THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Enzymes for selective decoupling of woody biomass

From fundamentals to industrial potential

JENNY ARNLING BÅÅTH

Department of Biology and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Enzymes for selective decoupling of woody biomass
From fundamentals to industrial potential

JENNY ARNLING BÅÅTH

ISBN 978-91-7597-848-2
© Jenny Arnling Bååth, 2019

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 4529
ISSN 0346-718X

Division of Industrial Biotechnology
Department of Biology and Biological Engineering
Chalmers University of Technology
SE-412 96 GOTHENBURG
Sweden
Telephone: + 46 (0) 31 772 10 00

Cover: Wood polymers in a forest landscape.

Printed by Chalmers Reproservice,
Gothenburg, Sweden 2019

"As scientists we have to be prepared to let go all the fancy theories and hypotheses that we generate to explain things. In other words: reset our minds. So, from this angle, a reset is a proper scientific approach."

-Detlef Jensen

Enzymes for selective decoupling of woody biomass
From fundamentals to industrial potential

JENNY ARNLING BÅÅTH

Division of Industrial Biotechnology
Department of Biology and Biological Engineering
Chalmers University of Technology

ABSTRACT

The need for an economy based on renewable materials has resulted in growing interest in the use of woody biomass in a wider field of application. However, the chemical complexity of lignocellulose and the dense structure of wood pose challenges in its processing. The aim in a materials biorefinery is to extract the individual wood-components in as native and intact form as possible; and highly specific enzymes could be used for this. In this work, lignocellulosic enzymes with the potential of selectively decoupling and decomposing wood polymers were investigated and characterised.

One of the studies was devoted to assessing enzyme accessibility in wood structures and evaluating the importance of pre-treatment to open up the dense wood structure, in order to improve enzyme performance. It was shown that the physical contribution during steam explosion is crucial for enhanced enzymatic hydrolysis of wood. Moreover, the performance of two *endo*-mannanases produced by the bacterium *Cellvibrio japonicus* (*Cj*Man5A and *Cj*Man26A) were compared on mannan polymers with dissimilar backbone structures and decorations, including the industrially relevant spruce galactoglucomannan. The enzymes were shown to be differently affected by the backbone heterogeneity and the presence of side groups on the substrates, demonstrating the variation in substrate preferences among mannanases. It was further shown that chemical acetylation of mannans reduced substrate hydrolysis significantly. Acetylation was therefore suggested as a tool to limit the biodegradation of mannan-based material.

The major part of the work described in this thesis was dedicated to investigating the role and function of glucuronoyl esterases (GEs), which are enzymes that hydrolyse the ester linkages between lignin and hemicelluloses that contribute to the recalcitrance of woody biomass. Extensive structure-function studies of bacterial GE candidates contributed to our understanding of the diversity of this relatively unexplored enzyme family. Both similarities and differences in substrate preferences among the GEs studied revealed enzymes more promiscuous than their characterised fungal counterparts. GE activity was further assayed on lignin-carbohydrate complexes isolated from woody biomass, and GE-mediated ester cleavage was demonstrated with advanced and complementary tools, including size-exclusion chromatography, ^{31}P NMR and 2D NMR. These findings not only confirm the suggested biological role of GEs, but also demonstrate the potential of these enzymes in decoupling lignin from hemicelluloses in industrial settings.

Keywords: materials biorefinery, lignocellulose, woody biomass, carbohydrate esterase family 15, glucuronoyl esterase, lignin-carbohydrate bonds, mannanase, galactoglucomannan, polysaccharide acetylation, enzyme accessibility

Preface

This doctoral thesis partially fulfils the requirements for a PhD degree at the Department of Biology and Biological Engineering, Chalmers University of Technology, Sweden. The work presented in this thesis was performed between 2014 and 2019, and was funded by the Knut and Alice Wallenberg Foundation within the Wallenberg Wood Science Center (WWSC). The WWSC is a joint research centre involving Chalmers University of Technology and the Royal Institute of Technology (KTH) with the mission "to create knowledge and build competence that can form the basis for an innovative future value creation from forest raw material". The main part of the work was carried out at the Division of Industrial Biotechnology at Chalmers University of Technology under the supervision of Professor Lisbeth Olsson and Dr Johan Larsbrink.

Part of the research, that on the enzymatic cleavage of mannan polymers and lignin-carbohydrate complexes (LCCs), was performed in collaboration with Drs Francisco Vilaplana and Martin Lawoko at KTH. Initial experiments on the detection of glucuronoyl esterase activity on isolated LCCs were performed at Toronto University in collaboration with Professor Emma Master and at Kyoto University in collaboration with Professor Takashi Watanabe and Dr Hiroshi Nishimura. Fluorescence recovery after photobleaching was studied at the Product Design and Perception Section at RISE (Research Institutes of Sweden) with the help of Professor Niklas Lorén and Dr Annika Krona.

Jenny Arnling Bååth
January 2019

List of publications

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Muzamal M, **Arnling Bååth J**, Olsson L and Rasmuson A. (2016) Contribution of structural modification to enhanced enzymatic hydrolysis and 3-D structural analysis of steam-exploded wood using X-ray tomography. *Bioresources,* 11(4): 8509-8521.

II **Arnling Bååth J**, Martínez-Abad A, Berglund J, Larsbrink J, Vilaplana F and Olsson L. (2018) Mannanase hydrolysis of spruce galactoglucomannan focusing on the influence of acetylation on enzymatic mannan degradation. *Biotechnology for Biofuels*, 11: 114.

III **Arnling Bååth J***, Giummarella N*, Klaubauf S, Lawoko M and Olsson L. (2016) A glucuronoyl esterase from *Acremonium alcalophilum* cleaves native lignin-carbohydrate ester bonds. *FEBS Letters,* 590: 2611-2618.

IV **Arnling Bååth J***, Mazurkewich S*, Meland Knudsen R, Navarro Poulsen J, Olsson L, Lo Leggio L and Larsbrink J. (2018) Biochemical and structural features of diverse bacterial glucuronoyl esterases facilitating recalcitrant biomass conversion. *Biotechnology for Biofuels*, 11: 213.

V **Arnling Bååth J**, Mazurkewich S, Navarro Poulsen J, Olsson L, Lo Leggio L and Larsbrink J. Biochemical and structural characterization of the glucuronoyl esterase *Tt*CE15A from *Teredinibacter turnerae* reveals interactions with a variety of carbohydrate and aromatic compounds. Manuscript in preparation.

*These authors contributed equally

Reprints were made with permission from the respective publishers.

Author's contributions

Paper I. Second author. I planned and performed the enzymatic hydrolysis and the HPAEC-PAD analysis, I interpreted the hydrolysis data and wrote minor parts of the manuscript.

Paper II. First author. I conceived the study, and planned and performed the enzymatic hydrolysis, the quantitative measurements on the reducing sugars, and the HPAEC-PAD analysis. I identified and quantified the reaction products and was the main author.

Paper III. First author (shared). I conceived the study, designed the experiments, performed the enzymatic reactions, analysed the data and prepared parts of the manuscript.

Paper IV. First author (shared). I participated in conceiving the project. I carried out the transcriptional analysis and the enzyme boosting study, produced and characterised all of the enzymes, participated in solving the *Su*CE15C structure and prepared parts of the manuscript.

Paper V. First author. I conceived the study and planned the experimental design. I produced and characterised *Tt*CE15A, created enzyme mutations, performed inhibition studies, interpreted the data and was the main author.

Abbreviations

4-*O*-MeGlcA	4-*O*-methylglucuronic acid
AA	auxiliary activities
AllylGlcA	allyl glucuronoate
Ara	arabinose
BnzGlcA	benzyl glucuronoate
CAZy	carbohydrate-active enzymes database
CAZymes	carbohydrate-active enzymes
CBM	carbohydrate-binding module
CE	carbohydrate esterase
CE15	carbohydrate esterase family 15
D	diffusion constant
DNSa	3, 5-dinitrosalicylic acid
DP	degree of polymerisation
DSac	degree of substitution by acetylation
ESI-MS	electrospray-ionisation mass spectrometry
FRAP	fluorescence recovery after photobleaching
Gal	galactose
GalA	galacturonic acid/galacturonoate
GE	glucuronoyl esterase
GFP	green fluorescent protein
GGM	galactoglucomannan
GH	glycoside hydrolase
GH5	glycoside hydrolase family 5
GH26	glycoside hydrolase family 26
Glc	glucose
GlcA	glucuronic acid/glucuronoate
HPAEC-PAD	high-performance anion-exchange chromatography with pulsed amperometric detection
HPLC	high-performance liquid chromatography
HSQC	heteronuclear single quantum coherence spectroscopy
KGM	konjac glucomannan
LBG	locust bean gum galactomannan
LC	lignin-carbohydrate
LCC	lignin-carbohydrate complex
Man/M1	mannose
M2	mannobiose
M3	mannotriose
M4	mannotetraose
MeGalA	methyl galacturonoate
MeGlcA	methyl glucuronoate
MS	mass spectrometry
NMR	nuclear magnetic resonance
PDB	protein data bank
*p*NP	4-nitrophenol

*p*NP-Ac	4-nitrophenyl acetate
PUL	polysaccharide utilisation locus
PULDB	polysaccharide utilisation loci database
ROI	region of interest
SEC	size-exclusion chromatography
SpGGM	spruce galactoglucomannan
STEX	steam explosion
TLC	thin-layer chromatography
UDH	uronate dehydrogenase
Xyl	xylose

Contents

1. Introduction

1.1 Forest as a resource

The massive environmental and resource-related challenges facing society necessitate the transformation from a linear and fossil-based economy into a circular bioeconomy, where plant biomass (lignocellulose) is used as a resource to replace petroleum and to mitigate climate change. Lignocellulose has a heterogeneous chemical composition, which makes it promising as a raw material for complex products of high value, such as biochemicals and materials in a biorefinery where all its constituents can be used. There are local, national and global driving forces for improvements in economic and environmental sustainability (Esselin, 2017). In the Nordic countries, including Sweden, wood constitutes the major source of plant biomass, and the development of forest products has been identified as important to ensure economic growth. The wood-based materials biorefinery is therefore expected to be a keystone in Sweden's bioeconomy (Esselin, 2017; Persson, 2016). This thesis presents studies on lignocellulose, with the focus on woody biomass.

1.2 The biorefinery concept

In a biorefinery, plant biomass is converted by integrated processes into materials, chemicals, fuels, heat and energy carriers. The aims are to upgrade the biomass, to use all of its components, and to minimise waste using efficient and sustainable processes (Berntsson et al., 2013). Woody biomass, with its chemical complexity, has a particularly high potential as a raw material for the production of materials and value-added chemicals. In a materials biorefinery, the ambition is to keep the polymeric wood components (cellulose, hemicelluloses and lignin) as intact as possible (Figure 1). Selective and efficient process steps are of key importance in avoiding complete degradation of the polymers.

A number of high-value products can be derived from the polymeric and oligomeric forms of the components of wood. Cellulose can be extracted as pulp fibres, or used in textiles, but it could also be further processed into nanocellulose with suitable properties for composites (Grishkewich et al., 2017). Lignin is often combusted to produce heat and electricity, but should be considered as a raw material for the production of materials such as carbon fibres, adhesives and activated carbon (Norgren & Edlund, 2014). Hemicelluloses have traditionally been an underutilised wood component, mainly being degraded into fermentable monosaccharides. However, this class of polysaccharides has considerable potential in the polymeric form for the production of films, coatings and gels, and in the oligomeric form as prebiotics and food additives (Ebringerová, 2006; Zhang et al., 2013).

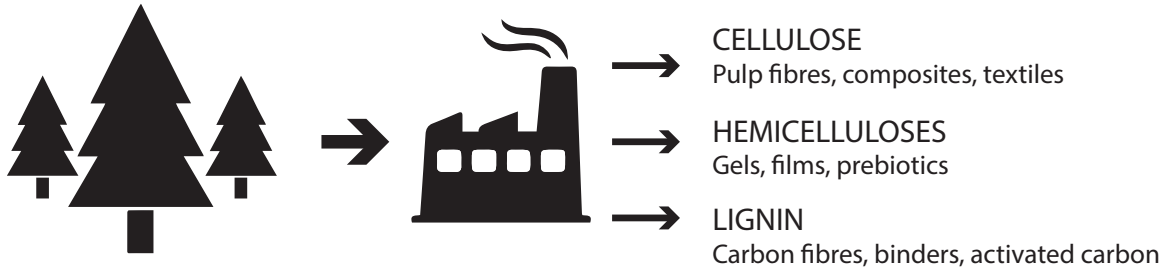


Figure 1. The wood-based materials biorefinery concept, in which the polymeric and oligomeric wood components are separated into different streams for further modification and upgrading. Examples of high-value applications are listed under each wood component.

1.3 Enzymes in a wood-based biorefinery

The density, complex structure and heterogeneous chemistry of woody biomass not only makes it a promising resource, but also recalcitrant and difficult to process. In order to extract and separate polymers and oligomers for use as high-value building blocks for various applications, mild pre-treatment and targeted separation techniques must be developed. Enzymes could play a key role in these processes.

Benefits of using enzymes

Enzymes are capable of catalysing highly specific reactions under ambient temperatures and moderate pH values. They are specific, and their capability to act under relatively mild conditions contributes to sustainable industrial processes without unwanted side reactions. A range of lignocellulose-degrading fungi and bacteria that produce enzymes necessary to completely degrade the polymers in woody biomass can be found in nature. Such microorganisms are promising sources of the specific enzymes that will be required in a wood-based biorefinery.

Although some enzymes are efficient in nature, several challenges must be overcome when choosing enzymes and applying them in industrial settings. Firstly, the enzyme must be compatible with the industrial environment (e.g. temperature, pH, pressure, chemical conditions) and secondly, it must be possible to produce in high amounts. Furthermore, the complexity and density of woody biomass cause problems associated with enzyme accessibility and unproductive binding to lignin, reducing the efficiency of the processes.

Enzymes for different end products

Woody biomass can be degraded, modified and upgraded in different ways depending on the choice of enzyme or enzymes. This allows for controlled and specific design of the end products. Different enzymatic approaches are listed in Box 1.

Box 1. Industrial approaches for enzymatic decomposition, modification and upgrading of woody biomass

1. Complete decomposition of wood polysaccharides into fermentable monosaccharides.
2. Partial decomposition of wood polysaccharides into oligosaccharides, to create oligomeric building blocks for the production of prebiotics and food additives.
3. Separation of lignin and hemicelluloses, to enable the extraction of polymeric building blocks.
4. Polishing and removal of side chains from branched hemicelluloses to create uniform polymeric building blocks.
5. Functionalisation of oligosaccharides and polysaccharides.
6. Upgrading of biomass-derived molecules.

Approaches 2 and 3 were the subject of this thesis.

From fundamentals to industrial potential

The forest industry faces several challenges in the transition to material-driven biorefineries. Research is therefore required on potential high-value products and materials, as well as energy-efficient processes, including the use of enzymes in these procedures. In this thesis, three industrially relevant aspects are considered: the accessibility of wood-derived substrates to enzymes, the way in which hemicelluloses, especially mannan polymers, can be enzymatically processed, and the possible enzymatic decoupling of lignin from hemicelluloses. Before enzymes can be used in wood processing, it is necessary to elucidate where and how they act on biologically and industrially relevant substrates. In this work, pre-treated wood chips were used for enzyme accessibility studies, and wood-derived lignin-carbohydrate complexes and galactoglucomannan for studies on enzymatic cleavage. Assaying enzymes on complex woody biomass is often complicated due to the lack of appropriate analytical tools. The evaluation and establishment of suitable detection methods was therefore an important part of this work.

Due to the difficulty in assaying the action of enzymes on lignocellulosic materials, research is often carried out on simpler model substrates, intended to mimic the native substrates. In this work, several model substrates were used, including small synthetic monomers and well-defined, commercially available polymers extracted from other types of plant material than wood. Highly accurate kinetic parameters can often be measured on these models and information on substrate structures important for enzyme specificity can be obtained.

Functional enzyme studies in combination with known enzyme structures further make it possible to elucidate enzyme-substrate interactions on an atomic level. Structure-function studies on glucuronoyl esterases (GEs) were a major part of this work, and led to new detailed insight on substrate specificities and the way in which these enzymes bind substrates in their active site. These types of studies are important in understanding the nature of the substrates on which the enzyme acts, and to provide information on the diversity between enzymes targeting the glucuronoyl ester bond. Fundamental research is not only of scientific interest, but also necessary to identify the enzymes best suited to act on specific lignocellulosic materials under the prevailing process conditions.

1.4 Aims and outline

The main objective of the work presented in this thesis was to understand, describe and discuss the potential of specific enzymes that could be used in a wood-based biorefinery for the production of value-added polymeric and oligomeric components. In the first part of the thesis, I summarise the research results and put them in a larger context. The second part of the thesis consists of scientific papers, addressing the specific aims discussed below.

One of the aims was to investigate **how the accessibility of the wood substrate to enzymes is influenced by biomass pre-treatment** (Chapter 4). The impact of the physical contribution of steam explosion (STEX) was examined by comparing the effects of STEX and steam treatment (without the explosion step) on the enzymatic hydrolysability of softwood (**Paper I**). Fluorescence recovery after photobleaching (FRAP) was then investigated as a means of evaluating enzyme accessibility in native and pre-treated wood chips. The results of this study are presented, although they were not conclusive, and the future potential of FRAP as a detection method for enzyme accessibility is briefly discussed.

The second aim was to **evaluate *endo*-mannanases for selective processing of galactoglucomannan (GGM),** the major hemicellulose in softwood (Chapter 6). The lack of information on the way in which mannanases from different enzyme families act on heterogeneous, complex mannan polymers was addressed in the study described in **Paper II**. The action of two mannanases from the bacterium *Cellvibrio japonicus* were compared on mannan polymers with dissimilar backbone structures, and with different degrees of galactosylation and acetylation, in order to expand our knowledge on substrate specificity. Substrate hydrolysability was quantified and mannooligosaccharide products were detected using various analytical tools to study the effect of the substrate structure on the enzymatic performance. Chemically acetylated mannan substrates were included in the study to obtain deeper insight into the influence of acetylation on degradation by mannanases.

