem tissues (Barratt et al., 2009).

In Paper III we addressed the role of SUS in wood–forming tissue by generating transgenic aspen trees that had downregulated levels of *SUS1* and *SUS2*, the major SUS isoforms in developing wood (*Paper III, Figure S1*). The almost complete downregulation of *SUS1* and *2* in three transgenic lines resulted in a much reduced SUS activity (*Paper III, Figures 1 and 7*). Despite this, the trees developed and grew almost like the wild type controls (*Paper III, Table 1*), and no difference in wood anatomy was observed under the light microscope (*Paper III, Table 4*). To further investigate any role of SUS in wood formation, we conducted a more in–depth analysis of the wood chemistry and ultrastructure in the transgenic lines. Initially, we applied Py–GC/MS combined with class–separation analysis by OPLS–DA for wood chemotyping. The samples from all three lines clearly separated in a two class model (*Paper III, Figure 2*). By peak integration of the Py–GC/MS data, we found an increase in the S/G ratio and a trend suggesting

reduction of the carbohydrate fraction in the transgenic lines compared to wildtype control trees (*Paper III, Table 2*). Cell–specific chemotyping with FTIR micro-spectroscopy (Gorzsás et al., 2011) confirmed this difference, and also showed that all cell types in the wood were chemically modified in a similar way (*Paper III, Figure S2*). However, quantification of cellulose, hemicelluloses and lignin fractions by wet chemical analysis showed that their relative abundance was, by large, not altered (*Paper III, Figure 4A*).

A major modification was however found in a reduced bulk density and large differences in mechanical properties of the wood (*Tables 3 and 4*). Surprisingly, the decrease in wood density could not be readily explained in a difference of the fiber to vessel ratio or a reduced fiber cell wall area in the wet sample. These data suggested that the density of the cell wall *per se* was decreased, meaning that the wall ultrastructure was more porous. Indeed, much thinner cell walls were observed by electron microscopy in dehydrated sections, suggesting a higher degree of collapse after dehydration, and therefore supporting this idea (*Paper III, Figure 5*). The density data also allowed us to estimate the amount of wood polymers per volume of wood (*Paper III, Figure 4B*). Hence reduced SUS activity caused an overall reduction of carbon flow to all wood polymers, and not specifically to cellulose.

Our study demonstrates the limitations of the common practice to normalise chemical analysis results of the cell walls against dry-weight. Without further analysis of wood density, we would have concluded that SUS did not have any major effect on the biosynthesis of cell wall polymers, as was the case with the *Arabidopsis* quadruple mutant. In fact we also demonstrated with Py-GC/MS and OPLS-DA, that these *Arabidopsis* mutants indeed were severely affected in their cell wall, which further proofs the strength and sensitivity of Py-GC/MS chemotyping. When browsing the recent literature on the function of cell wall biosynthesis genes, there are few studies where density has been assessed. This means that whenever a difference in polymer composition is found it can not be concluded which polymer is primarily affected, and which was affected just as a result of the percentage normalisation. However, in cases where gene function is known, an educated guess can be done. Despite this, there are examples of data interpretation where a reduction of for example lignin was also interpreted as an increase in cellulose without any further investigations of density to make sure that this was not just an effect of normalisation (Hu et al., 1999).

The novel finding in this study is that SUS is important for carbon flux to all wood polymers, and does not seem to have an exclusive role in providing UDP-Glc for cellulose biosynthesis. It is interesting to note that despite

the almost absent activity of SUS, most of the cellulose was still produced. This strongly suggests that INVs have an important function in sucrose metabolism in wood forming tissues. We also demonstrate on the requirement for wood density measurement to fully understand the effect of gene modification on cell wall polymer biosynthesis.

Function of Fructokinase in Wood Formation of Aspen Sucrose is delivered from its transport in the phloem to the wood forming tissues, where it is hydrolysed by INV or SUS to enter primary carbon metabolism for biosynthesis of cell wall polymers (Sturm and Tang, 1999). In both cases, fructose is formed, which has to be phosphorylated by either a fructo– (FRK) or hexokinase (HXK) in order to be further metabolised (Sharples and Fry, 2007). The affinity of FRKs to fructose is about three orders of magnitudes higher than for HXKs (Renz and Stitt, 1993; Granot, 2007). This positions FRKs as a potentially very important enzyme in sucrose metabolism in wood forming tissues.

There is limited knowledge on the role of FRK in carbon metabolism to cell wall polymers in wood. In Paper IV we addressed this question by downregulating *FRK2A* and *FRK2B*, which were found to be the by far most important FRK isoforms in wood forming tissues (*Paper IV, Figure S2*). A single RNAi construct resulted in strong downregulation of both genes in three transgenic lines and reduced *FRK* transcript to 4–9 % in developing wood (*Paper IV, Figure 1 and Table S6*). The transgenic greenhouse grown lines showed only minor growth phenotype in comparison to the wildtype controls (*Paper IV, Table 1*).

An initial analysis of the transgenic lines with Py-GC/MS and OPLS-DA data analysis identified a characteristic chemotype in all transgenic lines (*Data not shown*). Compositional analysis by integration and summation of Py-GC/MS peaks showed an increase in the lignin S/G ratio as well as a moderate decrease in the carbohydrate to lignin ratio (*Paper IV, Table 2*). To obtain more detailed information on the amount of individual cell wall polymers, we conducted solution state 2D ^{13}C-^{1}H HSQC NMR, wet-chemical cellulose determination by Updegraff (Updegraff, 1969) and monosaccharide analysis of extractive–free wood hydrolysates by HPAEC-PAD (Perry and Currie, 2006). NMR confirmed the decrease in cellulose fraction which was quantified wet chemically to be about 20 % in all lines (*Paper IV, Figures 2 and 3*). The main finding from monosaccharide analysis was a decrease in glucose, which was interpreted as reduction of non-crystalline cellulose as also the total amount of cellulose was reduced (*Paper IV, Table S1*). FTIR micro-spectroscopy (Gorzsás et al., 2011) showed that

the modification in chemotype took place both in fibers and vessels (*Paper IV, Figure S4*). Wood anatomical features were evaluated by light– and electron microscopy. This showed significantly thinner cell walls in the transgenic lines, but generally unaffected anatomy (*Paper IV, Figure 4, Table 3*).

By using LC-MS/MS and robotic enzyme assays we assessed the general status of the carbohydrate metabolism in the wood forming tissue (Gibon et al., 2004, 2002; Lunn et al., 2006). Whereas sucrose, glucose, and fructose levels were increased, the assayed hexose phosphates were all reduced (*Paper IV, Figure 6*). These results confirm the important role of FRKs in carbon fluxes to the different wood polymers. Further metabolic data and microarray transcript profiling showed no signs of pleiotropic effects that would weaken the presented interpretation of our data (*Paper IV, Figure S5 and Tables S3, S4 and S5*).

Taken together, we presented novel data on the role of FRK in wood forming tissue. In–depth analysis of the cell wall chemistry and ultrastructure suggests FRK as an important enzyme for carbon partitioning to the cell wall in general and to cellulose biosynthesis in particular.