The third major aim was to **detect and understand the enzymatic decoupling of lignin and hemicelluloses** (Chapter 7). This part of the work provided results from research on GEs (**Papers III, IV, V**). The ability of a GE to cleave native lignin-carbohydrate (LC) ester linkages was evaluated by using complementary detection tools (**Paper III**). This study was conducted to demonstrate GE activity on wood-derived material, which had previously only been confirmed on synthetic substrates. Unpublished results from trials using additional analytical approaches for the detection of GE-mediated LC ester hydrolysis are included and discussed in the end of Chapter 7.

A substantial part of Chapter 7 describes the detailed characteristics of several novel GEs. The small number of GEs previously studied, and the limited information available on bacterial enzymes motivated the study reported in **Paper IV**, where ten putative and distantly related GEs of bacterial origin were investigated. Biochemical characterisation on model substrates and structure-function studies were conducted to expand our knowledge concerning the diversity and similarities between these bacterial GE candidates. Gene expression analysis and studies on synergistic actions with other lignocellulolytic enzymes in saccharification processes were included to investigate the biological function of these enzymes. **Paper V** describes the detailed biochemical and structural characterisation of a GE from the bacterium *Teredinibacter turnerae*. In an attempt to gain a more detailed understanding of enzyme-substrate specificity, the interaction of this GE to a variety of LC-mimicking compounds was investigated through genetic mutation and inhibition studies.

In the studies described in this thesis, enzymes were evaluated on native biomass, extracted wood polymers and synthetic model substrates, using a range of detection methods. Identifying and developing suitable and complementary analytical tools was crucial, and constituted a substantial part of the work. Chapter 5 is therefore dedicated to the methods that were used for the detection of enzyme activity, including the advantages and limitations of the different approaches.

2. Lignocellulose

Lignocellulose is the collective name of the polymeric, main carbon-based constituents of plant cell walls: cellulose, hemicelluloses and lignin. Chemical components useful for industrial production can be obtained by breaking these polymers into their constituent building blocks (Sanderson, 2011). A variety of industrially relevant lignocellulose sources are available in large quantities, or as waste products, e.g. trees (hard- and softwood species) and grasses (e.g. rice, corn, wheat straw, sugar cane). The lignocellulose content and proportions differ substantially, both quantitatively and qualitatively, depending on the source (Pauly & Keegstra, 2008). During the course of this work, various types of lignocellulosic material were used to evaluate the enzymes, but since the overall objective of this thesis is to describe enzymes for a wood-based biorefinery, this chapter focuses on the chemistry of wood. According to Fengel & Wegener, *"Wood is not merely a chemical substance, an anatomical tissue, or a material – it is a combination of all three…"* (Fengel & Wegener, 1989), which means that chemistry, composition and ultra-structure must be discussed in relation to each other.

2.1 From tree to wood cell

Wood is classified as softwood or hardwood, which refer to gymnosperms (coniferous trees) and angiosperms (deciduous trees), respectively. Typical hardwood species in the northern hemisphere are beech, birch and poplar, while spruce, pine and fir are common softwood species. Wood is composed of longitudinally and radially ordered cells (often called fibres) of different types and sizes. Most of the cells are tube-shaped with a hollow void (lumen) providing water transport. Softwood has a simpler structure than hardwood, with a lower number of cell types and more uniform cells (Fengel & Wegener, 1989).

All cells have a thick cell wall organised in layers, which are "glued" to each other via the lignin-rich middle lamella region. The outermost layer is the primary cell wall (P) followed by the three layers of the secondary cell wall (S1, S2 and S3), where S3 surrounds the lumen. All layers differ in their chemical composition, thickness and cellulose microfibril orientation (Daniel, 2009). The S2 layer is by far the thickest, accounting for 80 % of the cell wall, and is therefore chemically most important in an industrial context. It consists mainly of cellulose, followed by hemicelluloses and lignin with the micro-fibrils oriented at a slight vertical angle (Figure 2) (Gibson, 2012). The ultra-structure of the different cell wall layers is described in detail in Chapter 4, in relation to enzyme accessibility in wood structures (**Paper I**).

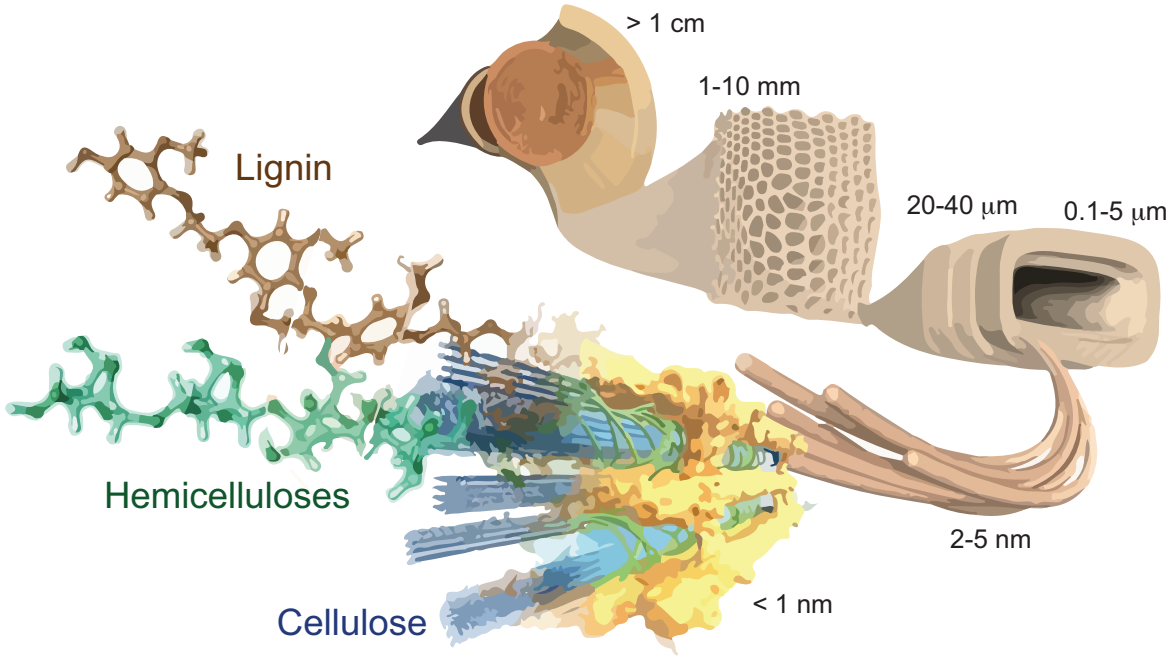


Figure 2. The hierarchical structure and scales of wood showing the stem, wood cells, the cell wall layers and the arrangement of lignin and hemicelluloses around the organised cellulose microfibrils. (Figure courtesy of Nishimura et al., 2018.)

2.2 Chemical composition of wood

Lignocellulosic plant biomass, including wood, consists of three main polymers: cellulose (40-50 %), hemicelluloses (20-40 %) and lignin (20-30 %) (Figure 3) (Sjöström, 1993). Minor components include pectins (mainly in the primary cell wall), proteins, lipids and inorganic components such as minerals (usually referred to as ash). The chemical composition differs both qualitatively and quantitatively, depending on the type of wood and the sample within a wood species, (Pauly & Keegstra, 2008).

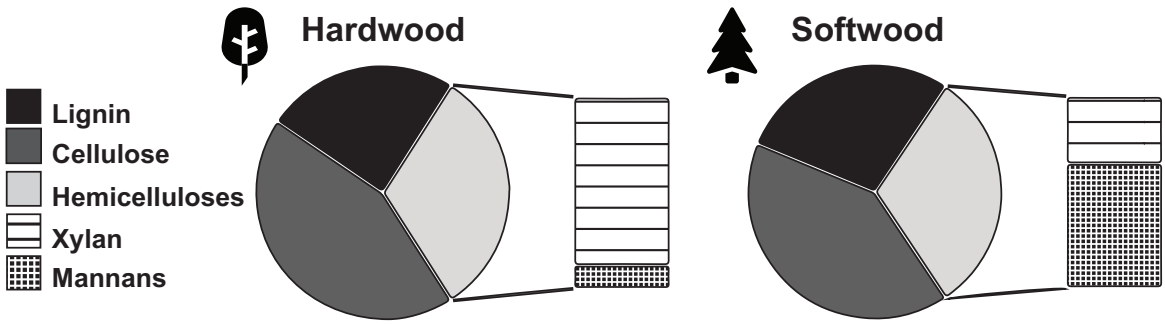


Figure 3. The chemical composition of hardwood (beech) and softwood (spruce) according to Fengel & Wegener, (1989). Wood consists mainly of cellulose, followed by hemicelluloses and lignin. The major difference between hardwood and softwood is the hemicellulose composition, predominantly represented by xylan in hardwood species and by mannans in softwood species.

Cellulose

The major component of wood is cellulose. Cellulose is a linear homopolysaccharide made up of chains of β-1,4-linked glucose units (Figure 4). Every second glucose unit in the cellulose chain is rotated 180°, and the disaccharide cellobiose is thus the smallest repeating

unit. The degree of polymerisation (DP) is high, between 5000 and 15 000 repeating units, depending on the source and extraction method employed. Cellulose polymers are aligned in a parallel manner, and held together with hydrogen bonds in so-called microfibrils. The microfibrils consist of highly ordered and crystalline regions and amorphous regions that are less organised. The microfibrils are bundled into fibrils that form the cellulose fibres (Henriksson & Lennholm, 2009; Sjöström, 1993).

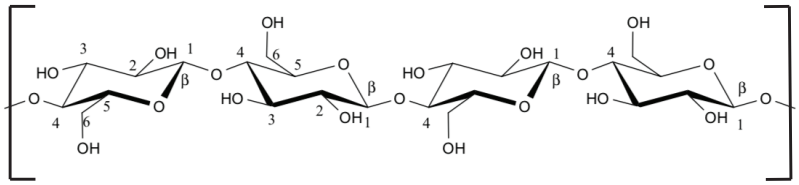


Figure 4. Chemical structure of cellulose, made up of β-1,4-D-glucan. Since every second glucose unit is rotated 180°, cellobiose is the smallest repeating unit.

Hemicelluloses

Hemicelluloses are a group of branched polysaccharides with large variation in chemical composition and abundancy among different plant species and cell types. They are heteropolymers with backbones of different monosaccharide units, named after the major sugar component. Hemicelluloses are embedded around the cellulose fibrils, providing contact between them, which contributes to the strength of the cell wall. They are also, to some extent, covalently linked to lignin (Scheller & Ulvskov, 2010). The two major hemicelluloses in woody biomass are different variants of xylans and mannans, found in the secondary cell wall layers (Table 1). Smaller amounts of other types of hemicelluloses, including glucans, galactans and arabinans, are found in certain wood species. Xyloglucan is the major hemicellulose in all primary cell walls, but is still only a minor component of the wood tissue due to the considerably thicker secondary wall (Fengel & Wegener, 1989; Telemann, 2009).

Table 1. The major wood hemicelluloses and their chemical composition (Sjöström, 1993)

Hemicellulose type	Wood type	Amount %*	Units	Molar ratio	Linkages
Galactoglucomannan	Softwood	5-8	Man	3	β-1,4
			Glc	1	β-1,4
			Gal	1	α-1,6
			O-Acetyl	1	
(Galacto)glucomannan	Softwood	10-15	Man	4	β-1,4
			Glc	1	β-1,4
			Gal	0.1	α-1,6
			O-Acetyl	1	
Arabinoglucuronoxylan	Softwood	7-15	Xyl	10	β-1,4
			4-*O*-MeGlcA	2	α-1,2
			Ara	1.3	α-1,3
Glucomannan	Hardwood	2-5	Man	1.5	β-1,4
			Glc	1	β-1,4
			O-Acetyl	1	
Glucuronoxylan	Hardwood	15-35	Xyl	10	β-1,4
			4-*O*-MeGlcA	1	α-1,2
			O-Acetyl	7	

*By dry weight

Xylan is the major hemicellulose in hardwood (15-35 %, w/w), present as *O*-acetyl-(4-*O*-methylglucurono)-xylan with reported DP between 100 and 200. Xylan consists of a main

chain of β-1,4-D-xylan, substituted by 4-*O*-methylglucuronic acid (4-*O*-MeGlcA) groups, on average, at every 10th xylose unit. The xylose residues are partly *O*-acetylated at the C2 and/or the C3 positions, with approximately seven acetyl groups per ten xylose residues. Softwood xylan represents only 5-10 % of the dry weight, and is present as arabinoglucuronoxylan. It has a larger number of 4-*O*-MeGlcA decorations than hardwood, regularly distributed (every 5-6 xylose residues), and additionally contains arabinosyl side groups (every 8-9 xylose residues). Unlike hardwood xylan, no acetyl groups are present (Figure 5A). Softwood xylan is slightly shorter than xylan from hardwood, with DP values typically reported to be between 70 and 130 (Fengel & Wegener, 1989; Sjöström, 1993; Telemann, 2009). Xylan is also a major cell wall polysaccharide in many grasses and annual plants, where it is found in different chemical and structural variants. The cell walls of corn cob, for example, are rich in heteroxylan, contributing approx. 30 % to the dry weight (Da Silva et al., 2015). Ball-milled corn cob was used in this work as a substrate to investigate the boosting effect of glucuronoyl esterases on xylanases and cellulases (**Paper IV**).

Mannans from wood consist of a heteropolymer backbone of β-1,4-linked mannose and glucose moieties, and are therefore correctly referred to as glucomannans (Fengel & Wegener, 1989). Glucomannans are the principal hemicelluloses in softwood, present as galactoglucomannans, containing β-1,6-linked D-galactosyl side groups (Figure 5B). Softwood galactoglucomannans are divided into a galactose-rich fraction and a fraction low in galactose, with Gal:Glc:Man molar ratios of 1:1:3.5 and 0.1:1:3.5, respectively. Hardwood glucomannan makes up only 2-5 % (w/w) of the cell wall. The Glc:Man ratio varies within 1:1-2, and no galactose units are present. The lack of galactosyl side groups makes hardwood glucomannan less water soluble than softwood galactoglucomannan. *O*-acetylation of the C2 and/or C3 positions of the mannose residues are seen in both wood types, with a degree of substitutions between 0.17 and 0.36, and DP values between 50 and 100 (Sjöström, 1993; Telemann, 2009). Spruce-derived galactoglucomannan was used in this work and subjected to enzymatic hydrolysis to study the substrate specificity of *endo*-mannanases (**Paper II**).

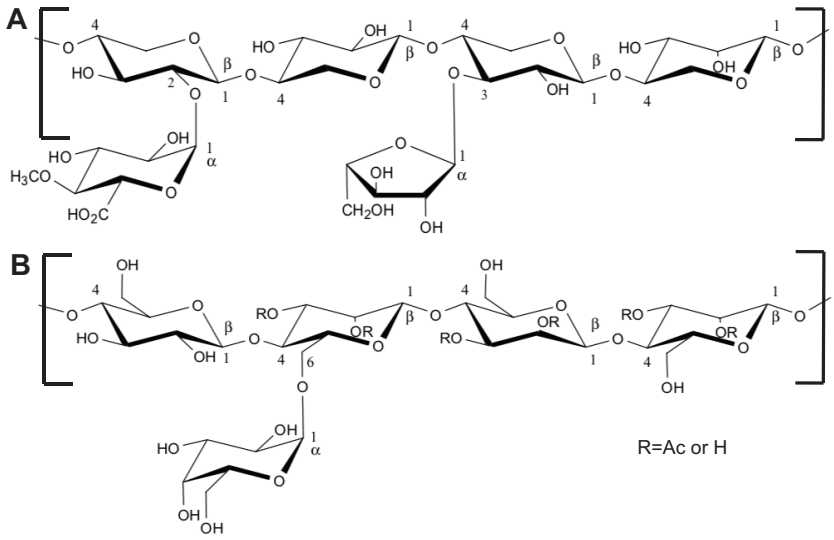


Figure 5. Chemical structures of the major hemicelluloses in softwood: arabinoglucuronoxylan (**A**) and galactoglucomannan (**B**). Arabinoglucuronoxylan is made up of β-1,4-D-xylan, decorated with arabinosyl and 4-*O*-MeGlcA moieties. Galactoglucomannan consists of a backbone of β-1,4-linked mannose and glucose units, decorated with galactosyl moieties. The mannose units can additionally be *O*-acetylated at the C2 and/or C3 positions.

Pectins

Pectins are a class of heterogeneous branched polysaccharides, not classified as hemicelluloses. They consist predominantly of a backbone of α-1,4-linked galacturonic acid (GalA) units, sometimes intermixed with rhamnose residues. The galacturonans are partly methylated at the C6 position and *O*-acetylated at the C2 and/or C3 positions. The high content of carboxylic groups contributes to their charged nature. The pectin structure can be more complex, being composed of several different types of monosaccharides and linkages, and their complete structure has still not been fully elucidated (Telemann, 2009). Pectins are common in many plant species, including fruits and vegetables, but constitute only a minor component (approx. 5 %) of wood tissue, found predominantly in the thin primary cell wall and the middle lamella (Voragen et al., 2009).

Lignin

Lignin is the most abundant source of aromatics in nature and makes up approximately one third of the plant cell wall. In contrast to the long polysaccharide chains, the structure of lignin can be described as a three-dimensional network, with randomly distributed monomers covalently connected by carbon-carbon bonds (Figure 6A). There are many types of bond, but the most important is the β-*O*-4 ether bond. The main monomers are three types of monolignols, methoxylated to various degrees: *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol (Figure 6B). The relative amounts of these monolignols depend on the source of the lignin. Softwood lignin consists almost exclusively of coniferyl alcohol (95 %), while hardwood lignin contains a mixture of coniferyl (25-50 %) and sinapyl alcohol (46-75 %). *p*-coumaryl alcohol is only present at small amounts in woody biomass, but is relatively abundant in grass lignin. It has been suggested that lignin contributes to cell wall stiffness, serving as a glue between cell walls, offering protection against microbial degradation and providing water resistance to the cell wall (Duval & Lawoko, 2014; Henriksson, 2009).

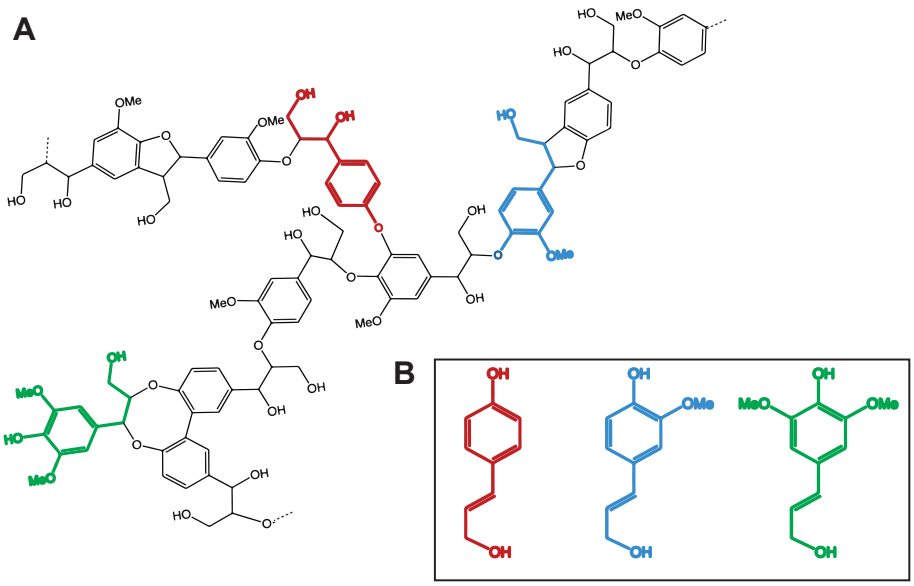


Figure 6. The aromatic, three-dimensional structure of lignin, with randomly distributed monolignols covalently connected by carbon-carbon bonds (**A**), and the major monolignols: *p*-coumaryl alcohol (red), coniferyl alcohol (blue) and sinapyl alcohol (green) (**B**). Coniferyl alcohol is the predominant building block in softwood, whereas hardwood lignin consists of both coniferyl and sinapyl alcohol. *p*-coumaryl alcohol is mainly found in grass lignin.

2.3 Linkages between lignin and carbohydrates

Lignin is connected to hemicelluloses via covalent bonds, forming so-called lignin-carbohydrate complexes (LCCs). The LC bonds are estimated to be few, but contribute to the recalcitrance of plant biomass. This might be crucial for the cell wall structure, but complicates delignification and depolymerisation in industrial wood processing (Henriksson, 2017). The presence of LC linkages was suggested already in the 19th century by Erdmann, due to his difficulties in separating lignin (which he called drupose) from carbohydrates (Erdmann, 1866). The term LCC was first introduced by Björkman, who developed a method for isolating milled wood lignin in the 1950s (Björkman, 1957).

Numerous investigations have been conducted to elucidate the structure and nature of LCCs in plant cell walls, and several indirect observations have been made of their presence, based on the solubility properties of wood polymers, isolation of LCCs, degradation methods, and analysis with size-exclusion chromatography and 1D nuclear magnetic resonance (NMR) spectroscopy (Brunow & Lundquist, 2010; Henriksson, 2017; Koshijima & Watanabe, 2003). However, the presence of LC bonds has been debated over the years due to the lack of direct evidence. The development of different 2D NMR techniques, able to detect LC linkages directly, and advances in LCC enrichment through improved extraction methods have, therefore, been key in proving their existence (Balakshin et al., 2011; Balakshin et al., 2007; Giummarella et al., 2016; Nishimura et al., 2018; Yuan et al., 2011).

Three types of LC bonds have been proposed: ester bonds to the glucuronic acid (GlcA) moieties of xylan (Figure 7A), ether bonds to arabinose and xylose residues in xylan (Figure 7B) and phenyl glycoside bonds to the reducing end of glucomannan (Figure 7C) (Brunow & Lundquist, 2010; Fengel & Wegener, 1989; Henriksson, 2017; Koshijima & Watanabe, 2003).

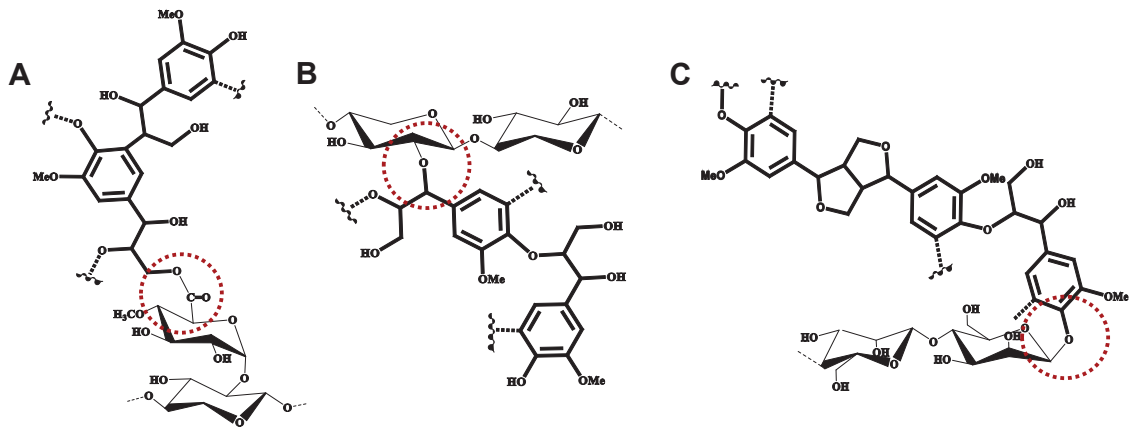


Figure 7. Fragments of LCCs with the carbohydrate fraction linked to lignin by the three suggested LC linkages (indicated in red): LC ester- (**A**), LC ether- (**B**) and LC phenyl glycoside (**C**) bonds. (Figure courtesy of Nicola Giummarella.)

LC ester bonds

In the current work, a major research effort was devoted to investigating the enzymatic hydrolysis of LC ester bonds (**Paper III**). Ester bonds are linked to 4-*O*-MeGlcA moieties on xylan (Figure 8), but might also be present on GalA residues in pectin (Wang et al., 2013). For a long time, it was believed that LC ester bonds were present as benzyl esters, linked to the α-carbon of the uronic acid. This suggestion was based on the results of indirect detection

methods, such as derivatisation and 1D NMR (Fengel & Wegener, 1989; Koshijima & Watanabe, 2003; Watanabe, 1995; Xie et al., 2000). In more recent studies, 2D NMR spectroscopy has been applied, allowing for direct detection of the chemical structures in lignocellulosic materials. In the majority of these studies, α-linked LC esters have not been identified, but instead γ-linked LC esters (Balakshin et al., 2011; Balakshin et al., 2007; Giummarella & Lawoko, 2016, 2017; Yuan et al., 2011). A large part of the research community is now convinced that the γ-form is the major one. However, there is still some scepticism as the γ-ester signals are heavily overlapped by signals from other types of γ-esters and γ-acetylated lignin structures in the NMR spectra (Nishimura et al., 2018).

Based on studies using synthetic model substrates, it has been proposed that LC esters are formed by nucleophilic attack of the α-carbon of the uronic acid, followed by acyl migration to the more stable γ-form (Evtuguin et al., 2005; Li & Helm, 1995). Direct detection and knowledge on LC ester formation indicate that the γ-ester is the biologically relevant type. However, both ester forms have been detected in a study on a synthetic LCC material (Giummarella, 2018). Furthermore, signals assigned to benzyl esters were clearly observed in 2D NMR spectra from LCCs extracted from Japanese beech in the present work (presented in Sections 5.7 and 7.8).

In my opinion, these observations, and the fact that enzymes are able to act on α-linked model substrates (discussed in Section 7.5) provide sufficient evidence for the existence of benzyl esters in lignocellulose. LC esters may exist in both the benzyl and γ-form, but the presence and ratio of the two types of ester may depend on the biomass source and processing steps/conditions (ball milling, solvent extraction, high temperature, etc.). More knowledge is required on LC ester formation, together with the development of non-destructive extraction procedures and analysis methods before we can fully understand the nature of the LC ester bond.

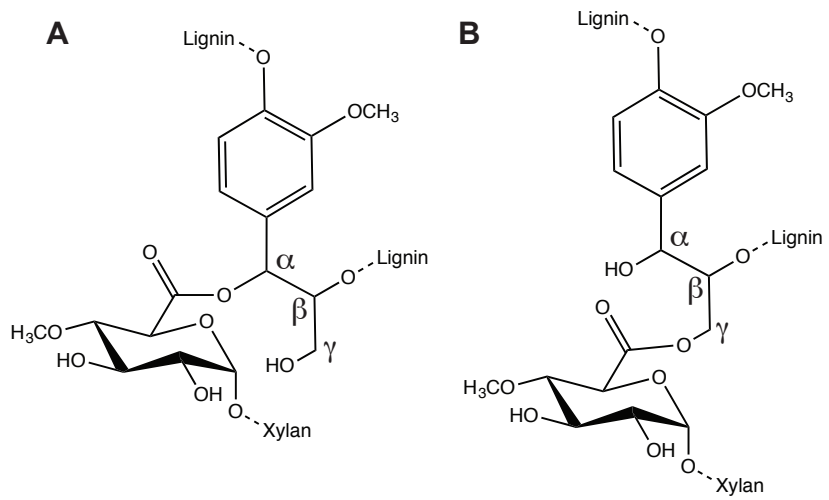


Figure 8. Structure of LC esters, either α-linked (benzyl ester) (**A**) or γ-linked (**B**) to the lignin alcohol. Direct detection suggests the γ-ester is more abundant, but the presence and the ratio of the two types of ester may depend on the biomass source and the kind of processing.

3. Carbohydrate-active enzymes

Enzymes that act on lignocellulose are classified as carbohydrate-active enzymes (CAZymes). CAZymes are grouped into enzyme families, based on similarities in amino acid sequence, protein fold and mode of action, in the carbohydrate-active enzymes database (CAZy) (www.cazy.org; (Lombard et al., 2014)). The families are grouped into five different enzyme classes (glycoside hydrolases (GHs), glycosyl transferases, polysaccharide lyases, carbohydrate esterases (CEs), auxiliary activities (AAs)) and a class of non-catalytic carbohydrate-binding modules (CBMs). GHs are additionally divided into clans, i.e., groups of families, that share a structural fold and catalytic machinery. Biochemically characterised enzymes are identified by an enzyme commission (EC) number, based on the chemical reaction they catalyse. Enzymes from different CAZy families that have the same activity will therefore share EC numbers, and enzymes with more than one activity will have multiple EC numbers (Lombard et al., 2014). In the present work, enzyme activities belonging to GHs and CEs were investigated, and the CAZy classification system is used when describing the enzymes.

3.1 Glycoside hydrolases

GHs are the most important carbohydrate-degrading enzymes in nature, and represent the majority of all entries in CAZy. To date, over 150 GH families have been identified (www.cazy.org; January 2019), some of which only containing a single activity (e.g. GH11, containing xylanases only), whereas other families are very broad with several different specificities (e.g. GH5, comprising over 10 activities) (Lombard et al., 2014). GHs catalyse the hydrolysis of glycosidic bonds via acid catalysis, normally between two sugar molecules, which involves a proton donor and a nucleophile (Davies & Henrissat, 1995).

GHs that act on oligosaccharides and polysaccharides can be grouped into *exo*- and *endo*-types, based on how they cleave their substrate (Naumoff, 2011). *Endo*-acting GHs cleave internal linkages and produce oligosaccharides, whereas *exo*-acting GHs target the reducing or the non-reducing end of the substrate backbone and side chains, which results in monosaccharide or shorter oligosaccharide products (Davies & Henrissat, 1995). *Endo*-acting GHs recognise and accommodate a portion of the polymeric substrate in their active site in order to facilitate hydrolysis. Binding of the substrate is assisted by a subsite system in the cleft-like active site, where each subsite binds to specific sugar units in the polysaccharide chain. The subsites are comprised of one or several amino acid residues and the interactions with the polysaccharide are important for substrate recognition, catalytic action, product release and processivity (Zhang et al., 2015). They are labelled from the non-

reducing end starting with -*n*, going to +*n* closest to the reducing end. Hydrolysis occurs between -1 and +1 (Figure 9) (Davies et al., 1997). The subsites of *endo*-acting mannanases are further described in Chapter 6 and **Paper II** in relation to the hydrolysis of mannan polymers.

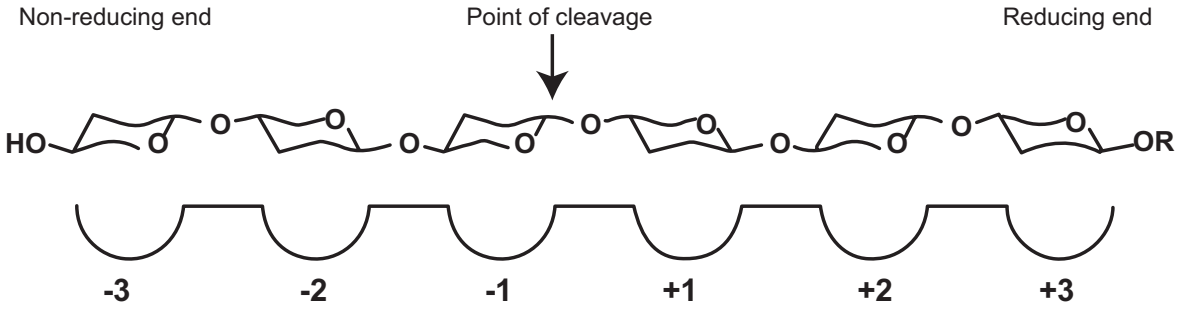


Figure 9. Schematic of the nomenclature used for the sugar-binding subsites in *endo*-acting GHs. The arrow indicates the glycosidic bond that is susceptible to enzymatic hydrolysis.

3.2 Carbohydrate esterases

CEs catalyse the hydrolysis of ester bonds of *N*- or *O*-acylations linked to carbohydrates, where the sugar moiety plays the role of either the alcohol (e.g. acetylated mannan) or of the acid in the form of uronic acids. CEs are currently classified into 15 families (www.cazy.org; January 2019), including activities such as acetyl esterases, feruloyl esterases, pectin acetyl esterases, pectin methyl esterases and GEs (Biely, 2012; Nakamura et al., 2017). The most common catalytic mechanism involves a serine-histidine-aspartate/glutamate catalytic triad, analogous to classical serine hydrolases (further described in Section 7.6). Enzymes with catalytic dyads, where catalysis is mediated by serine and histidine residues only, have also been reported (Biely, 2012). CEs are important for lignocellulose degradation through the hydrolysis and removal of ester side groups from the polysaccharides that hamper the action of GHs. In this way, CEs can boost and facilitate the action of GHs (discussed in **Paper II**). CEs are also useful for decoupling lignin from the carbohydrate fractions, through GE-mediated hydrolysis of covalent ester linkages (**Papers III, IV**, **V** and Chapter 7).

3.3 Lignocellulolytic enzymes

Cellulolytic enzymes

It is generally accepted that three types of GH activity are involved in the hydrolysis of cellulose; *exo*-1,4-β-glucanases (cellobiohydrolases), *endo*-1,4-β-glucanases and β-glucosidases. Cellobiohydrolases attack the reducing or the non-reducing ends, producing cellobiose units, while 1,4-β-glucanases primarily reduce the DP of the polymer by randomly cleaving internal bonds in the amorphous and less ordered parts of the cellulose chain. β-glucosidases convert the cellobiose and cello-oligosaccharides produced into glucose. In addition to these activities, enzymes that break glycosidic bonds in the crystalline and highly ordered cellulose by oxidation (Lytic polysaccharide monooxygenases, classified as AAs), and aiding proteins such as swollenins and expansins have been reported to be important for efficient cellulose degradation (Horn et al., 2012; Van Dyk & Pletschke, 2012).

Hemicellulolytic enzymes
A large number of enzyme activities are required to efficiently hydrolyse the heterogeneous hemicelluloses into monosaccharides, including a range of GHs and CEs. For mannan polymers, the core enzymes involved in hydrolysis are *endo*-1,4-β-mannanases (**Paper II**, further described in Section 6.3), producing shorter oligosaccharides, and *exo*-acting β-mannosidases that convert shorter oligosaccharides into mannose. The corresponding enzyme activity for xylan degradation are *endo*-1,4-β-xylanases and *exo*-1,4-β-xylosidases, respectively. Not only depolymerising enzymes that act on the backbone are involved, but also enzymes that remove substituents. Depending on the type of substituents and decoration linked to the backbone, various additional enzymes are required, such as α-L-arabinofuranosidases, α-glucuronidases, α-galactosidases and acetyl esterases (Van Dyk & Pletschke, 2012). Linkages connecting hemicelluloses to the lignin polymer must also be considered in association with hemicellulose degradation (or decoupling), where GEs play a key role (**Papers III, IV, V** and Chapter 7).

Ligninolytic enzymes
Ligninolytic enzymes are included in the CAZy database despite the fact that they do not act directly on carbohydrate structures. The close association between lignin and polysaccharides in the plant cell wall makes the cooperation of these enzymes with classical polysaccharide depolymerases important, and they are therefore classified as AAs. Examples of lignin-depolymerising enzymes are laccases, manganese peroxidases and lignin peroxidases. They all catalyse oxidation or reduction reactions and, therefore, normally carry metal ions necessary for their function (Teeri & Henriksson, 2009).

Carbohydrate-binding modules
CBMs are non-catalytic proteins that bind to different types of carbohydrate structures. They are normally found within a longer protein sequence, together with catalytic GH modules, but in rare cases can be found on their own. Their primary functions are to position the catalytic modules close to the substrate and prolong the enzyme-substrate interaction, thereby enhancing the catalytic efficiency. CBMs are grouped into several families in CAZy based on amino acid similarity, and a large variety of folds and substrate specificity have been identified, e.g. affinity for cellulose, xylan, mannan and β-glucan (Guillen et al., 2010). Some CBMs are very specific, whereas others are more promiscuous and have evolved to recognise more than one type of monosaccharide or glycosidic bond (Armenta et al., 2017). CBMs that target insoluble polysaccharides can also play a significant role in non-specific adsorption to lignin, leading to an undesired reduction in enzyme activity during hydrolysis (Varnai et al., 2013).

3.4 Discovery and characterisation of lignocellulolytic enzymes

Fungi and bacteria that degrade plant biomass efficiently are excellent sources of CAZymes. Proteomic, transcriptomic and biochemical studies of these microorganisms are therefore common in the search for novel lignocellulolytic enzymes. Bioinformatic analyses are also important, as genomic or metagenomic information is useful in identifying sequence similarities to other CAZymes. Enzyme sequences are listed and classified in protein databases (e.g. CAZy and the Protein Data Bank (PDB)), however, the majority of these entries have not been characterised. These databases are therefore good starting points in the search for appropriate enzymes to express and investigate (Lombard et al., 2014). The GEs described in this thesis (**Papers IV**, **V)** were selected from uncharacterised sequences in the CE family 15 in CAZy.