4.3.2 Functional Genomics of Genes involved in Lignin Biosynthesis

Function of Cinnamate 4–Hydroxylase for lignification, wood structure and wood mechanics in aspen The lignin pathway has been extensively studied and is rather well understood (Vanholme et al., 2008). Cinnamate 4–Hydroxylase (C4H) is an enzyme in the early lignin pathway, that hydroxylates cinnamic–acid to p–coumaric–acid (Boerjan et al., 2003). Functional analysis of C4H has previously been conducted in *Arabidopsis*, tobacco and alfalfa xylem. It was consistently found that a reduction in C4H resulted in reduced lignin content and in most cases altered lignin composition (Sewalt et al., 1997; Blee et al., 2001; Shadle et al., 2007; Schilmiller et al., 2009)

In paper V we evaluated the role of C4H for the first time during wood formation in a tree (aspen). Two independent transgenic lines with about 80 % reduction of C4H expression were identified and the wood in the transgenic lines was characterised for chemical, mechanical and ultrastructural properties.

In the greenhouse the transgenic lines showed a slight reduction in height growth, but plant development and wood anatomy was similar to wildtype control trees. Wet chemical procedures revealed a severe reduction of Klason lignin, from 22 % to 15 %, with corresponding increase in cellulose due to weight normalisation (Ona et al., 1995)(*Paper V, Table 1*). Py-GC/MS analysis and FTIR spectroscopy confirmed the severe reduction of lignin

(*Paper V, Table 1*). The lignin S/G ratio, determined by Py-GC/MS, was not affected.

To assess the ultrastructural consequences of such a severe reduction of a major wood constituent, we first measured the bulk density of the transgenic trees, which was reduced by 10 % (*Paper V, Figure 1*). This was supported by AFM measurements, which showed a looser arrangement of hemicelluloses/lignin lamellas, eventually swollen and hydrated due to the low abundance of hydrophobic lignin (*Paper V, Figures 2 and Table 2*). Cellulose MFA and crystallinity, the most determining parameters for the mechanical properties of wood, were measured by X-ray diffraction and were not found to be affected (*Paper V, Figure 1*).

To investigate the influence of the modified ultrastructure on the mechanical properties of the wood, we conducted a number of static and dynamic micro-mechanical tests. From theory, a reduction in wood density should result in a decrease of stiffness (Kohler and Spatz, 2002). Measurements by static tensile testing showed that this was indeed the case (*Paper V, Figures 1 and 3*). However the ultimate tensile strength was not affected (*Paper V, Figure 1*). This is inline with the textbook knowledge, stating that cellulose microfibrils are the by far most important load bearing molecules in uni-axial tensile strength tests (Salmén and Burgert, 2008). However, it still seems surprising that a relative reduction of as much as 30 % in the lignin polymer did not result in more significant changes in strength parameters. Dynamic mechanical testing under varied humidity and temperature environment did not reveal any change in the glass transition temperature of the wood polymer (*Paper V, Figures 4,5,6 and 8*).

In summary, this was the first time that C4H was functionally analysed in a tree species, which facilitated novel findings on the effect of lignin reduction on wood structure and wood mechanics. We demonstrate that a strong reduction of lignin amount reduces wood density, which is the most likely cause of the concomitant reduction in stiffness. The most surprising result of our study was however that the ultimate tensile strength of the transgenic wood was not affected despite the severe reduction in lignin.

Functional Analysis of the Transcription Factor MYB103 in *Arabidopsis* Secondary cell wall biosynthesis has been shown to be regulated by a transcriptional network (Caño Delgado et al., 2010). According to current understanding, there are 'master switches' for secondary wall formation that are specific for fibers and vessels. These 'master switches' in turn control a network of common downstream transcriptional regulators, and eventually the biosynthetic genes for cellulose, hemicelluloses and lignin. In the case of the lignin

biosynthetic pathway, several MYB transcription factors are suggested to be key regulators (Zhao and Dixon, 2011)

In Paper VI, we investigated the cell wall chemistry of two *MYB103* null mutant alleles. MYB103 was initially identified by Zhong et al. (2008) and shown by several authors to be a direct target of the fiber and vessel master switches (Yamaguchi et al., 2008; Nakano et al., 2010; Yamaguchi et al., 2010). Because MYB103 was found to activate CesA8 in a protoplast transactivation assay, it was speculated to have a role in regulation of cellulose biosynthesis (Zhong et al., 2008).

We identified two T-DNA null alleles for *MYB103* that initially were evaluated by Py-GC/MS. Peak integration for compositional analysis revealed that a lack of MYB103 resulted in up to 75 % reduction in S–type lignin. G–type lignin was increased, leaving the total proportion of lignin by large unaffected (*Paper VI, Table 1*). Further in–depth chemical cell wall characterisation included solution state 2D ^{13}C-^{1}H HSQC NMR on solubilised whole cell wall material, Klason lignin, Updegraff cellulose and monosaccharide composition by HPAEC-PAD. NMR analysis supported the Py-GC/MS data, showing severely reduced S– and increased G–type lignin, but also minor reductions in the proportion of cellulose and glucomannan and a minor increase in xylans (*Paper VI, Figure 4*). Equally, wet chemical analysis also showed a trend (non significant) for a decrease in the proportion of cellulose in *myb103* mutants, whereas Klason lignin was the same as for wildtype (*Paper VI, Table 1*). A difference in hemicelluloses between *myb103* mutants and wildtype plants was also indicated in the HPAEC-PAD data (*Paper VI, Table 2*).

In *Arabidopsis* interfascicular fibers, xylem fibers and vessels all form lignified secondary cell walls (Chaffey et al., 2002). FTIR micro-spectroscopy with OPLS-DA data analysis (Gorzsás et al., 2011) revealed that all these types had different chemotypes (*Paper VI, Figure 5*). We confirmed the enrichment in S–type lignin in interfascicular fibers, earlier shown by Mäule staining (Zhong et al., 2008), but also demonstrated in differences in the carbohydrate matrix of the cell wall. The effect of *myb103* mutation on S–type lignin could be observed in all cell types indicating a general function of this transcription factor (*Paper VI, Figure 5*).

We tried to identify potential downstream interactors of MYB103 in order to better understand the function of this transcription factor. We found that the transcript levels of F5H, the enzyme that branches into the synapyl alcohol synthesis, were significantly reduced in *myb103* mutants (*Paper VI, Figure 7*). We further demonstrated that the effect of oligolignols were similar in a F5H mutant and in *myb103* mutants (*Paper VI, Figure 8*). Together,

these data strongly indicate that MYB103 is required for F5H expression, hence S–lignin biosynthesis. However, a protoplast transactivation assay did not show any interaction between MYB103 and the promoter of F5H (*Paper VI, Table 4*).

Transcript profiling of *myb103* mutants revealed a number of transcription factors that were significantly downregulated in both mutant alleles (*Paper VI, Table 3*). We hypothesised that the effect of MYB103 on F5H expression was mediated through any of these transcription factors. However, none of the tested TFs could activate the F5H promoter in our protoplast system. Thus, we have no simple explanation for the molecular mechanism connecting MYB103 with F5H expression.