Biochemical assays are performed on relevant substrates to investigate the actual function of an expressed enzyme (as further described in Chapter 5). Assays can be run for quick screening purposes to identify an enzyme, or for detailed characterisation to quantify enzyme activity and determine enzyme-substrate interactions. The catalytic potential is the essential feature of enzymes and, therefore, the quantification of an enzyme is expressed in terms of its activity rather than the mass of protein. Detection of substrate consumption or product formation is used to obtain a quantitative estimate of enzyme performance and substrate preferences, and kinetic constants such as k_{cat} (substrate turnover number) and K_M (substrate concentration required for half-maximal velocity) can be calculated (Bisswanger, 2014). Functional studies are often combined with structural determinations. Elucidating the three-dimensional structure of an enzyme in combination with knowledge on its activity gives more detailed information on its ability to accommodate substrates and confirms its catalytic mechanism. Structure-function studies of GEs constituted a major part of this work (**Papers IV**, **V)**.

4. Enzyme accessibility in wood structures

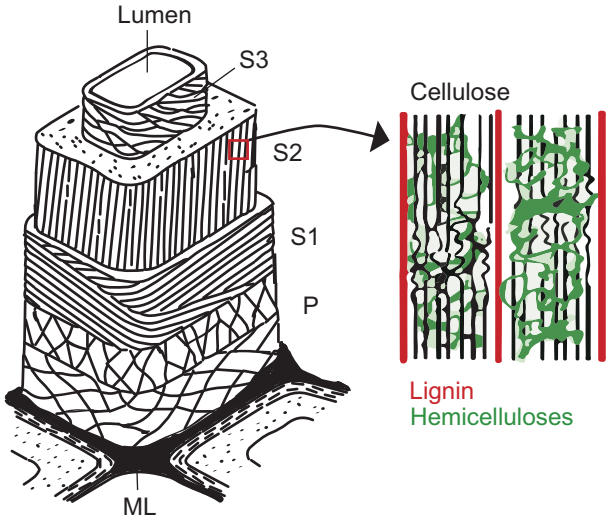
CAZymes are able to depolymerise complex carbohydrate structures, but in order to function they must be able to reach their target substrate. One challenge associated with woody biomass is the dense and chemically complex ultra-structure, which hampers accessibility. Different forms of biomass pre-treatment prior to enzymatic hydrolysis are therefore crucial to open up the wood structure and increase the surface area, thereby facilitating contact between the enzymes and the chemical linkages on which they act. There is thus a need to develop analytical methods that can detect enzyme mobility and attachment to different parts of the cell walls in order to evaluate how the physico-chemical effects of pre-treatment contribute to enzyme performance and substrate accessibility. In this chapter, the results obtained concerning the effect of steam explosion (STEX) pre-treatment in enhancing the enzymatic hydrolysis of wood substrates (**Paper I**) are presented, and the potential of using FRAP as a direct method of evaluating enzyme accessibility in the same substrates is discussed.

4.1 The complex ultra-structure of wood

The ultra-structure of wood refers to the organisation of polymers in the different cell wall layers, and is the hierarchic level between the molecular (chemical composition) and the cell wall level (Figure 2). The cellulose microfibrils have different arrangements depending on the layer in the plant cell wall (Figure 10). In the thin primary layer (P), they are randomly aligned in a matrix of hemicelluloses, pectins and glycoproteins. In the three secondary layers (S1, S2 and S3) the microfibrils are ordered in a highly aligned manner, at different angles depending on the layer. S2 is structurally the most important layer due to its thickness (1-5 µm), and the almost vertical orientation of its microfibrils contributes to its strength. The microfibrils can be further aggregated into lamellar macro-fibrils that afford support and rigidity to the wood cell. The hemicelluloses are amorphous and oriented along the cellulose, while lignin surrounds the cellulose and the hemicelluloses (Figure 10).

The entanglement and tight connection between the wood polymers makes processing of woody biomass demanding (Brändström, 2001). The ultra-structure also reduces the accessibility of the substrate, and thus the performance of specific enzymes. Furthermore, structural factors such as cellulose crystallinity, DP and lignin distribution/chemistry (contributing to unproductive enzyme binding) are frequently associated with limited enzyme efficiency (see Box 2, Figure 10) (Chandra et al., 2009; Nakagame et al., 2010). Some of these factors have been considered more influential for biomass recalcitrance than others, but contradictory trends have been reported. The relationship between various substrate characteristics and recalcitrance is clearly non-linear and since the alteration of one

substrate characteristics most likely influences other substrate properties, it can be complicated to separate the individual effects (Foston & Ragauskas, 2012).



Box 2. Factors described to influence the accessibility and efficiency of enzymes in woody biomass.

-The dense wood structure
-Closely connected and entangled polymers
-Cellulose crystallinity
-DP
-Lignin distribution
-Lignin chemical composition

Figure 10. The wood cell wall layers and the middle lamella (ML) showing the cellulose microfibrils in different orientations depending on the layer. In the thick S2 layer, cellulose is vertically aligned, and embedded in a matrix of lignin (red) and hemicelluloses (green). Factors described to influence the enzymatic hydrolysability of woody biomass are listed in Box 2.

4.2 Pre-treatment to open up the wood structure

Native wood is recalcitrant to enzymatic degradation, and various pre-treatment strategies can be applied to open up the wood structure in order to increase the surface area of the polymers accessible to the enzymes. Pre-treatment is often classified into chemical and physical methods, where physical disruption of the biomass is normally achieved by grinding or milling processes. The breakdown into smaller particles increases the overall surface area, but does not result in the removal of lignin or hemicelluloses (Chandra et al., 2007).

Chemical pre-treatment often involves the hydrolysis of hemicelluloses and/or the redistribution of lignin, which exposes the polysaccharides. The uncovered cellulose fibrils and the partial solubilisation of hemicelluloses allow for the cellulolytic and hemicellulolytic enzymes to reach their respective targets more easily. The removal of lignin leads to a reduction in unproductive binding of enzymes, however, chemical pre-treatment can also lead to depolymerisation and repolymerisation (condensation) of lignin. Condensed lignin tends to be more hydrophobic, and has been observed to increase the level of unproductively bound enzyme (Nakagame et al., 2010).

Chemical pre-treatment is commonly performed by applying acid or alkaline reagents to the biomass. One problem associated with this approach is the formation of lignocellulose-derived by-products such as weak acids, phenolic compounds and furan derivatives, which can inhibit the enzymes (Jönsson & Martin, 2016). Liquid hot water pre-treatment is, in this sense, a milder method that contributes to chemical modification without the need of chemicals and with reduced formation of inhibitors. Drawbacks of hot water pre-treatment are the high demands for water and energy, which make the process costly (Zhuang et al., 2016). STEX is another alternative that combines physical and chemical disruption of the biomass. The formation of inhibitors is low, and the energy demand is significantly lower

than in liquid hot water pre-treatment (Alvira et al., 2010). STEX was used as pre-treatment for spruce wood chips in the study presented in **Paper I**.

4.3 STEX pre-treatment

STEX is the most widely used physico-chemical pre-treatment for woody biomass, resulting in ultra-structural, polymeric and chemical modifications. The biomass, in the form of wood chips, is initially fed into a vessel and treated with steam under high pressure (steam treatment). This is followed by a sudden release of pressure (explosion) (Figure 11A). The high temperature and the presence of water promote the autohydrolysis of acetyl groups present in the hemicelluloses, resulting in partial solubilisation of the hemicelluloses, while the pressure drop leads to explosive decompression and mechanical separation of the wood fibres. Other chemical modifications that have been reported are the condensation of lignin and an increase in the crystallinity of the cellulose in the residual material (Nakagame et al., 2010). The effect of the steam treatment is dependent on temperature (T [°C]) and residence time (t [min]). Changing these process parameters allows the level of pre-treatment severity, R, to be changed (Eq. 1) so as to avoid undesirable chemical reactions (Alvira et al., 2010; Jacquet et al., 2015).

$$Log\ (R) = Log\ t^{*}\exp\left((T - 100)/14.75\right) \qquad \text{(Eq. 1)}$$

A high value of R indicates severe process conditions, i.e. high temperatures and/or long residence time. A low value of severity has been defined as Log (R) < 3, whereas Log (R) > 4 is considered as high severity (Wu et al., 1999). High severity conditions result in a material with the highest cellulose hydrolysability, but low severity is to preferable for the recovery of the hemicellulosic and cellulosic components.

STEX pre-treatment in a materials biorefinery

Hemicelluloses, especially galactoglucomannan, are easily degraded and solubilised during acidic hydrolysis reactions during steam treatment. This has previously been considered problematic in terms of inhibitory degradation products (such as furfural and hydroxymethyl furfural). In a materials biorefinery, the destruction of intramolecular structures is an additional problem, since the aim is to keep the individual wood polymers as intact as possible. It is therefore necessary to develop and adapt STEX pre-treatment to avoid complete degradation of the hemicelluloses, but still improve the accessibility of the substrate to specific enzymes.

STEX at lower temperatures and shorter residence times has been reported to be efficient in terms of enhanced enzyme performance, albeit with reduced hydrolysis of the hemicelluloses (Jedvert, Wang, et al., 2012). Combining STEX with the reducing agent $NaBH_4$ has been described as an approach to further limit the degradation of galactoglucomannan (Jedvert, Saltberg, et al., 2012). Reduced breakdown of oligosaccharides through limited acid hydrolysis of hemicelluloses has also been observed, when buffers where introduced into STEX reactions (Michalak et al., 2018). Moreover, milder pre-treatment strategies should be combined with efficient recovery of partially hydrolysed hemicelluloses, in order to minimise the loss of raw material.

4.4 Contribution of STEX to enhanced hydrolysability

STEX has been used as pre-treatment for a variety of lignocellulosic materials, and the effects of varying the factors contributing to the severity of the process (e.g. residence time, pressure, temperature, moisture content, sample size) have been extensively investigated (Jacquet et al., 2015). Several studies have shown how the chemical effect of steam treatment alone leads to changes in the levels of enzymatic hydrolysis and unproductive binding (Pielhop et al., 2016).

Auxenfans and co-workers investigated both structural and chemical modifications of diverse biomass sources resulting from steam treatment. They suggested that the redistribution of lignin and the removal of hemicelluloses was important for improved cellulolytic enzyme performance, while changes in cellulose crystallinity and lignin composition were considered to have less impact on the enzyme-catalysed hydrolysis (Auxenfans et al., 2017). Other studies have focused on how changes in the lignin chemistry contribute to changes in unproductive enzyme binding (Nakagame et al., 2011; Nakagame et al., 2010).

Despite many studies on the effect of steam treatment, the importance of the explosion step for the biochemical conversion has been less well explored. It has even been reported that enhanced substrate hydrolysability is predominantly the result of the steam treatment step, and not due to the physical disruption resulting from explosive decompression (Duff & Murray, 1996).

The importance of physical disruption for enzyme performance

The aim of the present study (**Paper I**) was to identify the contributions from the chemical and physical components of STEX on enzyme performance, by differentiating between the steam treatment and the explosion step. The individual contributions of the STEX process were studied by applying a cellulolytic enzyme cocktail to differently treated wood chips of Norway spruce (*Picea abies*).

Three cases were compared: (1) untreated wood chips, (2) steam-treated wood chips and (3) STEX-treated wood chips (Figure 11B-D). The STEX-treated wood was saturated with steam in the reaction vessel, followed by a rapid pressure drop to atmospheric pressure. Steam treatment, on the other hand, was performed with a slow pressure release, to avoid the expansion of vapour in the wood. The steam-treated wood chips were additionally held in a wire cage, to prevent them from colliding against each other, and the wall of the reaction vessel (Figure 11A). The STEX-treated wood chips were both chemically and physically modified and this material provided information on the contribution of the physical disruption.

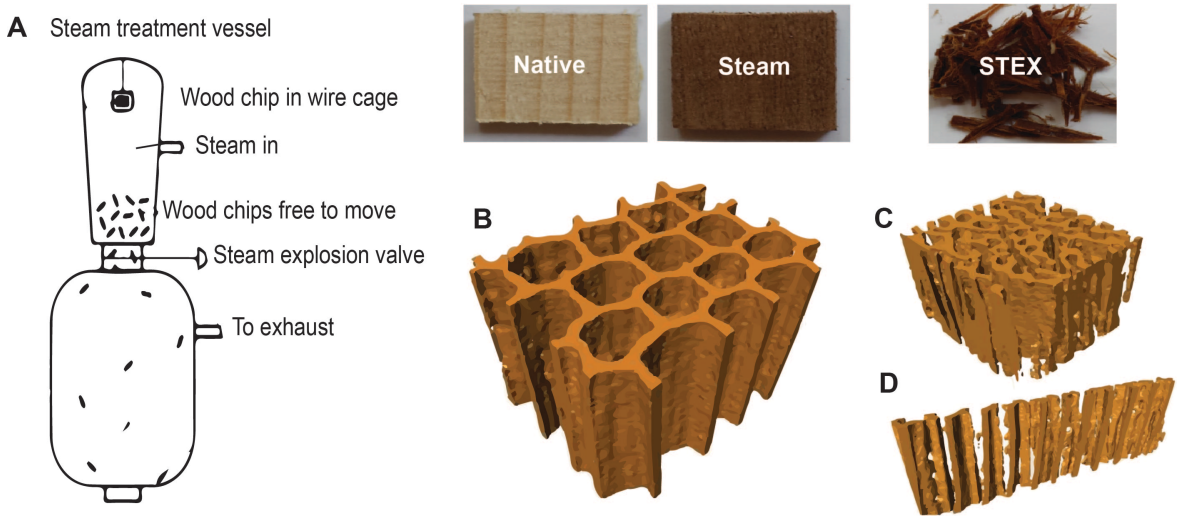


Figure 11. Simplified illustration of the STEX equipment used for steam treatment or STEX-treatment of wood chips (**Paper I**) (**A**). Wood samples and X-ray tomography images of the internal structures of intact wood (native or steam-treated) (**B**) or STEX-treated and structurally deformed wood with the formation of microcracks in the cell walls (**C-D**). (Figures courtesy of Muhammad Muzamal.)

The glucose released by enzymatic hydrolysis was analysed and compared with the initial amount of glucose present in the woody biomass. The results showed a very low glucose yield from the hydrolysis of native wood. Steam treatment led to a twofold increase in the glucose released and STEX led to an additional twofold increase (Figure 12). The enhanced glucose yield from STEX-treated wood illustrates the importance of explosive decompression for enzymatic cellulose digestibility. The effect of explosion has been reported to be the result of the reduced particle size of the biomass. However, explosion also contributed to the formation of microcracks in the cell walls and the structural deformation of cells (identified by X-ray tomography, **Paper I**) (Figure 11C-D).

Shortly after **Paper I** was published, additional studies on the contribution of explosive decompression became available (Pielhop et al., 2016; Seidel et al., 2017). They also confirmed the importance of the explosion step for enhanced hydrolysis of spruce, and reported similar results for hardwood and other plant species. It is thus clear that STEX contributes to enhanced enzyme performance, probably through improved enzyme accessibility, but the effects of structural and chemical changes in enzyme accessibility have not yet been directly assessed.

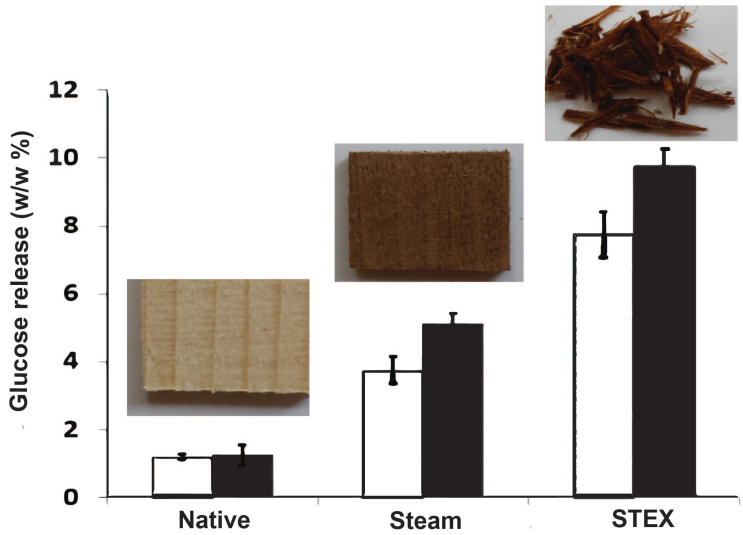


Figure 12. Enzymatically catalysed glucose release after 30-h (white) and 72-h (black) incubation with native, steam-treated or STEX-treated wood chips. The twofold increase in glucose release from STEX-treated wood compared to steam-treated wood demonstrates the importance of the explosive decompression for enhanced enzyme performance.

4.5 Evaluating enzyme accessibility

It is essential to understand the mechanisms that control the accessibility of different substrates to enzymes in order to evaluate the activity of the enzymes and to facilitate the design of biomass pre-treatment. Detection methods that are capable of measuring and quantifying the diffusion and binding of enzymes in lignocellulose directly are therefore highly desirable. Imaging and spectral fluorescence analysis are highly sensitive, non-invasive methods that have the potential to provide such information.

Fluorescent recovery after photobleaching

FRAP is a technique used for studying dynamic processes in biological cells and materials. A confocal microscope is used to direct a high-power laser beam into a region of interest (ROI), resulting in the irreversible bleaching of fluorescent molecules (probes). The surrounding fluorescent molecules can thereafter freely diffuse into the ROI and the fluorescence recovery is recorded until it reaches a plateau (Figure 13). The fraction of probes that can move between the non-bleached area and the bleached ROI is called the mobile fraction, while the fraction that cannot is called the immobile fraction. The diffusion constant (D) of the mobile fraction can be calculated from the half-life and the radius of the bleached ROI. The value of D provides information on the mobility of the probe molecules (Paës, 2014). The probes may be either fluorophores themselves, dyed molecules or molecules coupled to a fluorophore. One commonly used probe for FRAP is green fluorescent protein (GFP), a non-catalytic protein that exhibits bright green fluorescence when exposed to light in the blue to ultraviolet range (Lorén et al., 2015). FRAP has been used successfully as a tool for determining the diffusion and interaction/binding of proteins, lipids and other molecules in living cells (Lorén et al., 2015). Based on earlier applications, FRAP was considered promising for studying enzymes in complex lignocellulosic materials.