The novel finding in this study is a first clue on the molecular regulation of 'S–type' lignin formation in lignocellulose. Although an earlier study claims that the master regulator SND1 directly binds and regulates F5H (Zhao et al., 2010) we demonstrate that this is an effect of using the *Arabidopsis* gene and the *Medicago* promoter in the transactivation study. Thus, at present MYB103 is a main candidate for transcriptional regulation of F5H, but our data suggest that this regulation also involves other players.

5 Conclusions and Future Perspectives

5.1 Concluding Remarks

An automated data processing pipeline for Py-GC/MS data allowed the processing of a large number of complex chromatograms in a short time. This opened a whole new range of applications for this technique that is complementary to direct pyrolysis and spectroscopic techniques such as FTIR and NIR. While spectroscopic techniques yield information about chemical bonding characteristics in a sample, mass–spectrometry results in identification and quantitation of molecules. Both types of data can easily be used as chemical fingerprints, but the chemical interpretation differs widely. Py-GC/MS data allows a very direct and uncomplicated interpretation of the lignin fraction in lignocellulosic samples. Combined with the possibility for high sample throughput, this allows for replacing tedious wet-chemical methods such as the 'Klason lignin' or 'thioacidolysis' procedure. It can also be applied on small sample amounts. Compared to direct Py-MS techniques, the chromatographic separation resulting in higher sensitivity of Py-GC/MS allows for detection of minute differences between, for example, transgenic plants.

We used transgenic plants to study both structure and biosynthesis of

the plant cell wall. Transgene modifications can result in structural cell wall features beyond the natural variation for a species. Thus, transgenic material is useful for testing hypothesis about function of specific cell wall polymers and wood structures and about the control of cell wall biosynthesis. For this purpose a range of techniques was applied that included both chemical analysis and mechanical testing, assessment of bulk properties as well as assessment of detailed anatomical parameters to have a broader understanding of the plant cell wall. Our findings showed that the use of several, sometimes overlapping, experiments is very valuable to build and verify more detailed models about the cell wall structure and the role of each individual polymer.

5.2 Future Perspectives

To reach the goal of tailor–made raw material from forest production systems using biotechnology tools, it is not sufficient to understand the biosynthetic pathways and the molecular biology of wood formation. To develop transgenic approaches that successfully target and modify specific properties of interest for wood processing and the final product will also require an understanding of the chemical network of wood polymers and the resulting ultrastructure. This in turn requires that a large range of wood properties can be measured.

In particular the interaction of the different wall polymers and the resulting bulk properties are still poorly understood. Cell wall material density is one of them. In transgenic trees, downregulated in genes functional in either lignin biosynthesis or primary carbon metabolism, bulk density was reduced. In SUS and C4H mutant lines, we found that this was to a certain extent also due to looser or more porous cell wall material. While this seems logic in the case of reduced lignin, it is more intriguing in the case of the SUS transgenics where the abundance of all polymers was reduced. These findings are very encouraging for biotechnological applications: There seems to be a substantial range of cell wall modifications possible while maintaining the structural integrity of the wall.

To further explore the range and possible targets of cell wall modifications, novel approaches and methodologies in cell wall analysis will be needed, that in combination with traditional tools will increase the resolution of our knowledge of the plant cell wall. A special focus should be put on computational modelling to integrate and connect various data-sets numerically. In terms of current analytical developments it seems that the recent generous funding of renewable biomass research has created a surge in the availability of cutting edge tools to the laboratories in this field. Hence,

there has recently been an impressive number of cell wall analysis related methods published that will produce novel information. A general trend in bio-analytical chemistry is the increasing spatial resolution of many analytical methods. This should also benefit cell wall research as many structural details need to be assessed at sub-micron resolution. For example, highly resolved x-ray micro-tomography could become instrumental to quickly obtain a numerical evaluation of wood anatomy. It can also be expected that imaging mass–spectrometry techniques will be used more readily in cell wall analysis to obtain detailed chemical maps of specific morphological regions of the cell wall.

Acknowledgements

I would like to thank:

Supervisors: Björn Sundberg, Thomas Moritz. **Co–workers:** Current and past members of the Sundberg, Niittylä and Fischer research groups. **Office mates:** Steve Bruce, Shuzen Men, Peter Immerzeel, Melissa Roach, Judith Felten. **SLU / University:** Gerhard Gröbner, Lasse Lundmark, Markus Grebe, Harry Wu, Kjell Olofsson, Junko Takahashi-Schmidt, Rishi Bhalerao, László Bakó, Erik Björn, Slim Gerge Molnar, Bo Zhang, Inga–Lis, Gun–Britt, Krister Lundgren, Inga–Britt Carlsson, Margareta Zetherström. **Collaborators:** Johan Trygg, Rui C. Pinto, Mattias Eliasson, Mattias Hedenström, András Gorzsás, Ingo Burgert, Bo Zhang, Lennart Salmén, David Öhman, Manoj Kumar, Ingela Bjur-hager, Lars A. Berglund, Wout Boerjan and group, Marc Stitt and group, Brian Ellis, Geoffrey Daniel and group. **Funding Agencies of my PhD position:** FORMAS, Bio4Energy (the Swedish program for renewable energy). **Friends:** Urs Fischer, Ilka Abreu, Chrigu Bigler, Mätthu Grimm, Mustafa Assim. **Families in Switzerland and Vietnam**, **Duyên**.

References

Aberg, K.M., Torgrip, R.J.O., Kolmert, J., Schuppe-Koistinen, I., and Lindberg, J., 2008. Feature detection and alignment of hyphenated chromatographic-mass spectrometric data - extraction of pure ion chromatograms using kalman tracking. *Journal of Chromatography A*, 1192(1):139–146.

Åkerholm, M. and Salmén, L., 2005. The oriented structure of lignin and its viscoelastic properties studied by static and dynamic FT-IR spectroscopy. *Holzforschung*, 57(5):459–465.

Alves, A., Simoes, R., Stackpole, D.J., Vaillancourt, R.E., Potts, B.M., Schwanninger, M., and Rodrigues, J., 2011. Determination of the syringyl/guaiacyl ratio of *Eucalyptus globulus* wood lignin by near infrared-based partial least squares regression models using analytical pyrolysis as the reference method. *Journal of Near Infrared Spectroscopy*, 19(5):343–348.

Amigo, J.M., Popielarz, M.J., Callejon, R.M., Morales, M.L., Troncoso, A.M., Petersen, M.A., and Toldam-Andersen, T.B., 2010. Comprehensive analysis of chromatographic data by using parafac2 and principal components analysis. *Journal of Chromatography A*, 1217(26):4422–4429.

Amor, Y., Haigler, C.H., Johnson, S., Wainscott, M., and Delmer, D.P., 1995. A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 92(20):9353–9357.

Andersson-Gunnerås, S., Mellerowicz, E.J., Love, J., Segerman, B., Ohmiya, Y., Coutinho, P.M., Nilsson, P., Henrissat, B., Moritz, T., and Sundberg, B., 2006. Biosynthesis of cellulose–enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. *The Plant Journal*, 45(2):144–165.