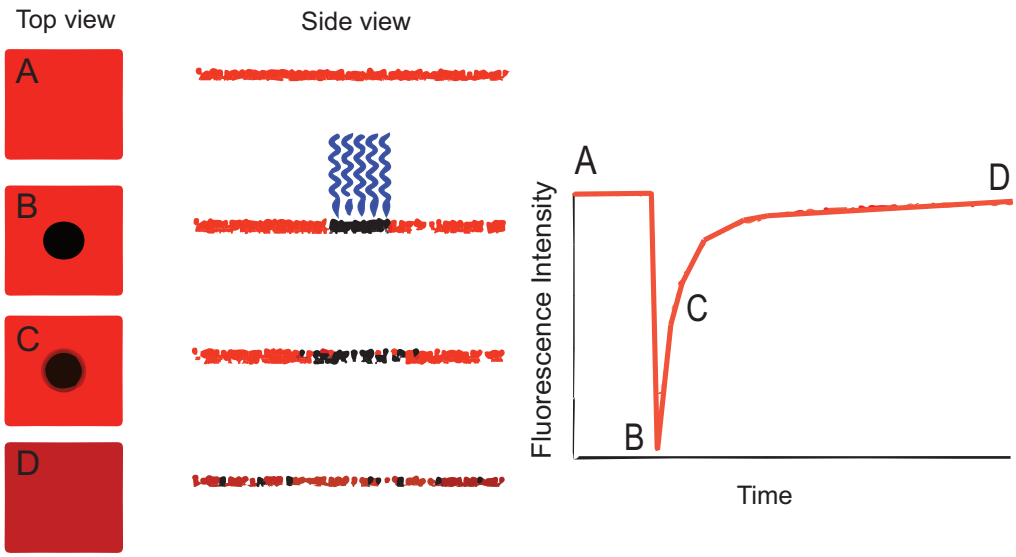


Figure 13. The principle of FRAP showing the top view and the side view of the bleached material (in this work, slices of wood) with the ROI/bleached molecules coloured in black, as well as the FRAP recovery curve of the mean intensity over time: initial fluorescence intensity (**A**), photobleaching by a high-intensity laser pulse (**B**), the intensity within the ROI is monitored as the bleached molecules diffuse out and others diffuse in (**C**), and the intensity recovery is reaching a plateau (**D**).

FRAP for assessing enzyme mobility

FRAP has in recent years become established as a technique for studying enzyme mobility. One group has evaluated the diffusion of xylanases in different xylan substrates. By comparing a wild-type xylanase, inactive xylanase mutants and xylanases with modified CBMs, the authors reported on the importance of both catalytically active enzymes (substrate hydrolysis) and binding affinity (substrate targeting) for enzyme mobility (Cuyvers et al., 2011). The reversibility of cellulase binding in microcrystalline cellulose has been investigated in another study (Moran-Mirabal et al., 2011). Lignocellulosic model assemblies (containing feruloylated arabinoxylans and cellulose nanocrystals) have also been designed and used in FRAP studies to measure the binding properties of CBMs (Paës, 2017; Paës et al., 2015). The first studies measuring the mobility of molecular probes in native lignocellulose have recently been published (Herbaut et al., 2018; Paës et al., 2017). However, no FRAP studies in which enzymes and plant biomass are combined have yet been reported.

Accessibility of wood structures with a fluorescent CBM

In the present work, FRAP was evaluated as a potential method to detect changes in enzyme accessibility in spruce wood resulting from STEX. A GFP-linked CBM with an affinity for mannan (CBM27) was selected as the probe to investigate diffusion and binding in wood structures. GFP is in the same size range (27 kDa) as many CAZymes, and therefore appropriate as an enzyme model. GFP-CBM27 and GFP (used as the control) were applied to thin slices (200-300 µm) of native and STEX-treated spruce wood. It was hypothesised that the binding of CBM27 would be reduced in STEX-treated wood as the mannan polymers are partly degraded. The aim was to answer following questions:

1. Is the diffusion of GFP-CBM27 affected by STEX?
2. Is the distribution (i.e., in the lumen, on cell walls) of GFP-CBM27 changed by STEX?
3. Is the diffusion of GFP-CBM27 affected by changes in binding to the cell wall?

In preliminary tests (unpublished data), it was confirmed that GFP-CBM27 was detected (in the µM range) by confocal laser scanning microscopy and FRAP. As expected, the diffusion of GFP-CBM27 in native wood was significantly reduced compared to that in free solution (water). Also, binding of GFP-CBM27 to the surface of the inner cell walls of native spruce was clearly seen (Figure 14A), whereas GFP without the binding module was mostly seen in the lumen (Figure 14B). However, applying both probes to STEX-treated spruce samples gave rise to puzzling results. The distribution of GFP-CBM27 (Figure 14C) appeared to be identical to that of untreated wood, with most of the probes bound to the cell walls. Surprisingly, GFP was observed on the cell walls of the STEX-treated spruce, similar to the case with GFP-CBM27 (Figure 14D), but was not seen in the lumen.

The value of D was slightly lower for GFP-CBM27, and almost twofold lower for GFP, in STEX-treated wood than in native wood, probably as a result of increased irreversible and unproductive binding to the cell walls. No clear differences were observed concerning the diffusion and distribution of CBM27 before and after STEX treatment, but the mobility of both probes was indeed affected by the unproductive binding to the cell walls of the STEX-treated samples. Steam treatment has been reported to increase the hydrophobicity of lignin, which could be one explanation of the increased unproductive binding of both probes (Nakagame et al., 2011).

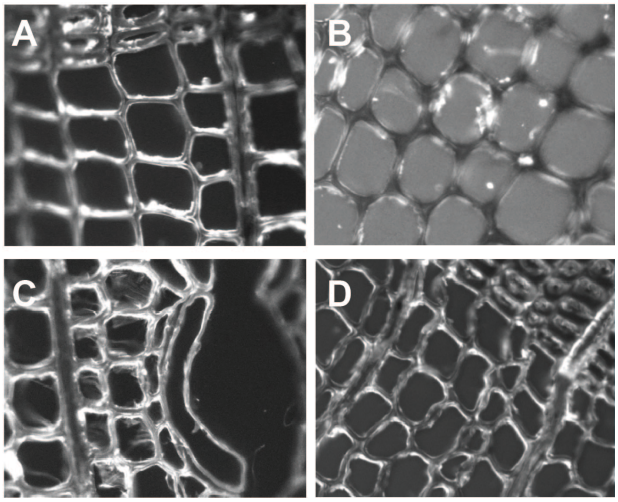


Figure 14. Distribution of GFP-CBM27 and GFP in wood samples detected by confocal laser scanning microscopy. GFP-CBM27 in native spruce (present on cell walls) (**A**), GFP in native spruce (mainly present in the lumen) (**B**), GFP-CBM27 in STEX-treated spruce (present on cell walls) (**C**), and GFP in STEX-treated spruce (present on cell walls) (**D**).

Despite the ambiguity of the results obtained from these preliminary tests with FRAP, this technique should be further investigated as a tool for analysing enzyme accessibility in wood structures. However, it might be crucial to reconsider the choice of probe and/or material. Both binding and hydrolysis have been reported to be important mechanisms for the mobility of CAZymes in lignocellulose (Cuyvers et al., 2011), and it may therefore be necessary to use catalytically active enzymes as probes in future FRAP studies. Enzymatic hydrolysis, which actively removes the hydrolysed substrate from the active site of the enzyme, could possibly result in a reduction in unproductive binding. It is also important to bear in mind that lignin autofluorescence and the heterogeneity of the wood material may influence the results provided by FRAP. Investigating enzymes with simpler and more well-defined substrates could therefore be useful as control experiments.

5. Detecting lignocellulosic enzyme activity

Detailed characterisation of enzyme activity is key to understanding the action of enzymes in different contexts. Depending on the substrate (e.g. simple synthetic substrates or complex biomass) and the aim of the study (quantitative or qualitative detection of enzyme activity), various detection methods are suitable. It is also important to use complementary analytical tools to confirm enzyme activity, especially if precise methods are lacking. This chapter briefly presents the analytical methods that were used to assay enzyme activity in this work, with emphasis on their applications, advantages and limitations.

5.1 Substrates

The enzymes used in this work were assayed on a range of substrates, including simple monomeric molecules, well-defined polysaccharides, extracted LCCs and complex plant biomass. The substrates were of different sizes, origin and chemical complexity, and gave information on various levels of detail regarding substrate specificity and enzyme activity. In the context of lignocellulosic enzymes, the chemical linkages present in the cell wall of plants are referred to as native substrates. However, many of these linkages cannot be quantified in the intact biomass, and thus various methods of isolation and fractionation will have to be applied. These processes might lead to modification of the chemical composition and the physical, supramolecular structure, and the resulting substrates will therefore not be completely "native". Bearing this in mind, all wood-derived materials are defined here as native substrates (e.g. isolated LCCs (**Paper III**) and galactoglucomannan extracted from spruce (**Paper II**)).

A model is a simple representation of a more complex system, and the same applies to model substrates. The model substrates used in these studies represent simpler or more accessible versions of wood-derived materials. They are small, well-defined chemical compounds that share key structures with the native substrate, e.g. GE model substrates and 4-nitrophenyl acetate (*p*NP-Ac) (**Papers IV, V**). However, in the work on mannanases (**Paper II**), polymeric mannans from konjac and locust bean gum served as model substrates. These polymers are extracted from plants, but have similar structural features to the target substrate spruce galactoglucomannan, and are commercially available and well-characterised. Ball-milled corn cob biomass was used as a model to woody biomass to investigate the auxiliary action of GEs on hemicellulase hydrolysis, due to its high content of xylan (**Paper IV**). Detailed information on all the substrates used in the studies reported in **Papers II-V** is given in Chapters 6 and 7.

5.2 Spectrophotometric assays

Spectrophotometric assays are commonly used to study enzyme kinetics by measuring the amount of light (in the UV/visible region) absorbed or scattered during a reaction. These assays are suitable when studying small, simple substrates that do not generate high levels of background absorbance, and can often be run continuously. The time dependency allows the reaction velocity to be determined, which is essential when evaluating the enzyme activity (Bisswanger, 2014). The restriction to simpler (often synthetic) substrates is a disadvantage, especially when the biologically relevant and native substrate is far more complex, as is the case with lignocellulose. The measurement of enzymatic parameters using model substrates therefore only provides an indication of the activity on the native substrate.

Reducing-sugar assay

Most of the spectrophotometric assays used to determine GH activities are based on the measurement of the amount of reducing sugars (free carbonyl groups (C=O)), upon oxidation of the aldehyde functional groups present in monosaccharides and oligosaccharides (new aldehyde groups are formed when glycosidic bonds in polysaccharides and oligosaccharides are enzymatically hydrolysed). There are several different methods based on similar principles, but the 3, 5-dinitrosalicylic acid (DNSa) assay is one of the most commonly used. The alkaline DNSa reagent involves the oxidation of free aldehyde groups, and DNSa is simultaneously reduced to 3-amino-5-nitrosalicylic acid, which is visualised by the colour transition from yellow to brown upon heating (absorption in the visible region) (Figure 15) (Miller, 1959). The darker the colour, the higher the concentration of reducing sugars, which indicates a high level of GH activity.

The DNSa method is quick and simple and allows for reactions with wood-derived substrates and native lignocellulose, but its relatively low sensitivity makes it necessary to regularly include blanks and standards. Overestimation of the activity due to polysaccharide degradation when heating the sample together with the alkaline DNSa reagent is also a problem. Furthermore, the assay cross-reacts with proteins (Gusakov et al., 2011). Detection of reducing sugars is therefore appropriate for qualitative measurements, but less suitable for quantitative purposes. The DNSa method was used to compare mannanase hydrolysis of mannan polymers (**Paper II**), to estimate the rate of enzymatic hydrolysis and the extent of saccharification (Table 2).

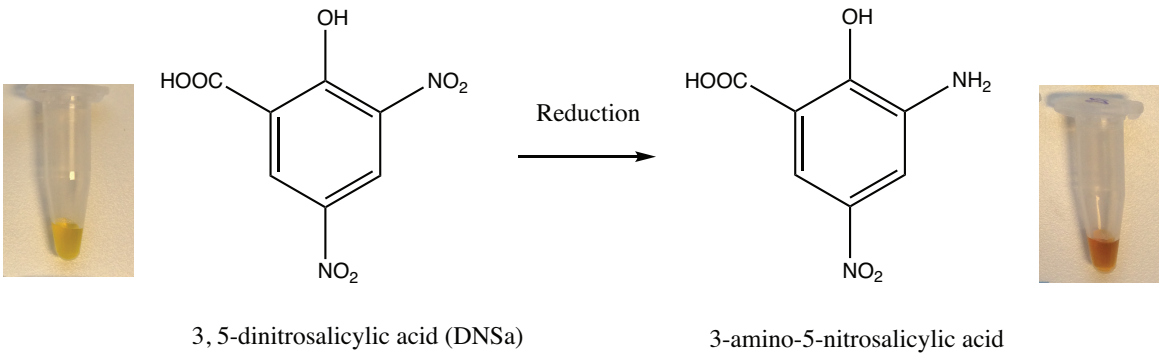


Figure 15. When the alkaline DNSa solution reacts with reducing sugars it is converted into 3-amino-5-nitrosalicylic acid. This is visualised by the change in colour from yellow to orange/brown upon heating, and the absorbance is detected at 595 nm.

Chromogenic substrate-based assays

It is necessary to use well-defined substrates for the screening of specific enzyme activities and to investigate substrate specificities. Chromogenic, 4-nitrophenol (*p*NP)-linked substrates are convenient models for many CAZymes, and are commercially available. The enzymatically catalysed *p*NP product, which is yellow in colour, absorbs at around 405-410 nm, and the detected intensity is proportional to the enzyme activity. Chromogenic substrates are suitable for rapid assays, but are small artificial compounds. Other drawbacks are the pH-dependence of the chromogenic properties of *p*NP (colour formation is only possible under alkaline conditions), and the fact that *p*NP-linked substrates might not be available for screening the desired activity. *p*NP-Ac (Figure 16) was used to screen for acetyl esterase activity among bacterial GEs to investigate enzyme promiscuity (**Papers IV, V**) (Table 2).

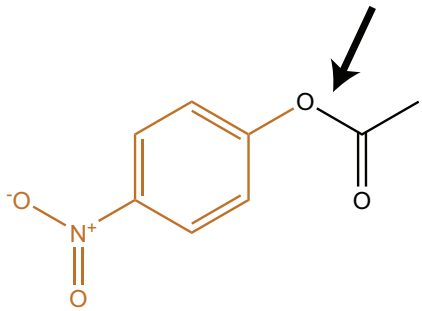


Figure 16. The structure of *p*NP-Ac, the substrate used for screening of acetyl esterase activity (**Papers IV, V**). The arrow indicates the site of enzymatic hydrolysis. Yellow colour formation only occurs upon enzymatic cleavage if the reaction conditions are sufficiently alkaline to convert the *p*NP product into its ionic form.

Coupled enzyme assays

When the action of an enzyme cannot be detected on its own, it is common to combine it with one or more secondary enzymes to generate a measurable product. The co-factor NADH is frequently used in these types of coupled spectrophotometric enzyme assays, since it absorbs UV light when reduced, but not in its oxidised form. In this work, GE activity was assayed in a coupled mode together with uronate dehydrogenase (UDH) on commercially available ester substrates of uronic acids (GlcA or GalA). The assay is based on esterase-catalysed uronic acid formation, followed by UDH-mediated oxidation to D-glucarate or D-galactarate in the presence of NAD^+, with the simultaneous formation of reduced NADH (Figure 17). NADH absorbs UV light at 340 nm, and the increase in absorbance is proportional to the GE activity (Sunner et al., 2015).

The GE assay as described by Sunner et al. (2015) was further developed for operation in continuous mode (**Paper IV**). It was used for screening of putative GEs and to quantify GE activity on four different model substrates (**Papers IV, V**) (Table 2). The assay is sensitive and rapid, and provides information on how small changes in the model substrates affect the activity and specificity of the enzymes. However, the substrates must be simple, soluble, stable (to avoid spontaneous hydrolysis) and not generate high background absorbance levels. It was not possible to use complex or biologically relevant substrates in this assay.

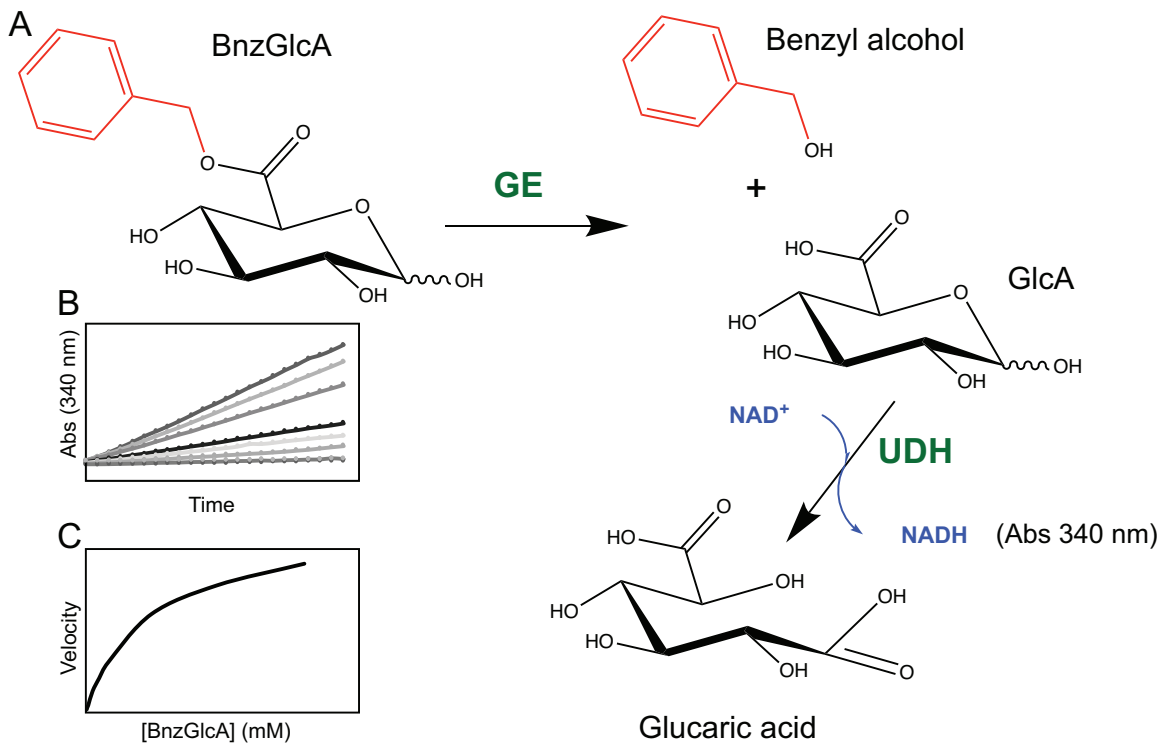


Figure 17. Overview of the reaction in the coupled GE assay using benzyl glucuronoate (BnzGlcA) as substrate (**A**). The GE-catalysed hydrolysis of the benzyl ester allows UDH to form glucaric acid, with the simultaneous formation of NADH. The increase in absorbance upon NADH formation is detected at 340 nm. The absorbance is monitored over time, with increasing concentrations of BnzGlcA at a fixed enzyme concentration (**B**). Plotting the reaction rate of GE-catalysed hydrolysis as a function of BnzGlcA concentration (obtained from linear regression of the functions in B), allows the kinetic parameters (K_M and k_{cat}) to be estimated using the Michaelis-Menten equation (**C**).