Aspeborg, H., Schrader, J., Coutinho, P.M., Stam, M., Kallas, Å., Djerbi, S., Nilsson, P., Denman, S., Amini, B., Sterky, F., Master, E., Sandberg, G., Mellerowicz, E., Sundberg, B., Henrissat, B., and Teeri, T.T., 2005. Carbohydrate-active enzymes involved in the secondary cell wall biogenesis in hybrid aspen. *Plant Physiology*, 137(3):983–997.

Atalla, R.H. and Agarwal, U.P., 1985. Raman microprobe evidence for lignin orientation in the cell walls of native woody tissue. *Science*, 227(4687):636–638.

Barbosa, L.C.A., Maltha, C.R.A., Silva, V.L., and Colodette, J.L., 2008. Determination of the siringyl/guaiacyl ratio in eucalyptus wood by pyrolysis-gas chromatography/mass spectrometry (Py-GC/MS). *Quimica Nova*, 31:2035–2041.

Baroja-Fernández, E., Muñoz, F.J., Li, J., Bahaji, A., Almagro, G., Montero, M., Etxeberria, E., Hidalgo, M., Sesma, M.T., and Pozueta-Romero, J., 2011. Sucrose synthase activity in the *sus1/sus2/sus3/sus4 Arabidopsis* mutant is sufficient to support normal cellulose and starch production. *Proceedings of the National Academy of Sciences*, 109(1).

Barratt, D.H.P., Derbyshire, P., Findlay, K., Pike, M., Wellner, N., Lunn, J., Feil, R., Simpson, C., Maule, A.J., and Smith, A.M., 2009. Normal growth of *Arabidopsis* requires cytosolic invertase but not sucrose synthase. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31):13124–13129.

Bauer, S., 2012. Mass spectrometry for characterizing plant cell wall polysaccharides. *Frontiers in Plant Science*, 3(45):1–6.

Bhalerao, R., Nilsson, O., and Sandberg, G., 2003. Out of the woods: forest biotechnology enters the genomic era. *Current Opinion in Biotechnology*, 14(2):206–213.

Biermann, C. and McGinnis, G., 1989. *Analysis of Carbohydrates by GLC and MS*. CRC Press, 292 pages.

Blee, K., Choi, J.W., O'Connell, A.P., Jupe, S.C., Schuch, W., Lewis, N.G., and Bolwell, G., 2001. Antisense and sense expression of cDNA coding for CYP73A15, a class II cinnamate 4-hydroxylase, leads to a delayed and reduced production of lignin in tobacco. *Phytochemistry*, 57(7):1159–1166.

Bodig, J. and Jayne, B., 1993. *Mechanics of Wood and Wood Composites*. Krieger Publishing Company, Malabar, FL, USA, 712 pages.

Boerjan, W., 2005. Biotechnology and the domestication of forest trees. *Current Opinion in Biotechnology*, 16(2):159–166.

Boerjan, W., Ralph, J., and Baucher, M., 2003. Lignin biosynthesis. *Annual Review of Plant Biology*, 54:519–546.

Bonan, G.B., 2008. Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, 320(5882):1444–1449.

Booker, R. and Sell, J., 1998. The nanostructure of the cell wall of softwoods and its functions in a living tree. *Holz als Roh- und Werkstoff*, 56(1):1–8.

Brown, D.M., Zeef, L.A.H., Ellis, J., Goodacre, R., and Turner, S.R., 2005. Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell*, 17(8):2281–2295.

Browning, B., 1967. *Methods of wood chemistry*. Interscience Publishers, 498 pages.

Brunow, G. and Lundqvist, K., 2010. Functional groups and bonding patterns in lignin (including the lignin-carbohydrate complexes). In C. Heitner, D. Dimmel, and J.A. Schmidt, editors, *Lignin and lignans: advances in chemistry*, pages 267–299. CRC Press, Taylor & Francis Group, Boca Raton, FL, USA.

Burgert, I. and Fratzl, P., 2009. Plants control the properties and actuation of their organs through the orientation of cellulose fibrils in their cell walls. *Integrative and Comparative Biology*, 49(1):69–79.

Bylesjö, M., Rantalainen, M., Cloarec, O., Nicholson, J.K., Holmes, E., and Trygg, J., 2006. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics*, 20(8-10):341–351.

Caño Delgado, A., Lee, J.Y., and Demura, T., 2010. Regulatory mechanisms for specification and patterning of plant vascular tissues. *Annual Review of Cell and Developmental Biology*, 26(1):605–637.

Caffall, K.H. and Mohnen, D., 2009. The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydrate Research*, 344(14):1879–1900.

Chaffey, N., Cholewa, E., Regan, S., and Sundberg, B., 2002. Secondary xylem development in *Arabidopsis*: a model for wood formation. *Physiologia Plantarum*, 114(4):594–600.

Chakar, F.S. and Ragauskas, A.J., 2004. Review of current and future softwood kraft lignin process chemistry. *Industrial Crops and Products*, 20(2):131–141.

Chourey, P.S., Taliercio, E.W., Carlson, S.J., and Ruan, Y.L., 1998. Genetic evidence that the two isozymes of sucrose synthase present in developing maize endosperm are critical, one for cell wall integrity and the other for starch biosynthesis. *Molecular and General Genetics*, 259(1):88–96.

Ciucanu, I. and Kerek, F., 1984. A simple and rapid method for the permethylation of carbohydrates. *Carbohydrate Research*, 131(2):209–217.

Coimbra, M., Delgadillo, I., Waldron, K., and Selvendran, R., 1996. Isolation and analysis of cell wall polymers from olive pulp. In H. Linskens and J. Jackson, editors, *Plant Cell Wall Analysis*, Modern Method in Plant Analysis, pages 19–43. Springer.

Cosgrove, D.J., 2000. Expansive growth of plant cell walls. *Plant Physiology and Biochemistry*, 38:109–124.

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

Dimmel, D., 2010. Overview. In C. Heitner, D. Dimmel, and J.A. Schmidt, editors, *Lignin and lignans: advances in chemistry*, pages 1–10. CRC Press, Taylor & Francis Group, Boca Raton, FL, USA.

Dimmel, D. and Gellerstedt, G., 2010. Chemistry of alkaline pulping. In C. Heitner, D. Dimmel, and J.A. Schmidt, editors, *Lignin and lignans: advances in chemistry*, pages 349–391. CRC Press, Taylor & Francis Group, Boca Raton, FL, USA.

Ding, S. and Himmel, M., 2006. The maize primary cell wall microfibril: A new model derived from direct visualization. *Journal of Agricultural and Food Chemistry*, 54(3):597–606.

Dixon, S.J., Brereton, R.G., Soini, H.A., Novotny, M.V., and Penn, D.J., 2006. An automated method for peak detection and matching in large gas chromatography-mass spectrometry data sets. *Journal of Chemometrics*, 20(8–10):325–340.

Donaldson, L., 2008. Microfibril angle: Measurment, variation and relationships - a review. *IAWA Journal*, 29(4):345–386.

Dubois, M., Gilles, K., Hamilton, J., Rebers, P., and Smith, F., 1956. Colorimetric method for determination of sugars and related substances. *Analytical Chemistry*, 28(3):350–356.

Eriksson, L., Johansson, E., N., K.W., and Wold, S., 2001. *Multi- and Megavariate Data Analysis: Principles and Applications*. Umetrics Academy, 533 pages.

Fagundes Lopes, F.J., Silverio, F.O., Ferreira Baffa, D.C., Loureiro, M.E., and Pereira Barbosa, M.H., 2011. Determination of sugarcane bagasse lignin S/G/H ratio by pyrolysis GC/MS. *Journal of Wood Chemistry and Technology*, 31(4):309–323.

Faix, O., Fortmann, I., Bremer, J., and Meier, D., 1991. Thermal-degradation products of wood - gas-chromatographic separation and mass-spectrometric characterization of polysaccharide derived products. *Holz Als Roh-Und Werkstoff*, 49(5):213–219.

Faix, O., Meier, D., and Fortmann, I., 1990. Thermal-degradation products of wood - gas-chromatographic separation and mass-spectrometric characterization of monomeric lignin derived products. *Holz Als Roh-Und Werkstoff*, 48(7-8):281–285.

Fengel, D. and Wegener, G., 1989. *Wood: chemistry, ultrastructure, reactions*. W. de Gruyter, Berlin, Germany, 613 pages.

Fernandes, A.N., Thomas, L.H., Altaner, C.M., Callow, P., Forsyth, V.T., Apperley, D.C., Kennedy, C.J., and Jarvis, M.C., 2011. Nanostructure of cellulose microfibrils in spruce wood. *Proceedings of the National Academy of Sciences*, 108(47):E1195–E1203.

Fraga, C.G., Prazen, B.J., and Synovec, R.E., 2001. Objective data alignment and chemometric analysis of comprehensive two-dimensional separations with run-to-run peak shifting on both dimensions. *Analytical Chemistry*, 73(24):5833–5840.

Fukushima, R.S. and Hatfield, R.D., 2004. Comparison of the acetyl bromide spectrophotometric method with other analytical lignin methods for determining lignin concentration in forage samples. *Journal of Agricultural and Food Chemistry*, 52(12):3713–3720.

Gallezot, P., 2012. Conversion of biomass to selected chemical products. *Chemical Society Reviews*, 41(4):1538–1558.

Geisler-Lee, J., Geisler, M., Coutinho, P.M., Segerman, B., Nishikubo, N., Takahashi, J., Aspeborg, H., Djerbi, S., Master, E., Andersson-Gunneråas, S., Sundberg, B., Karpinski, S., Teeri, T.T., Kleczkowski, L.A., Henrissat, B., and Mellerowicz, E.J., 2006. Poplar carbohydrate-active enzymes.

gene identification and expression analyses. *Plant Physiology*, 140(3):946–962.

Gibon, Y., Blaesing, O., Hannemann, J., Carillo, P., Hohne, M., Hendriks, J., Palacios, N., Cross, J., Selbig, J., and Stitt, M., 2004. A robot-based platform to measure multiple enzyme activities in arabidopsis using a set of cycling assays: Comparison of changes of enzyme activities and transcript levels during diurnal cycles and in prolonged darkness. *Plant Cell*, 16(12):3304–3325.

Gibon, Y., Vigeolas, H., Tiessen, A., Geigenberger, P., and Stitt, M., 2002. Sensitive and high throughput metabolite assays for inorganic pyrophosphate, ADPGlc, nucleotide phosphates, and glycolytic intermediates based on a novel enzymic cycling system. *The Plant Journal*, 30(2):221–235.

Gorzsás, A., Stenlund, H., Persson, P., Trygg, J., and Sundberg, B., 2011. Cell-specific chemotyping and multivariate imaging by combined FT-IR microspectroscopy and orthogonal projections to latent structures (OPLS) analysis reveals the chemical landscape of secondary xylem. *The Plant Journal*, 66(5):903–14.

Granot, D., 2007. Role of tomato hexose kinases. *Functional Plant Biology*, 34(6):564–570.

Haigler, C.H., Ivanova-Datcheva, M., Hogan, P.S., Salnikov, V.V., Hwang, S., Martin, K., and Delmer, D.P., 2001. Carbon partitioning to cellulose synthesis. *Plant Molecular Biology*, 47:29–51.

Hakamori, S.I., 1964. A rapid permethylation of glycolipid, and polysaccharide catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide. *Journal of Biochemistry*, 55(2):205–208.

Hantao, L.W., Aleme, H.G., Pedroso, M.P., Sabin, G.P., Poppi, R.J., and Augusto, F., 2012. Multivariate curve resolution combined with gas chromatography to enhance analytical separation in complex samples: A review. *Analytica Chimica Acta*, 731:11–23.

Harfouche, A., Meilan, R., and Altman, A., 2011. Tree genetic engineering and applications to sustainable forestry and biomass production. *Trends in Biotechnology*, 29(1):9–17.

Harris, D.C., 2007. *Quantitative Chemical Analysis*. W.H. Freeman and Company, New York, USA, 7th edition, 533 pages.

Hatfield, R. and Vermerris, W., 2001. Lignin formation in plants. the dilemma of linkage specificity. *Plant Physiology*, 126(4):1351–1357.

Hedenström, M., Wiklund-Lindström, S., Oman, T., Lu, F.C., Gerber, L., Schatz, P., Sundberg, B., and Ralph, J., 2009. Identification of lignin and polysaccharide modifications in *Populus* wood by chemometric analysis of 2D NMR spectra from dissolved cell walls. *Molecular Plant*, 2(5):933–942.

Henriksson, G., Lawoko, M., Martin, M.E.E., and Gellerstedt, G.G., 2007. Lignin-carbohydrate network in wood and pulps: A determinant for reactivity. *Holzforschung*, 61:668–674.

Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A., Erlandsson, R., Blomqvist, K., Bhalerao, R., Uhlen, M., Teeri, T.T., Lundeberg, J., Sundberg, B., Nilsson, P., and Sandberg, G., 2001. A transcriptional roadmap to wood formation. *Proceedings of the National Academy of Sciences of the United States of America*, 98(25):14732–14737.

Hon, D.N.S. and Shiraishi, N., editors, 2001. *Wood and Cellulosic Chemistry*. Marcel Dekker, New York, 2nd edition, 914 pages.

Houtman, C.J. and Atalla, R.H., 1995. Cellulose-lignin interactions (a computational study). *Plant Physiology*, 107(3):977–984.

Hu, W.J., Harding, S.A., Lung, J., Popko, J.L., Ralph, J., Stokke, D.D., Tsai, C.J., and Chiang, V.L., 1999. Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nature Biotechnology*, 17.

Huang, H.J., Ramaswamy, S., Al-Dajani, W., Tschirner, U., and Cairncross, R.A., 2009. Effect of biomass species and plant size on cellulosic ethanol: A comparative process and economic analysis. *Biomass and Bioenergy*, 33(2):234 – 246.

Huntley, S., Ellis, D., Gilbert, M., Chapple, C., and Mansfield, S., 2003. Significant increases in pulping efficiency in C4H-F5H-transformed poplars: Improved chemical savings and reduced environmental toxins. *Journal of Agricultural and Food Chemistry*, 51(21):6178–6183.

Johnson, K.J., Wright, B.W., Jarman, K.H., and Synovec, R.E., 2003. High-speed peak matching algorithm for retention time alignment of gas chromatographic data for chemometric analysis. *Journal of Chromatography A*, 996(1–2):141–155.