5.3 Thin-layer chromatography

Thin-layer chromatography (TLC) is a qualitative chromatographic method, in which a mobile phase (the reaction mixture) flows through a stationary phase (typically a coating of silica gel), carrying the components of the mixture with it. Different components in the mobile phase travel at different rates and can be visualised as spots by staining or UV light (Spangenberg et al., 2010). TLC was used to detect GE activity coupled with a GH115 α-glucuronidase on extracted LCC fractions (Table 2) (unpublished results from the collaboration with Professor Master's group at Toronto University). The 4-*O*-glucuronic acid produced was visualised with aniline-diphenylamine-phosphoric acid stain (Figure 18). (The results are further discussed in Chapter 7.) TLC is useful as a qualitative detection method, but is not suitable for quantitative or kinetic measurements. Moreover, standards are required to identify the products formed. The advantage of TLC in this work, was the rapid confirmation of GE activity on native LCC material with an easily detectable product. TLC also served as a complement to more advanced and quantitative analyses.

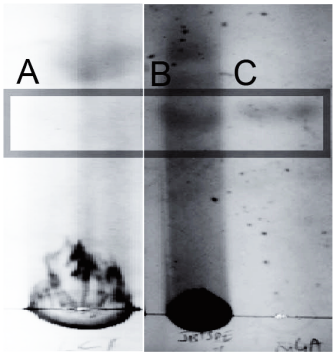
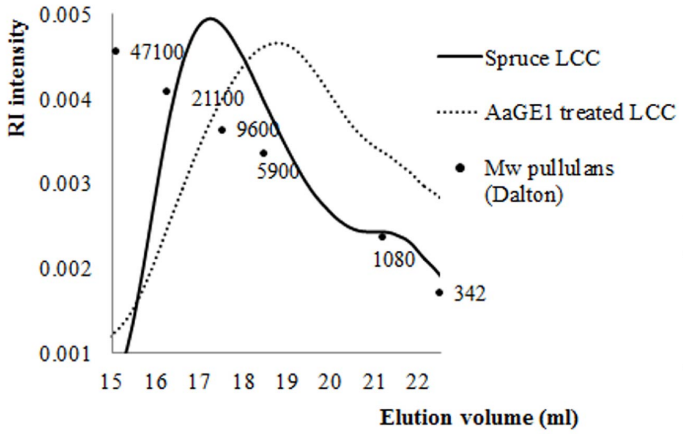


Figure 18. Qualitative TLC analysis of the 4-*O*-MeGlcA released (indicated by the grey box). Untreated LCCs (without 4-*O*-MeGlcA) (**A**) are compared to LCCs incubated with a GE and an α-glucuronidase (**B**), where the band corresponding to the 4-*O*-MeGlcA standard (**C**) is visible.

5.4 Size-exclusion chromatography

The molecular weight distribution of reaction products can be investigated by size-exclusion chromatography (SEC). SEC separates molecules by size as they pass through a porous matrix in a column, and the size distribution can be monitored by refractive index (RI) or UV detectors. Small molecules flow slowly through the column, since they penetrate deep into the pores of the matrix material, whereas larger molecules that are unable to enter the pores pass through the column faster. This means, the smaller the molecule, the longer the retention time. In the work described in **Paper III**, SEC equipped with an RI detector was used to qualitatively confirm the separation of lignin from carbohydrate fractions upon GE treatment (Figure 19). SEC was also used to confirm mannan polymer degradation and the production of mannooligosaccharides by two *endo*-mannanases (**Paper II**) (Table 2). SEC is an appropriate tool for detecting enzyme action that results in larger degradation products, distributed over a wide molecular weight range. It is not possible to obtain detailed information on substrate specificity, but SEC is a useful tool to confirm enzyme activity, and is a complement to other detection methods. A limitation of the technique is that the standards required (often linear pullulan polymers) do not mirror the polydispersity of lignocellulosic materials (Baumberger et al., 2007).



	M_w (Dalton)
Spruce LCC	9600
AaGE1 treated LCC	7500

Figure 19. RI detector response from SEC analysis of spruce LCC, before (solid line) and after (dotted line) GE hydrolysis. The peak shifts towards lower molecular weight (Mw) (higher elution volume), which indicates the release of lignin from carbohydrates as a result of GE action. The table gives the molecular weight of the LCC fraction before and after GE treatment.

5.5 Anion-exchange chromatography

High-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) is a highly sensitive method for qualitative and quantitative carbohydrate analysis. Analysis is performed under strongly alkaline conditions, where carbohydrates are deprotonated, followed by interaction with the anionic column (Cataldi et al., 2000). HPAEC-PAD analyses were performed for the quantification of enzymatically released monosaccharides from native and STEX-treated wood chips (**Paper I**), and from GE-supplemented hemicellulose reactions with corn cob biomass (**Paper IV**). HPAEC-PAD was also used to obtain qualitative product profiles of mannanase-hydrolysed mannan polymers, and to quantify selected mannooligosaccharide products (**Paper II**) (Table 2). HPAEC-PAD is a useful tool to verify and quantify enzyme activity on complex polymers and biomass material, but known standards must be included to obtain qualitative and quantitative information about the products formed. The availability of standards is therefore often a limiting factor (as further discussed in **Paper II)**.

5.6 Mass spectrometry

Mass spectrometers convert analyte molecules to a charged state, with subsequent analysis of the ions and fragment ions that are produced during the ionisation process. There are three key stages in mass spectrometry (MS): ionisation, acceleration/deflection and detection. In the first stage, molecules are vaporised and converted into ions. The subsequent acceleration and deflection sort the ionised analytes according to mass. The detection of the analytes is thus based on the mass-to-charge ratio (*m/z*), visualised in a mass spectrum with signals for each *m/z*. MS is commonly used to analyse carbohydrates, and ionic species can be obtained from the liquid or solid state. Several technologies are available for both ionisation and ion analysis, resulting in many different types of mass spectrometer (Bauer, 2012). It is also common to combine MS as a detector in chromatographic analyses, e.g. liquid chromatography, to exploit the benefit of MS, being more sensitive and specific than other chromatographic detectors (Pitt, 2009).

There are many advantages of using MS to investigate enzyme action on complex carbohydrates, for example the possibility of obtaining detailed information on enzymatic cleavage patterns in terms of fragment size, linkages and branching. Already fragmented ions can go through a second round of MS (MS^2), to elucidate additional structural features of enzyme substrates/products (e.g. the position of acetyl groups). Quantification of oligosaccharides by MS is, however, difficult, as ionisation efficiencies can differ significantly. For absolute quantification, internal standards and pure model standards of the analytes are required. Electrospray ionisation-MS (ESI-MS) was applied to study the acetylation pattern of enzymatically produced mannooligosaccharides (**Paper II**). The 4-*O*-MeGlcA released after the hydrolysis of LCCs by a GE and a GH115 α-glucuronidase was analysed and quantified using liquid chromatography combined with ESI-MS (unpublished data, described in Chapter 7) (Table 2).

5.7 Nuclear magnetic resonance spectroscopy

NMR spectroscopy is a technique that detects chemical structures by observing local magnetic fields around atomic nuclei (nuclei with an odd atomic number). The spins of these nuclei are randomly oriented, but when placed in an external magnetic field, the spins will align parallel or antiparallel to the field. If the nuclei are subjected to electromagnetic radiation of the correct frequency, the nuclei aligned with the external field will absorb

energy and "spin-flip" to align themselves against the field. When the spin returns (relaxes), radio waves are emitted. However, as the magnetic properties of a nucleus depend on its chemical environment, the emitted radio wave frequency of a chemically bound nucleus will be shifted compared to that of a free nucleus. This so-called "chemical shift" is used to identify the chemical composition of a molecule, and is expressed as the ratio of the shifted frequency to a reference frequency of the same nucleus. NMR spectra are commonly shown as the absorption of radio waves as a function of chemical shift in parts per million, and each chemical structure will have a unique NMR spectrum (James, 1998).

Different NMR methods are commonly used for structural determination of LCCs, and are promising in detecting key structures in heterogeneous complex materials (Ralph & Landucci, 2010). However, a drawback of NMR is its relatively low sensitivity, which limits the detection of less abundant lignocellulosic structures such as LC bonds. Extraction procedures are therefore often required to enrich the LC bonds in the sample, and sometimes derivatisation is necessary. In the work described in this thesis, ^{31}phosphorus NMR spectroscopy (^{31}P NMR) and heteronuclear single quantum coherence spectroscopy (HSQC) were used to study GE-mediated cleavage of LC ester bonds (**Paper III** and unpublished data presented in Chapter 7).

^{31}P NMR

^{31}P NMR is useful in determining the overall distribution of functional groups in a lignin/LCC sample. This 1D NMR technique allows for the detection of OH groups attached to aliphatic, phenolic and carboxylic moieties (COOH), after derivatisation with a phosphitylation reagent (Argyropoulos, 2010) (Figure 20). The ability to quantify COOH is particularly interesting when studying the enzymatic hydrolysis of LC esters (formation of new GlcA-ended groups). However, it is not possible to determine which structures these esters are attached to with ^{31}P NMR, or whether they are of α [benzyl]- or γ-type. Complementary analyses must be included to confirm, for example, that acetyl-linked esters do not contribute to the COOH signals in the spectrum. ^{31}P NMR was used to measure the increase in carboxylic acid content in LCCs as a result of GE hydrolysis (**Paper III**) (Table 2).

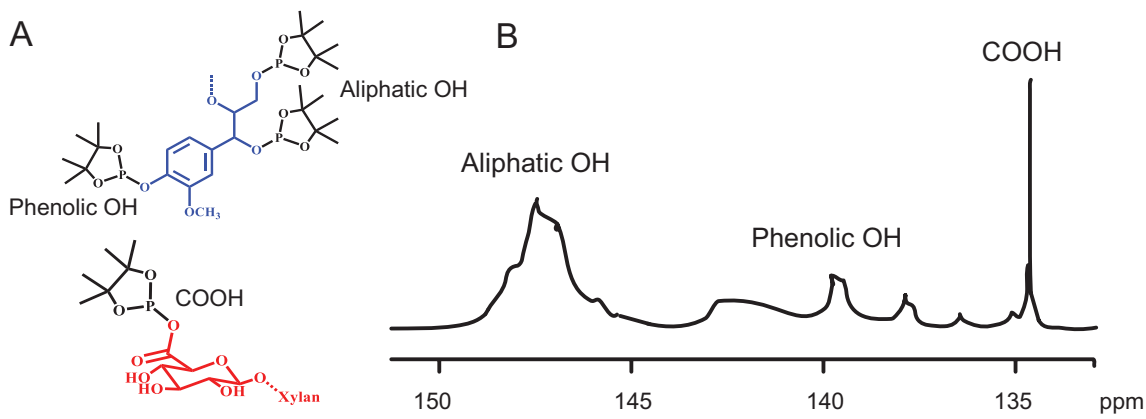


Figure 20. Derivatisation products after phosphitylation of LCCs (after GE-mediated ester cleavage) (**A**), and schematic representation of peaks assigned to aliphatic, phenolic and carboxylic (COOH) OH groups in a ^{31}P NMR spectrum (**B**).

HSQC

The development of different types of correlating 2D NMR techniques, e.g. HSQC, has been important for the analysis of LCC structures. HSQC correlates chemical shifts of directly bound nuclei of different types (normally ^{1}H and ^{13}C), where the chemical shift of the carbon lies in one dimension (F1) and the chemical shift of its attached proton in the other dimension (F2). Correlating NMR techniques make it possible to decode structural details and linkages of LCCs. Moreover, the 2D-dimensionality of the method reduces the number of overlapping signals from the LCC structure from one type of nucleus (^{1}H) by correlating it with separated signals from the other nucleus (^{13}C), resulting in better resolution than in 1D NMR (Ralph & Landucci, 2010).

2D NMR could, in theory, provide structural information on LC ester bonds that cannot be obtained with ^{31}P NMR. However, the signal assigned to the γ-ester is still heavily overlapped by other lignocellulosic structures in the analysed samples, and the detection of benzyl esters has seldom been reported (Balakshin et al., 2011; Balakshin et al., 2007; Giummarella & Lawoko, 2016; Yuan et al., 2011). LCC-extraction procedures and/or the sensitivity of the analysis will have to be improved in order for 2D NMR to be feasible in assaying GE activity. HSQC was applied in the study described in **Paper III** to verify the ^{31}P NMR results, by demonstrating that the increase in COOH content was not the result of acetyl esterase activity. Attempts were also made to apply HSQC for the detection of GE hydrolysis of LC esters extracted from Japanese beech (*Fagus crenata*) (unpublished data from the collaboration with Professor Watanabe's group at Kyoto University, presented in Section 7.8) (Table 2). HSQC spectra obtained from this LCC sample surprisingly displayed a clear signal identified as the benzyl ester (Figure 21).

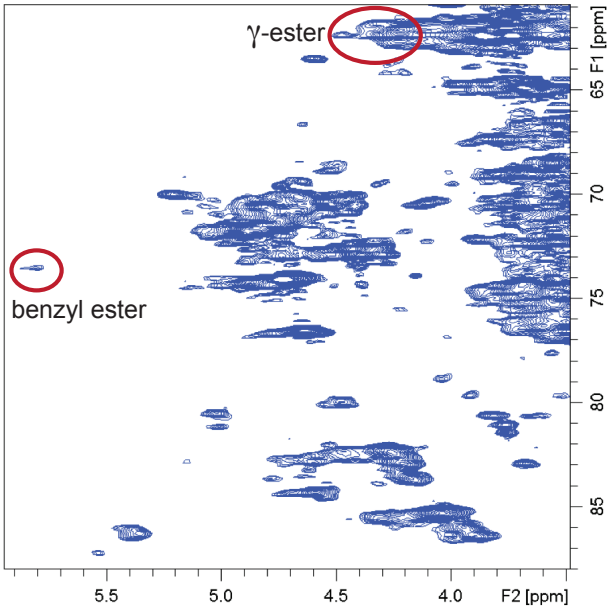


Figure 21. HSQC spectrum of LCCs isolated from Japanese beech (F1=^{13}C chemical shift, F2=^{1}H chemical shift). The red ellipses indicate the signals assigned to LC ester bonds of α [benzyl]- and γ-type. The signal from the γ-ester is heavily overlapped by other lignin-carbohydrate structures. (Figure courtesy of Hiroshi Nishimura.)

Table 2. Methods used to assay enzymatic activity in the present work

Method	Substrate	Purpose	Reported in:
Reducing-sugar assay	Mannan polymers	Obtaining time courses for mannanase hydrolysis	Paper II
Chromogenic assay	*p*NP-Ac	Screening for acetyl esterase activity of GEs	Papers IV, V
Coupled assay	Esters of uronic acids	Screening and quantification of GE activity	Papers IV, V
TLC	LCCs	Detection of GE-catalysed release of 4-*O*-MeGlcA	Chapter 7
SEC	Mannan polymers/LCCs	Detection of decrease in molecular weight upon mannanase/GE hydrolysis	Papers II/ III
HPAEC-PAD	Wood chips	Quantification of enzymatically released monosaccharides from wood chips	Paper I
	Mannan polymers	Detection and quantification of enzymatically produced mannooligosaccharides	Paper II
	Corn cob biomass	Quantification of the increase in monosaccharide release upon GE action	Paper IV
ESI-MS	Mannan polymers	Detection of acetylation patterns of mannanase-hydrolysed oligosaccharides	Paper II
	LCCs	Quantification of GE-catalysed release of 4-*O*-MeGlcA	Chapter 7
^{31}P NMR	LCCs	Detection of LC ester cleavage upon GE treatment	Paper III
HSQC	LCCs	Detection of LC ester cleavage upon GE treatment/	Chapter 7
		Check for acetyl esterase activity of a GE	Paper III

6. Enzymes for selective mannan processing

Upgrading woody biomass includes the valorisation of hemicelluloses, which are today often completely or partially degraded in industrial processes, in favour of the extraction of cellulose fibres. Process alternatives in which hemicelluloses are utilised must be further developed, and enzymes that specifically modify or decompose hemicelluloses into valuable building blocks could serve as important tools in this context. To evaluate the potential of hemicellulolytic enzymes, it is necessary to understand how these enzymes act on complex heterogeneous polysaccharides of industrial relevance. The major hemicellulose in softwood is galactoglucomannan (GGM), making up almost 20 % of the dry weight (Sjöström, 1993), and attention was therefore directed to studies of *endo*-mannanases for the selective degradation of GGM (**Paper II**). This chapter presents and discusses the results regarding the way in which the chemical structure of spruce GGM, including backbone heterogeneity, galactosylation and acetylation, influences mannanase action.

6.1 Mannan polymers

Softwood galactoglucomannan

Mannans exist in all plant cell walls at various amounts, and they are structurally and chemically diverse. Softwood GGM consists of a backbone of β-1,4-linked mannose (Man) and glucose (Glc) units decorated with α-1,6-linked galactose (Gal) moieties, believed to be exclusively bound to the mannose units. Softwood GGM is usually divided into Gal-rich and a Gal-poor types, with Man:Glc:Gal ratios of 3:1:1 and 4:1:0.1, respectively (Sjöström, 1993). The mannose units are partly *O*-acetylated at C2 and C3 at a degree of 0.2-0.3 (Figure 22A), but these ratios can vary significantly depending on source and extraction method (Capek et al., 2002; Hannuksela & Herve du Penhoat, 2004; Willför et al., 2008; Xu et al., 2010). In the study reported in **Paper II**, hot-water-extracted GGM from spruce wood (SpGGM), with a low degree of substitution by acetylation (DSac) of 0.13, was used as substrate.