Jonsson, P., Johansson, A.I., Gullberg, J., Trygg, J., A, J., Grung, B., Marklund, S., Sjöström, M., Antti, H., and Moritz, T., 2005. High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analyses. *Analytical chemistry*, 77(17):5635–42.

Kim, J. and Daniel, G., 2012. Distribution of glucomannans and xylans in poplar xylem and their changes under tension stress. *Planta*, pages 1–16.

Kohler, L. and Spatz, H., 2002. Micromechanics of plant tissues beyond the linear-elastic range. *Planta*, 215(1):33–40.

Kollmann, F.F.P., 1951. *Technologie des Holzes und der Holzwerkstoffe*, volume 1. Springer, 2 edition, 2233 pages.

Koshijima, T. and Watanabe, T., 2003. *Association Between Lignin and Carbohydrates in Wood and Other Plant Tissues*. Wood Science. Springer, Berlin, 329 pages.

Kumar, M., Thammannagowda, S., Bulone, V., Chiang, V., Han, K.H., Joshi, C.P., Mansfield, S.D., Mellerowicz, E., Sundberg, B., Teeri, T., and Ellis, B.E., 2009. An update on the nomenclature for the cellulose synthase genes in *Populus*. *Trends in Plant Science*, 14(5):248–254.

Lai, Y. and Sarkanen, K., 1971. Isolation and structural studies. In K. Sarkanen and C. Ludwig, editors, *Lignins. Occurrence, Formation, Structure and Reactions*, pages 165–240. Wiley Interscience, New York.

Lapierre, C., 2010. Determining lignin structure by chemical degradations. In C. Heitner, D. Dimmel, and J.A. Schmidt, editors, *Lignin and lignans: advances in chemistry*, pages 11–48. CRC Press, Taylor & Francis Group, Boca Raton, FL, USA.

Lapierre, C., Pollet, B., Petit-Conil, M., Toval, G., Romero, J., Pilate, G., Leplé, J., Boerjan, W., Ferret, V., De Nadai, V., and Jouanin, L., 1999. Structural alterations of lignins in transgenic poplars with depressed cinnamyl alcohol dehydrogenase or caffeic acid o-methyltransferase activity have an opposite impact on the efficiency of industrial kraft pulping. *Plant Physiology*, 119(1):153–163.

Lerouxel, O., Choo, T.S., Séveno, M., Usadel, B., Faye, L., Lerouge, P., and Pauly, M., 2002. Rapid structural phenotyping of plant cell wall mutants by enzymatic oligosaccharide fingerprinting. *Plant Physiology*, 130(4):1754–1763.

Li, X.Y. and Eriksson, L.A., 2005. Molecular dynamics study of lignin constituents in water. *Holzforschung*, 59(3).

Liang, Y.Z., Kvalheim, O.M., Rahmani, A., and Brereton, R.G., 1993. A two-way procedure for background correction of chromatographic/spectroscopic data by congruence analysis and least-squares fit of the zero-component regions: comparison with double-centering. *Chemometrics and Intelligent Laboratory Systems*, 18(3):265–279.

Lu, F. and Ralph, J., 1997. DFRC method for lignin analysis. 1. new method for β-aryl ether cleavage: Lignin model studies. *Journal of Agricultural and Food Chemistry*, 45(12):4655–4660.

Lu, F. and Ralph, J., 2003. Non-degradative dissolution and acetylation of ball-milled plant cell walls: high-resolution solution-state NMR. *Plant Journal*, 35(4):535–544.

Lu, F.C. and Ralph, J., 2011. Solution-state NMR of lignocellulosic biomass. *Journal of Biobased Materials and Bioenergy*, 5(2):169–180.

Lüdtke, M. and Liefländer, M., 1970. Quantitative Bestimmung der Aminosäuren im Wandprotein von pflanzlichen Faserzellen. *Holzforschung*, 24(2):33–37.

Lunn, J., Feil, R., Hendriks, J., Gibon, Y., Morcuende, R., Osuna, D., Scheible, W., Carillo, P., Hajirezaei, M., and Stitt, M., 2006. Sugar-induced increases in trehalose 6-phosphate are correlated with redox activation of ADPglucose pyrophosphorylase and higher rates of starch synthesis in *Arabidopsis thaliana*. *Biochemical Journal*, 397(1):139–148.

Ma, J., Zhang, Z., Yang, G., Mao, J., and Xu, F., 2011. Ultrastructural topochemistry of cell wall polymers in populus nigra by transmission electron microscopy and raman imaging. *Bioresources*, 6(4):3944–3959.

Mägdefrau, K., 1951. *Botanik. Eine Einfürung in das Studium der Pflanzenkunde*. Winter, Heidelberg, 239 pages.

Masuko, T., Minami, A., Iwasaki, N., Majima, T., Nishimura, S.I., and Lee, Y.C., 2005. Carbohydrate analysis by a phenol-sulfuric acid method in microplate format. *Analytical Biochemistry*, 339(1):69–72.

McNeil, M., Darvill, A., Fry, S., and Albersheim, P., 1984. Structure and function of the primary-cell walls of plants. *Annual Review of Biochmistry*, 53:625–663.

Mellerowicz, E.J., Baucher, M., Sundberg, B., and Boerjan, W., 2001. Unravelling cell wall formation in the woody dicot stem. *Plant Molecular Biology*, 47(1-2):239–274.

Mellerowicz, E.J. and Gorshkova, T.A., 2012. Tensional stress generation in gelatinous fibres: a review and possible mechanism based on cell-wall structure and composition. *Journal of Experimental Botany*, 63(2):551–565.

Merewether, J.W.T., 1957. A lignin-carbohydrate complex in wood. *Holzforschung*, 11(3):65–96.

Mitchell, R.A., Dupree, P., and Shewry, P.R., 2007. A novel bioinformatics approach identifies candidate genes for the synthesis and feruloylation of arabinoxylan. *Plant Physiology*, 144(1):43–53.

Moldoveanu, S.C., 1998. Analytical pyrolysis of natural organic polymers. *Techniques and Instrumentation in Analytical Chemistry*, 20:510.

Molhoj, M., Pagant, S., and Hofte, H., 2002. Towards understanding the role of membrane-bound endo-beta-1,4-glucanases in cellulose biosynthesis. *Plant and Cell Physiology*, 43(12):1399–1406.

Mouille, G., Robin, S., Lecomte, M., Pagant, S., and Höfte, H., 2003. Classification and identification of arabidopsis cell wall mutants using fourier-transform infrared (FT-IR) microspectroscopy. *The Plant journal*, 35(3):393–404.

Murakami, Y., Funada, R., Sano, Y., and Ohtani, J., 1999. The differentiation of contact cells and isolation cells in the xylem ray parenchyma of *Populus maximowiczii*. *Annals of Botany*, 84(4):429–435.

Mutwil, M., Debolt, S., and Persson, S., 2008. Cellulose synthesis: a complex complex. *Current Opinion in Plant Biology*, 11(3):252–257.

Nakano, Y., Nishikubo, N., Goue, N., Ohtani, M., Yamaguchi, M., Katayama, Y., and Demura, T., 2010. MYB transcription factors orchestrating the developmental program of xylem vessels in Arabidopsis roots. *Plant Biotechnology*, 27(3):267–272.

Neale, D. and Savolainen, O., 2004. Association genetics of complex traits in conifers. *Trends in Plant Science*, 9(7):325–330.

Nguyen, Q.B., Foyer, C.H., and Yelle, S., 1997. Overexpression of sucrose phosphate synthase increases sucrose unloading in transformed tomato fruits. *Plant Physiology*, 114(3):301–301.

Nielsen, N.P.V., Carstensen, J.M., and Smedsgaard, J., 1998. Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimised warping. *Journal of Chromatography A*, 805(1–2):17–35.

Nunes, C.A., Lima, C.F., Barbosa, L.C., Colodette, J.L., Gouveia, A., and Silvério, F.O., 2010. Determination of *Eucalyptus spp* lignin S/G ratio: A comparison between methods. *Bioresource Technology*, 101(11):4056–4061.

Obel, N., Erben, V., Schwarz, T., Kühnel, S., Fodor, A., and Pauly, M., 2009. Microanalysis of plant cell wall polysaccharides. *Molecular Plant*, 2(5):922–932.

Ona, T., Sonoda, T., Shibata, M., and Fukazawa, K., 1995. Small-scale method to determine the content of wood components from multiple eucalypt samples. *TAPPI Journal*, 78(3):121–126.

O'Neill, M.A., Ishii, T., Albersheim, P., and Darvill, A.G., 2004. Rhamnogalacturonan ii: Structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annual Review of Plant Biology*, 55(1):109–139.

Panshin, A. and de Zeeuw, C., 1980. *Textbook of Wood Technology*. McGraw-Hill, 772 pages.

Peña, M.J., Zhong, R., Zhou, G.K., Richardson, E.A., O'Neill, M.A., Darvill, A.G., York, W.S., and Ye, Z.H., 2007. *Arabidopsis irregular xylem8* and *irregular xylem9*: Implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell*, 19(2):549–563.

Peng, J., Lu, F., and Ralph, J., 1998. The DFRC method for lignin analysis. 4. lignin dimers isolated from DFRC-Degraded loblolly pine wood. *Journal of Agricultural and Food Chemistry*, 46(2):553–560.

Perry, C.C. and Currie, H.A., 2006. Resolution of complex monosaccharide mixtures from plant cell wall isolates by high pH anion exchange chromatography. *Journal of Chromatography A*, 1128(1-2):90–96.

Petrov, O.V. and Furo, I., 2009. NMR cryoporometry: Principles, applications and potential. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 54(2):97–122.

Pinto, R.C., Stenlund, H., Hertzberg, M., Lundstedt, T., Johansson, E., and Trygg, J., 2011. Design of experiments on 135 cloned poplar trees to map environmental influence in greenhouse. *Analytica Chimica Acta*, 685(2):127–131.

Pregitzer, K.S. and Euskirchen, E.S., 2004. Carbon cycling and storage in world forests: biome patterns related to forest age. *Global Change Biology*, 10(12):2052–2077.

Ralph, J. and Hatfield, R.D., 1991. Pyrolysis-Gc-Ms characterization of forage materials. *Journal of Agricultural and Food Chemistry*, 39(8):1426–1437.

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., and Boerjan, W., 2004. Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenyl- propanoids. *Phytochemistry Reviews*, 3:29–60.

Ralph, J., Peng, J., Lu, F., Hatfield, R., and Helm, R., 1999. Are lignins optically active? *Journal of Agricultural and food chemistry*, 47(8):2991–2996.

Ralph, S.G., 2009. Studying *Populus* defenses against insect herbivores in the post-genomic era. *Critical Reviews in Plant Sciences*, 28(5):335–345.

Ray, F.E., 2006. *Esau's Plant anatomy : meristems, cells, and tissues of the plant body: their structure, function, and development*. John Wiley & Sons, Inc., 3rd edition, 601 pages.

Renz, A. and Stitt, M., 1993. Substrate–specificity and product inhibition of different forms of fructokinases and hexokinases in developing potato–tubers. *Planta*, 190(2):166–175.

Rodrigues, J., Meier, D., Faix, O., and Pereira, H., 1999. Determination of tree to tree variation in syringyl/guaiacyl ratio of *Eucalyptus globulus* wood lignin by analytical pyrolysis. *Journal of Analytical and Applied Pyrolysis*, 48(2):121–128.

Rolando, C., Monties, B., and Lapierre, C., 1992. Thioacidolysis. In *Methods in Lignin Chemistry*, Springer Series in Wood Science. Springer.

Roussel, M.R. and Lim, C., 1995. Dynamic model of lignin growing in restricted spaces. *Macromolecules*, 28(1):370–376.

Rowe, J. and Connor, A., 1979. Extractives in eastern hardwoods: a review. Technical report, Forest Products Laboratory, Forest Service, U.S. Department of Agriculture.

Ruan, Y.L., Llewellyn, D.J., and Furbank, R.T., 2003. Suppression of sucrose synthase gene expression represses cotton fiber cell initiation, elongation, and seed development. *Plant Cell*, 15(4):952–964.

Salmén, L. and Burgert, I., 2008. Cell wall features with regard to mechanical performance. a review cost action e35 2004-2008: Wood machining - micromechanics and fracture. *Holzforschung*, 63(2):121–129.

Sannigrahi, P., Ragauskas, A.J., and Tuskan, G.A., 2010. Poplar as a feedstock for biofuels: A review of compositional characteristics. *Biofuels Bioproducts & Biorefining*, 4(2):209–226.

Scheffer, T.C., 1966. Natural resistance of wood to microbial deterioration. *Annual Review of Phytopathology*, 4(1):147–168.

Scheller, H.V. and Ulvskov, P., 2010. Hemicelluloses. In S. Merchant, W.R. Briggs, and D. Ort, editors, *Annual Review of Plant Biology*, volume 61 of *Annual Review of Plant Biology*, pages 263–289. Annual Reviews.

Schilmiller, A.L., Stout, J., Weng, J.K., Humphreys, J., Ruegger, M.O., and Chapple, C., 2009. Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in *Arabidopsis*. *The Plant Journal*, 60(5):771–782.

Seifert, G.J. and Roberts, K., 2007. The biology of arabinogalactan proteins. *Annual Review of Plant Biology*, 58:131–161.

Sewalt, V., Ni, W., Blount, J.W., Jung, H.G., Masoud, S.A., Howles, P.A., Lamb, C., and 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., Reddy, M.S., Jackson, L., Nakashima, J., and Dixon, R.A., 2007. Down-regulation of hydroxycinnamoyl CoA: Shikimate hydroxycinnamoyl transferase in transgenic *alfalfa* affects lignification, development and forage quality. *Phytochemistry*, 68(11):1521–1529.

Sharples, S.C. and Fry, S.C., 2007. Radioisotope ratios discriminate between competing pathways of cell wall polysaccharide and RNA biosynthesis in living plant cells. *The Plant Journal*, 52(2):252–262.