Mannans as model substrates

Mannan polymers are found not only in plant cell walls, but can also serve as storage polysaccharides, e.g. in the tubers of konjac and locust bean seeds. Konjac glucomannan (KGM) and locust bean gum galactomannan (LBG) are commercially available, well-defined mannans, and therefore commonly used as substrates for evaluating *endo*-mannanase activity (Malgas et al., 2015). KGM and LBG were used as complementary model substrates in this work (**Paper II**). They both share structural features with SpGGM; KGM having a heterogeneous glucomannan backbone and LBG containing galactose

decorations (Figure 22B-C). KGM contains a low number of mannose/glucose decorations on the backbone, and in this work was regarded as a linear polysaccharide compared to the highly substituted LBG and SpGGM. LBG lacks acetylation, but KGM has a DSac of 0.1 (Dey, 1978; Kato & Matsuda, 1969), similar to the DSac of the extracted SpGGM (DSac 0.13) (Table 3).

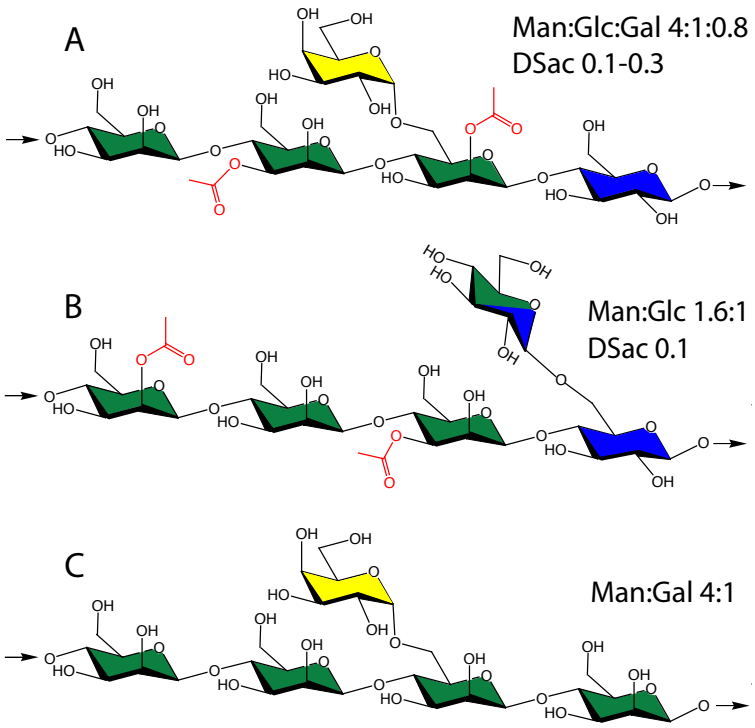


Figure 22. Chemical structures of mannan polymers. Spruce *O*-acetyl-galactoglucomannan (**A**), konjac glucomannan (**B**), and locust bean gum galactomannan (**C**). C2 and C3 are potential acetylation sites in native SpGGM and KGM (indicated in red). The average ratios of mannose (green), glucose (blue) and galactose (yellow), and the values of DSac are given for each mannan polymer. KGM and LBG were subjected to further chemical acetylation, making C6 also a possible acetylation site.

Chemical acetylation of mannans

KGM and LBG have both served as substrates for chemical acetylation to study the changes in gelation behaviour and microbial degradability of mannan polymers (Bi et al., 2016; Huang et al., 2002). In the present work, KGM and LBG were acetylated to different degrees with the purpose of investigating the enzymatic hydrolysability of artificially and highly acetylated mannan polymers in relation to natively acetylated SpGGM and KGM (**Paper II**). Naturally acetylated mannan substrates have acetyl groups at the C2 and C3 positions, but chemical acetylation allows the C6 position to also be acetylated. Two levels of acetylation were investigated. The resulting mannans are denoted KGM_A/LBG_A for DSac 0.7-0.8 and KGM_B/LBG_B for the higher DSac (~2) (Table 3). The solubility of KGM and LBG decreased upon increased chemical acetylation, and the chemically added acetyl groups were only present in the insoluble fractions. The hydrophobicity of the mannan substrates was suggested to be an effect of the unevenly distributed chemically introduced acetyl moieties on the KGM and LBG.

Table 3. DSac and solubility of native and chemically acetylated mannan substrates (**Paper II**)

Substrate	Total fraction	Soluble fraction		Insoluble fraction		MW (kDa)
	DSac[A]	Mass %[B]	DSac[C]	Mass %[D]	DSac[E]	
KGM	0.09	84	NM*	16	NM	1000
KGM_A	0.7	48	0.1	52	1.8	
KGM_B	2.1	15	NM	85	NM	
LBG	0	66	NM	34	NM	1000
LBG_A	0.8	41	0.1	59	1.5	
LBG_B	1.9	13	NM	85	NM	
SpGGM	0.13	NM	NM	NM	NM	30

[A] Quantified by HPLC: four individual measurements with standard deviations < 10 %.
[B, D] Gravimetrically measured: triplicate measurements with standard deviations < 10 %.
[C, E] Quantified by HPLC: triplicate measurements with standard deviations < 17 %.
*Not measured

6.2 Mannanolytic enzymes

Multiple enzyme activities are required for the complete degradation of the heterogeneous and highly decorated SpGGM. These include *endo*-acting β-mannanases, *exo*-acting β-mannosidases, *endo*-acting β-glucanases, *exo*-acting β-glucosidases, debranching (*exo*-acting) α-galactosidases and acetyl esterases; activities that can be found in several GH and CE families (Figure 23) (Srivastava & Kapoor, 2017; Van Zyl et al., 2010). Synergistic behaviour has been reported between debranching α-galactosidases and main-chain-cleaving β-mannanases (Malgas et al., 2015; Wang et al., 2010). Synergism between β-mannanases is also likely, despite the fact that they share the ability to cleave glycosidic bonds between mannose units, due to large variations in substrate specificity (as observed for endoglucanases) (Van Dyk & Pletschke, 2012).

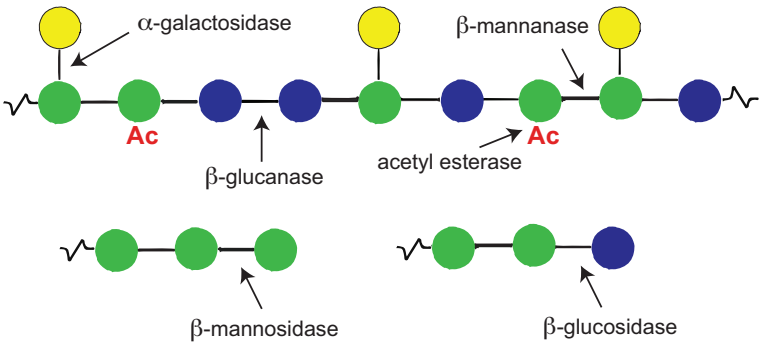


Figure 23. Schematic illustration of *O*-acetylated SpGGM (mannose=green, glucose=blue, galactose=yellow, acetyl groups=Ac) and the enzymes involved in the complete degradation of the polysaccharide, and the oligosaccharides released. Arrows indicate the possible sites of enzymatic attack.

6.3 β-Mannanases

Classification and function

β-Mannanases (hereafter referred to as mannanases) are the major backbone-cleaving enzymes that randomly hydrolyse mannans and heteromannans of DP > 3 (Figure 23), producing smaller oligosaccharide fragments (mainly mannobiose and mannotriose) (Dhawan & Kaur, 2007). Mannanases are omnipresent in a wide range of microorganisms isolated from different habitats, and have been classified into families GH5, GH26, GH113 and GH134 in CAZy (www.cazy.org; (Lombard et al., 2014)), where GH5 and GH26

comprise the majority of all characterised mannanases. Although mannanases from different families are diverse in terms of their protein sequence, they are similar in their three-dimensional arrangement, sharing the $(\beta/\alpha)_8$-barrel fold, and are grouped in clan GH-A (Dhawan & Kaur, 2007).

Both bacterial and eukaryotic mannanases have been classified as GH5, whereas GH26 mannanases are predominantly of bacterial origin (Dhawan & Kaur, 2007; Srivastava & Kapoor, 2017). Some organisms produce multiple mannanases belonging to both GH5 and GH26, such as the fungus *Podospora anserina* (Couturier et al., 2013) and the bacteria *Bacillus licheniformis* (Ethier et al., 1998; Songsiriritthigul et al., 2010) and *Cellvibrio japonicus* (Hogg et al., 2003; Hogg et al., 2001). The functional differences between mannanases encoded by the same species are often ambiguous. However, different functions and biological roles have been suggested for five mannanases encoded by *C. japonicus* (*Cj*Man5A-C and *Cj*Man26A&B) based on studies on oligosaccharides and well-defined polymers. The actions of *Cj*Man5A and *Cj*Man26A were further studied in this work, and evaluated in detail for the first time, employing polymeric and decorated mannan substrates (**Paper II**).

Structure and modularity

Structural determination of mannanases from both GH5 and GH26 has revealed solvent-exposed active sites, with two strictly conserved glutamate residues serving as acid/base and nucleophilic catalytic residues (Gilbert et al., 2008; Van Zyl et al., 2010). The open cleft-shaped active site comprises at least four subsites that bind to specific sugars in the polysaccharide, close to the site of the cleavage. The number of subsites and the substrate specificity of the subsites have been reported to vary between the mannanases investigated (Srivastava & Kapoor, 2017). An example is *Bs*Man26A (from *Bacillus subtilis*), which has a strong preference for mannose residues in subsite -2, whereas *Ba*Man5A (from *Bacillus agaradhaerens*) can accommodate either glucose or mannose in this subsite (Tailford et al., 2009). The same behaviour has been reported for *Cj*Man26A and *Cj*Man5A; *Cj*Man26A having a higher preference for mannose at subsite -2 than *Cj*Man5A. The binding of *Bs*Man26A and *Cj*Man26A to mannose is facilitated by polar interactions between the active site and the axial OH-group of C2, not present in their GH5 counterparts (Hogg et al., 2003). GH5 mannanases have therefore been suggested to be more promiscuous with regard to the backbone composition of mannan polymers (Table 4).

Mannanases can be modular, commonly with a catalytic module linked to one or several CBMs with affinity for the mannan backbone (e.g. CBM23, 27 and 35) or cellulose (e.g. CBM1) (Dhawan & Kaur, 2007; Gilbert et al., 2008; Von Freiesleben et al., 2018). A few mannanases are further fused to other catalytic modules (e.g. cellulase or xyloglucanase). Many mannanases in both GH5 and GH26 harbour only one individual catalytic module, and it is in general difficult to observe any patterns in modularity amongst mannanases (Gilbert et al., 2008).

Functional differences of mannanases

Variations in active site architecture and modularity have led to the suggestion of functional differences between GH5 and GH26 mannanases, but a larger number of mannanases from different species will have to be studied to identify and ascertain the suggested functions. Regarding the mannanases encoded by *C. japonicus*, all the GH5 enzymes are modular with CBMs. Additionally, they are secreted and display active sites with six or more subsites, allowing longer substrates to bind. These features suggest that they act on polymeric

mannan. *Cj*Man26A&B, on the other hand, lack CBMs, are membrane bound and have fewer subsites. They have therefore been proposed to favour soluble mannooligomeric substrates (Table 4) (Hogg et al., 2003). The actions of *Cj*Man5A and *Cj*Man26A have been investigated through studies on mannooligosaccharides and commercially available mannan polymers (Hogg et al., 2003; Hogg et al., 2001). However, in order to explain their real different biological roles, they will have to be evaluated on complex and biologically relevant mannan polymers, i.e., substrates that are of industrial interest.

6.4 Mannanase action on complex mannan polymers

The performance of *Cj*Man5A and *Cj*Man26A was evaluated on the biologically and industrially relevant SpGGM (**Paper II**). Detailed characterisation was also carried out on KGM and LBG, used as complementary model substrates to SpGGM. Mannanase reactions were performed with SpGGM, KGM or LBG, and the results were analysed using various complementary methods. Reducing-sugar measurements provided reaction rate profiles (discussed in detail in **Paper II**), while SEC and HPAEC-PAD allowed for the detection of both longer and shorter oligosaccharides produced by the enzymes. SEC analysis confirmed that both *Cj*Man5A and *Cj*Man26A hydrolysed the mannan polymers into mannooligosaccharides with DP < 10. HPAEC-PAD was used to detect smaller hydrolysis products, and quantification of these products was possible using standards for mannose (M1), mannobiose (M2), mannotriose (M3) and mannotetraose (M4). The limited availability of standards and problems associated with co-elution prevented the quantification of longer and heterogeneous oligosaccharide products.

Mannanase cleavage pattern of SpGGM

The product profiles provided by HPAEC-PAD differed significantly, not only between the two enzymes, but also between the substrates. Several unidentified product peaks were unique to the individual mannan polymers, in agreement with their different chemical structures (Figure 24A, C, E). The levels of M1-M4 released (the products for which standards were available) also varied between the enzymes and mannan polymers (Figure 24B, D, F). SpGGM was mainly hydrolysed to M3 and M4 by *Cj*Man5A, while *Cj*Man26A produced large amounts of M1 and M2, similar to the cases for KGM and LBG (Table 4).

The hydrolysis profiles and product quantification for SpGGM, KGM and LBG showed common features, as expected, due to the structural similarities of these substrates. However, some differences were also observed. *Cj*Man26A-mediated hydrolysis of SpGGM and LBG resulted in a significant increase of produced M2 compared to the hydrolysis of KGM. The ability of *Cj*Man26A to produce higher levels of M2 from SpGGM and LBG is likely to be an effect of the higher content of mannose in the backbone compared to KGM, which favours the affinity of *Cj*Man26A to the substrate. *Cj*Man26A generated a substantial amount of M3 and M4 from SpGGM, which was not observed for either KGM or LBG. The reason for this is not clear, but could be a combination of glucose in the backbone, branching galactosyl moieties and acetyl groups positioned in such a manner that the ability of the enzyme to produce smaller oligosaccharides is hindered.

The cleavage patterns of *Cj*Man5A differed only marginally between SpGGM and KGM/LBG, showing the release of M3>M4>M2>M1, in decreasing concentrations. This is in agreement with *Cj*Man5A having a low specificity to the backbone composition. The proposed functional differences between the enzymes, i.e., *Cj*Man26A being able to

hydrolyse shorter oligosaccharides, and *Cj*Man5A able to accommodate either mannose or glucose in its active site, were supported by the results from all product profiles reported in **Paper II.** When applying both mannanases together on KGM/LBG, the peaks identified as M3 and M4 disappeared, while the M1 and M2 peaks increased (unpublished data). This demonstrated the ability of *Cj*Man26A to act on shorter oligosaccharides (produced by *Cj*Man5A), as reported previously (Table 4). Most importantly, these results clearly demonstrate that mannanase action is highly substrate dependent, and underline the necessity of evaluating these enzymes on a range of different mannan polymers in order to understand their modes of action.

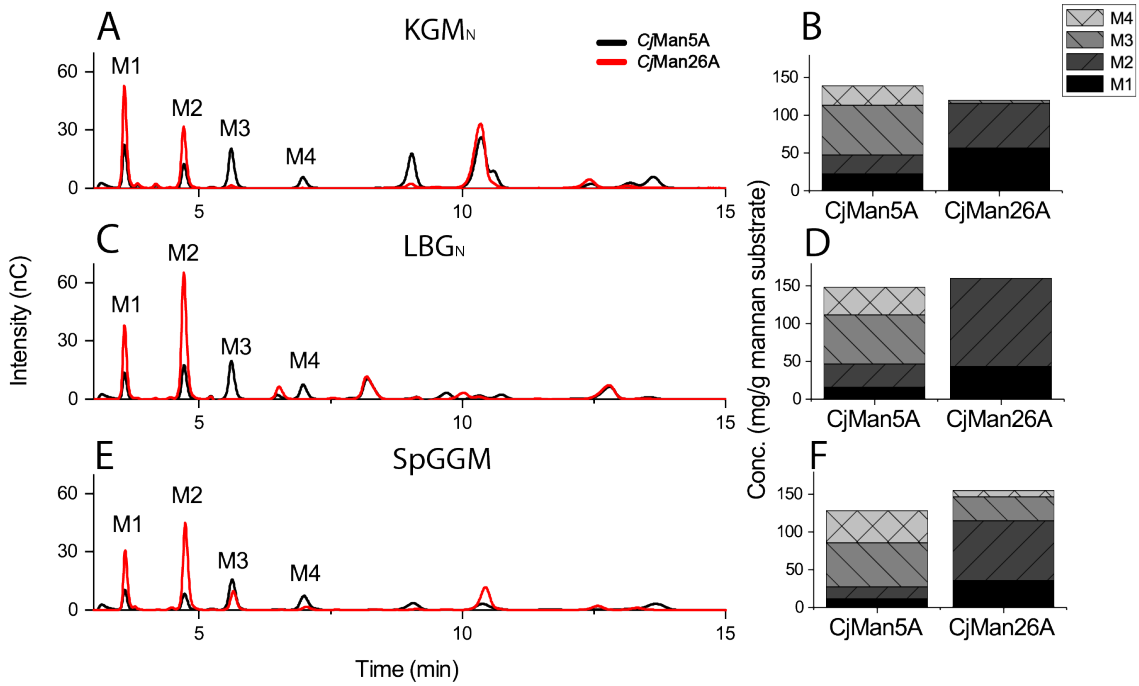


Figure 24. Oligosaccharide product profiles obtained from mannanase-catalysed reactions using HPAEC-PAD (left panel), and quantification of M1-M4 oligosaccharides (right panel) after 24-h reactions. The reactions were performed with either *Cj*Man5A (black) or *Cj*Man26A (red) on KGM (**A, B**), LBG (**C, D**) or SpGGM (**E, F**). The quantification data show that *Cj*Man26A produces mainly M1 and M2, whereas *Cj*Man5A produces mainly M3 and M4, from all the mannan polymers investigated. The product profiles show differences in both quantified and unidentified products between the substrates and the two enzymes.