Shmulsky, R. and Jones, D.P., 2011. *Forest Products and Wood Science – An Introduction*. Wiley-Blackwell, 496 pages.

Siegel, S.M., 1957. Non-enzymic macromolecules as matrices in biological synthesis: The role of polysaccharides in peroxidase-catalyzed lignin polymer formation from eugenol. *Journal of the American Chemical Society*, 79(7):1628–1632.

Simson, B.W. and Timell, T.E., 1978. Polysaccharides in cambial tissues of *Populus tremuloides* and *Tilia americana. Cellulose Chemistry and Technology*, 12(1):39–84.

Sjöström, E., 1993. *Wood chemistry: fundamentals and applications*. Academic Press, 293 pages.

Smith, C., Want, E., O'Maille, G., Abagyan, R., and Siuzdak, G., 2006. XCMS: Processing mass spectrometry data for metabolite profiling using Nonlinear peak alignment, matching, and identification. *Analytical Chemistry*, 78(3):779–787.

Stein, S., 1999. An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. *Journal of The American Society for Mass Spectrometry*, 10(8):770–781.

Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P., Brunner, A.M., Charbonnel-Campaa, L., Lindvall, J.J., Tandre, K., Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlén, M., Bhalerao, R.P., Nilsson, O., Sandberg, G., Karlsson, J., Lundeberg, J., and Jansson, S., 2004. A populus est resource for plant functional genomics. *Proceedings of the National Academy of Sciences of the United States of America*, 101(38):13951–13956.

Studer, M.H., DeMartini, J.D., Davis, M.F., Sykes, R.W., Davison, B., Keller, M., Tuskan, G.A., and Wyman, C.E., 2011. Lignin content in natural *Populus* variants affects sugar release. *Proceedings of the National Academy of Sciences of the United States of America*, 108(15):6300–6305.

Sturm, A. and Tang, G.Q., 1999. The sucrose-cleaving enzymes of plants are crucial for development, growth and carbon partitioning. *Trends in Plant Science*, 4(10):401–407.

Sun, L., Varanasi, P., Yang, F., Loqué, D., Simmons, B.A., and Singh, S., 2012. Rapid determination of syringyl: Guaiacyl ratios using FT-Raman spectroscopy. *Biotechnology and Bioengineering*, 109(3):647–656.

Sykes, R., Kodrzycki, B., Tuskan, G., Foutz, K., and Davis, M., 2008. Within tree variability of lignin composition in *Populus*. *Wood Science and Technology*, 42(8):649–661.

Syrjänen, K. and Brunow, G., 2001. Regioselectivity in oxidative cross-coupling of phenols. application to the synthesis of dimeric neolignans. *Tetrahedron*, 57(2):365–370.

Talbott, L. and Ray, P., 1992. Molecular-size and separability features of pea cell-wall polysaccharides - implications for models or primary wall structure. *Plant Physiology*, 98(1):357–368.

Tauler, R., 1995. Multivariate curve resolution applied to second order data. *Chemometrics and Intelligent Laboratory Systems*, 30(1):133–146.

Terashima, N., 1990. A new mechanism for formation of a structurally ordered protolignin macromolecule in the cell wall of tree xylem. *Journal of Pulp and Paper Science*, 16(5).

Tullus, A., Rytter, L., Tullus, T., Weih, M., and Tullus, H., 2012. Short–rotation forestry with hybrid aspen (Populus tremula L. × P. tremuloides Michx.) in northern europe. *Scandinavian Journal of Forest Research*, 27(1):10–29.

Turgeon, R., 1996. Phloem loading and plasmodesmata. *Trends in Plant Science*, 1(12):418–423.

Tuskan, G., West, D., Bradshaw, H.D., Neale, D., Sewell, M., Wheeler, N., Megraw, B., Jech, K., Wiselogel, A., Evans, R., Elam, C., Davis, M., and Dinus, R., 1999. Two high-throughput techniques for determining wood properties as part of a molecular genetics analysis of hybrid poplar and loblolly pine. *Applied Biochemistry and Biotechnology*, 77-9:55–65.

Tuskan, G.A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Dejardin, A., dePamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B.,

Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leplé, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y., and Rokhsar, D., 2006. The genome of black cottonwood, populus trichocarpa (torr. & gray). *Science*, 313(5793):1596–1604.

Updegraff, D.M., 1969. Semimicro determination of cellulose inbiological materials. *Analytical Biochemistry*, 32(3):420–424.

van Stokkum, I.H., Mullen, K.M., and Mihaleva, V.V., 2009. Global analysis of multiple gas chromatography-mass spectrometry (GC/MS) data sets: A method for resolution of co-eluting components with comparison to MCR-ALS. *Chemometrics and Intelligent Laboratory Systems*, 95(2):150 – 163.

Vanholme, R., Morreel, K., Ralph, J., and Boerjan, W., 2008. Lignin engineering. *Current Opinion in Plant Biology*, 11(3):278–285.

Wampler, T.P., editor, 2007. *Applied Pyrolysis handbook*. CRC Press, Taylor & Francis Group, Boca Raton, Florida, USA, 288 pages.

Wolf, S., Mouille, G., and Pelloux, J., 2009. Homogalacturonan methylesterification and plant development. *Molecular Plant*, 2(5):851–860.

Yamaguchi, M., Goue, N., Igarashi, H., Ohtani, M., Nakano, Y., Mortimer, J.C., Nishikubo, N., Kubo, M., Katayama, Y., Kakegawa, K., Dupree, P., and Demura, T., 2010. VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 Effectively Induce Transdifferentiation into Xylem Vessel Elements under Control of an Induction System. *Plant Physiology*, 153(3):906–914.

Yamaguchi, M., Kubo, M., Fukuda, H., and Demura, T., 2008. VASCULAR-RELATED NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in Arabidopsis roots and shoots. *Plant Journal*, 55(4):652–664.

Yang, X., Ye, C.Y., Bisaria, A., Tuskan, G.A., and Kalluri, U.C., 2011. Identification of candidate genes in *Arabidopsis* and *Populus* cell wall biosynthesis using text-mining, co-expression network analysis and comparative genomics. *Plant Science*, 181(6):675–687.

Zaia, J., 2004. Mass spectrometry of oligosaccharides. *Mass Spectrometry Reviews*, 23(3):161–227.

Zhao, Q. and Dixon, R.A., 2011. Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends in Plant Science*, 16(4):227–233.

Zhao, Q., Wang, H., Yin, Y., Xu, Y., Chen, F., and Dixon, R.A., 2010. Syringyl lignin biosynthesis is directly regulated by a secondary cell wall master switch. *Proceedings of the National Academy of Sciences*, 107(32):14496–14501.

Zhong, R., Lee, C., Zhou, J., McCarthy, R.L., and Ye, Z.H., 2008. A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell*, 20(10):2763–2782.

Zrenner, R., Salanoubat, M., Willmitzer, L., and Sonnewald, U., 1995. Evidence of the crucial role of sucrose synthase for sink strength using transgenic potato plants (*Solanum-Tuberosum L*). *Plant Journal*, 7(1):97–107.

Zugenmaier, P., 2008. *Crystalline Cellulose and Cellulose Derivatives*. Springer Series in Wood Science. Springer, 285 pages.