6.5 The influence of acetylation on mannanase activity

SpGGM is decorated with galactose moieties and acetyl groups. The effect of these main chain decorations on the action of different mannanases have been reported briefly (Tenkanen et al., 1997; Tenkanen et al., 1995; Von Freiesleben et al., 2016; Von Freiesleben et al., 2018), but it is still not known in detail how the acetylation of mannan polymers influences mannanase hydrolysis. Reactions with *Cj*Man5A and *Cj*Man26A on the naturally acetylated substrates SpGGM and KGM enabled the comparison of mannanase hydrolysis on branched and linear mannan polymers with similar DSac (~0.1) (**Paper II**). Oligomeric mass profiles (obtained using ESI-MS) provided information on enzymatically produced hexooligosaccharides in terms of length and the number of acetyl groups linked to each oligosaccharide.

The resulting SpGGM spectra showed that *Cj*Man26A was able to produce acetylated disaccharides and diacetylated trisaccharides, suggesting it was more tolerant to acetyl groups than *Cj*Man5A, which was only capable of producing monoacetylated oligosaccharides (Figure 25, Table 4). The product profiles for KGM, on the other hand, showed that *Cj*Man5A was capable of producing smaller acetylated mannooligosaccharides to a higher extent than *Cj*Man26A, whereas no diacetylated products were observed with either of the enzymes. These results are ambiguous, but indicate that the low DSac (~0.1) probably did not affect either of the enzymes significantly. Structural factors, such as the heterogeneity of the backbone or the presence of galactosyl side groups, probably have a greater influence on the substrate specificity and enzyme action.

The insignificant effect of acetylation on mannanase action was confirmed by studying the potential boosting effect of an acetyl esterase from *Clostridium thermocellum* (*Ct*Axe2A) on *Cj*Man5A or *Cj*Man26A action. In agreement with the ESI-MS results, a low degree of acetylation seemed to be a negligible obstacle for both *Cj*Man5A and *Cj*Man26A, especially in the presence of galactose moieties. Only a minor initial increase in activity was observed for *Cj*Man5A on linear KGM when supplemented with *Ct*Axe2A (Figure 5 in **Paper II**).

The boosting effect of an acetyl esterase on enzymatic SpGGM hydrolysis has been reported previously (Tenkanen et al., 1995), indicating that acetyl groups can affect the hydrolysability of decorated mannan polymers with higher DSac (~0.3) than the substrates used in the present work. The effect of acetyl groups on mannanase hydrolysis has also been demonstrated recently by Von Freiesleben et al. (2018). Ten fungal mannanases were applied to acetylated and deacetylated SpGGM, and a higher initial rate was observed for several of the mannanases on the deacetylated SpGGM than with acetylated SpGGM. Some of the enzymes were, however, not affected by the degree of acetylation, which indicates diverse substrate preferences amongst mannanases.

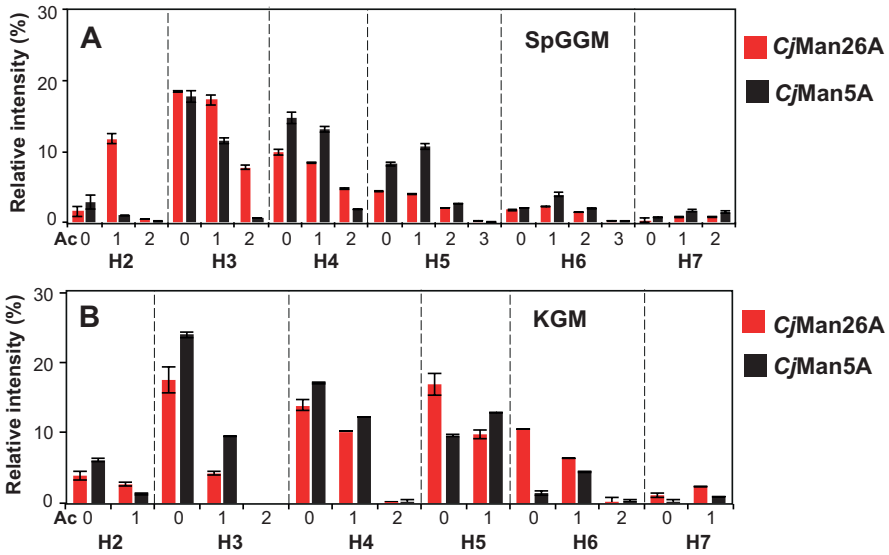


Figure 25. Oligomeric mass profiling of SpGGM (**A**) and KGM (**B**) after 24-h hydrolysis with either *Cj*Man5A (black) or *Cj*Man26A (red) (adapted from Figure 4, **Paper II**). H denotes the number of hexoses in the oligosaccharide products and Ac the number of acetylations. *Cj*Man26A was able to produce acetylated disaccharides and diacetylated trisaccharides from SpGGM, in contrast to *Cj*Man5A, which only produced monoacetylated oligosaccharides. The product profiles of KGM, on the other hand, showed that *Cj*Man5A was capable of producing a larger number of smaller acetylated mannooligosaccharides than *Cj*Man26A. No diacetylated products were observed for either of the enzymes on KGM.

Chemically acetylated mannans and the effect on hydrolysability

In order to study the effect of higher DSac on mannanase performance, acetylated mannan polymers were chemically prepared using commercially available KGM and LBG polymers (as described in Section 6.1). Chemical acetylation indeed resulted in reduced substrate hydrolysability compared to the native polymers (with both *Cj*Man5A and *Cj*Man26A), measured with the reducing-sugar assay (**Paper II**). Acetylated KGM_A, with DSac (~0.7), showed a 20 % reduction in hydrolysability compared with native KGM, and the highly acetylated KGM_B, with DSac (~2.1), could not be hydrolysed (within the detection limit of the assay). The reduction in hydrolysability was even more evident for the decorated LBG substrate, which clearly shows how acetylation, especially combined with galactosyl substitution, prevents the enzymes from acting on the polymers. The low substrate hydrolysability of the chemically acetylated KGM and LBG is likely to be an effect of the previously mentioned reduced solubility. It is also possible that the introduction of acetyl groups at the C6 position of the sugars, not acetylated in native mannans, may affect the ability of the enzymes to act. These results clearly show that chemical acetylation is an effective means of limiting biodegradation, but it changes the solubility of the mannans.

6.6 Mannanases in a wood-based biorefinery

SpGGM can be hydrolysed into fermentable sugars, but has great potential in oligomeric form in high-value applications such as prebiotics and food additives. Specific mannanases have the potential to perform partial and controlled hydrolysis to a desired DP, but also to alter the solubility and viscosity of mannan polymers. The two mannanases studied in this work (**Paper II**) produced a range of oligosaccharides from the softwood-derived SpGGM (from 1-10 hexose units). *Cj*Man5A generated high amounts of M3 and M4, and could therefore be a suitable candidate for the production of platform molecules. *Cj*Man26A, on the other hand, produced a large proportion of monosaccharides and disaccharides, already during the first few minutes of the reaction, and might be better suited for the production of fermentable sugars or for boosting the activity of other lignocellulolytic enzymes (Table 4).

The mannanases studied here may not be optimal for the valorisation of SpGGM, but it is important to use authentic lignocellulosic substrates rather than simplified and well-defined model substrates in order to identify mannanases appropriate for specific industrial applications, as demonstrated here. Continued characterisation of mannanases on wood-derived polymers is important in the search for enzymes with substrate specificities suitable for use in the processing of woody biomass to desired end products.

Table 4. General substrate preferences of *Cj*Man5A and *Cj*Man26A together with the main products generated by the hydrolysis of SpGGM (**Paper II**)

General substrate preferences	***Cj*Man5A**	***Cj*Man26A**	**References**
Backbone unit	Man/Glc	Man	Hogg et al., 2003
DP[1]	> 4	> 2	Hogg et al., 2003
Products from SpGGM	***Cj*Man5A**	***Cj*Man26A**	
Main mannooligosaccharides[2]	M3/M4	M2	Paper II
Monoacetylated hexooligosaccharides[3]	H3/H4	H2/H3/H4	Paper II
Diacetylated hexooligosaccharides[3]	-	H3/H4	Paper II

[1] Required for $kcat/K_M > 10^5$ $min^{-1}M^{-1}$
[2] Products for which standards were available, quantified with HPAEC-PAD
[3] With DP < 5, detected with ESI-MS

7. Enzymatic decoupling of lignin and hemicelluloses

The covalent linkages between lignin and hemicelluloses constitute a significant obstacle for efficient processing of woody biomass and the extraction of wood polymers. Glucuronoyl esterases (GEs), enzymes that specifically cleave lignin-carbohydrate (LC) ester linkages (Figure 26), could therefore serve as important tools to either extract polymeric hemicelluloses, or facilitate hydrolysis by other lignocellulosic enzymes. In order to evaluate the biotechnological potential of GEs, it is necessary to understand their function on the molecular level, as well as how they work in the plant cell wall. Research on this relatively novel type of enzymes also requires fundamental insight on sequence diversity, structure-function relationships and substrate specificity. Investigating the diversity and function of GEs, to spread light on their biological role, was the major part of this work (**Papers III, IV, V**), as discussed in this chapter.

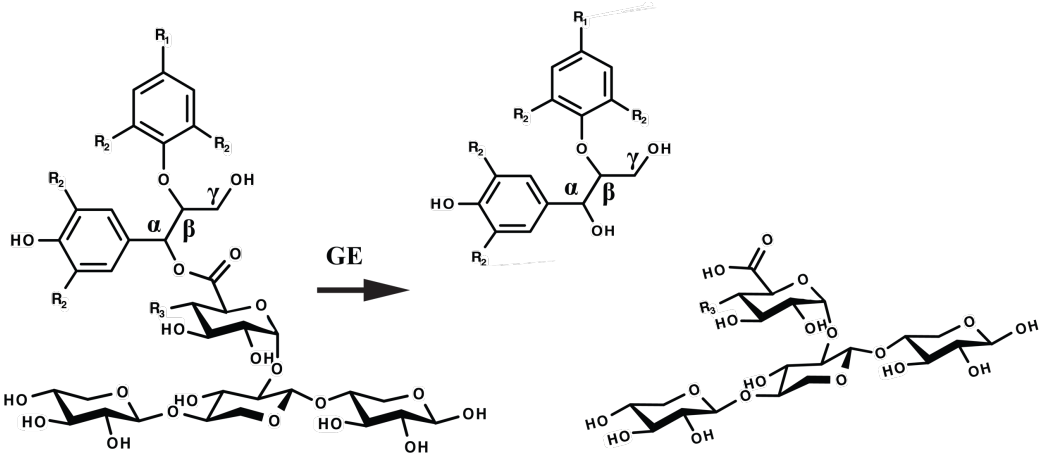


Figure 26. GE-mediated LC ester cleavage (represented by a benzyl ester) resulting in the separation of lignin from glucuronoxylan.

7.1 Discovery of GEs

The first glucuronoyl esterase was discovered during a stereochemical study of a GH67 α-glucuronidase from *Aspergillus tubingensis.* The α-glucuronidase released 4-*O*-MeGlcA from fragments of beechwood-derived glucuronoxylan, but was unable to act on corresponding methyl-esterified glucuronoxylan. This demonstrated the necessity of a free

carboxylic group on GlcA for the activity of the α-glucuronidase (Biely et al., 2000). Liberation of 4-*O*-MeGlcA from the ester-linked substrate was, however, observed after incubation with a crude extract from the wood-rotting fungus *Schizophyllum commune*. This suggested that *S. commune* produces an enzyme capable of hydrolysing methyl esters of glucuronic acids. This hypothesis was proven true, and implied the discovery of the first GE, *Sc*GE1 (Špániková & Biely, 2006). *Sc*GE1 was purified and found to be active on a range of synthetic substrates of ester-linked glucuronic acids, thus the name glucuronoyl esterase.

Shortly after the discovery of *Sc*GE1, the homologous sequence was found in the *cip2* gene of *Trichoderma reesei*. *Cip2* had been described as a gene encoding for a protein with unknown function, but was overexpressed in *T. reesei* under the induction of cellulolytic and hemicellulolytic enzymes (Foreman et al., 2003). The gene also contained a sequence coding for a cellulose-binding domain (CBM1), which strengthened the theory of its involvement in the degradation of plant biomass. The expression and characterisation of the catalytic Cip2 domain indeed confirmed it to be a GE (*Tr*GE) (Li et al., 2007). *Tr*GE was the second GE characterised and the first to be structurally solved (Pokkuluri et al., 2011). Homologous sequences have subsequently been found in both fungal and bacterial genomes, which have led to the creation of the carbohydrate esterase family 15 (CE15) in CAZy.

To date, 12 years after their discovery, 28 GEs of bacterial and fungal origin have been reported and characterised (Table 5). The first years of GE research were devoted solely to fungal enzymes, whereas bacterial GEs have only recently started to be explored. This work contributed by the characterisation of 11 bacterial GEs, i.e., more than a third of all characterised GEs and the vast majority of all reported bacterial GEs (**Papers IV, V**).

Table 5. All characterised GEs, starting with the first reported and listed in chronological order

GE	Organism	Origin	Characterisation*	References
***Sc*GE**	*Schizophyllum commune*	Fungal[B]	Model compounds Polymeric substrate	Špániková & Biely, 2006 Špániková et al., 2007 Wong et al., 2012 Ďuranová et al., 2009 Biely et al., 2015 D'Errico et al., 2015 D'Errico et al., 2016
***Tr*GE (Cip2)**	*Trichoderma reesei (Hypocrea jecorina)*	Fungal[A]	Model compounds Structure (PDB: 3pic) Polymeric substrate Activity on biomass	Li et al., 2007 Ďuranová et al., 2009 Pokkuluri et al., 2011 Biely et al., 2015 D'Errico et al., 2016
***Pc*GE1**	*Phanerochaete chrysosporium*	Fungal[B]	Model compounds	Ďuranová et al., 2009 Hüttner et al., 2017
***Pc*GE2**	*Phanerochaete chrysosporium*	Fungal[B]	Model compounds	Ďuranová et al., 2009
***St*GE1**	*Sporotrichum thermophile (Myceliophthora thermophila)*	Fungal[A]	Model compounds	Vafiadi et al., 2009
***St*GE2**	*Sporotrichum thermophile (Myceliophthora thermophila)*	Fungal[A]	Model compounds Structure (PDB: 4g4g, 4g4i, 4g4j)	Topakas et al., 2010 Charavgi et al., 2013 Katsimpouras et al., 2014 Sunner et al., 2015 Nylander et al., 2016

***Pca*GE**	*Phanerochaete carnosa*	Fungal[B]	Overexpressed in plants	Tsai et al., 2012 Latha Gandla et al., 2014
***Pa*GE1**	*Podospora anserina*	Fungal[A]	Model compounds	Katsimpouras et al., 2014 Sunner et al., 2015
CesA	*Ruminococcus flavefaciens*	Bacterial	Polymeric substrate	Biely et al., 2015
***Cu*GE**	*Cerrena unicolor*	Fungal[B]	Model compounds Activity on biomass	D'Errico et al., 2015 D'Errico et al., 2016 Mosbech et al., 2018
***Aa*GE1**	*Acremonium alcalophilum*	Fungal[A]	Model compounds Activity on LCCs	**Paper III** (2016) Hüttner et al., 2017
***Nc*GE**	*Neurospora crassa*	Fungal[A]	Model compounds	Huynh & Arioka, 2016 Huynh et al., 2018
MZ0003	Uncultured marine bacterium	Bacterial	Model compounds Structure (PDB: 6ehn)	De Santi et al., 2016 De Santi et al., 2017
***Wc*GE1**	*Wolfiporia cocos*	Fungal[A]	Model compounds	Hüttner et al., 2017
***Af*GE**	*Aspergillus fumigatus*	Fungal[A]	Model compounds	Huynh et al., 2018
***Cs*GE**	*Ceriporiopsis subvermispora*	Fungal[B]	Model compounds	Lin et al., 2018
***Pe*GE**	*Pleurotus eryngii*	Fungal[B]	Model compounds	Lin et al., 2018
***Ot*CE15A**	*Opitutus terrae*	Bacterial	Model compounds Structure (PDB: 6grw, 6gs0)	**Paper IV** (2018)
***Ot*CE15B**	*Opitutus terrae*	Bacterial	Model compounds	**Paper IV** (2018)
***Ot*CE15C**	*Opitutus terrae*	Bacterial	Model compounds	**Paper IV** (2018)
***Ot*CE15D**	*Opitutus terrae*	Bacterial	Model compounds	**Paper IV** (2018)
***Sl*CE15A**	*Spirosoma linguale*	Bacterial	Model compounds Activity on biomass	**Paper IV** (2018)
***Sl*CE15B**	*Spirosoma linguale*	Bacterial	Model compounds	**Paper IV** (2018)
***Sl*CE15C**	*Spirosoma linguale*	Bacterial	Model compounds	**Paper IV** (2018)
***Su*CE15A**	*Solibacter usitatus*	Bacterial	Model compounds Activity on biomass	**Paper IV** (2018)
***Su*CE15B**	*Solibacter usitatus*	Bacterial	Model compounds	**Paper IV** (2018)
***Su*CE15C**	*Solibacter usitatus*	Bacterial	Model compounds Structure (PDB: 6gu8, 6gry) Activity on biomass	**Paper IV** (2018)
***Tt*CE15A**	*Teredinibacter turnerae*	Bacterial	Model compounds Structure (PDB: 6hsw)	**Paper V** (2019)

*Including: biochemical characterisation/activity on model compounds, polymeric substrate, LCCs or biomass, overexpression in plants and structural determination.
[A] Ascomycete
[B] Basidiomycete

7.2 Phylogeny

All GEs characterised to date belong to the CE family 15 (CE15) in CAZy. CE15 enzymes conform to the α/β-hydrolase fold and often share conserved catalytic residues in the active site. They have distinctly different sequences and folds compared to enzymes in the other CE families, with acetyl, feruloyl and pectin methyl esterase activities, and so far, all characterised members have shown GE activity (Ďuranová, Špániková, et al., 2009). The family consists of bacterial and fungal members, with the exception of a CE15 sequence from the archaeon *Halorhabdus utahensis* (Anderson et al., 2009). Agger et al. (2017) recently proposed that plants have CE15-like proteins, due to the similar α/β-hydrolase fold, but further investigations will be needed to determine whether they are actual esterases or proteases (Agger et al., 2017).

The phylogeny of CE15 has been reported in a few articles. Several of these studies have shown that CE15 enzymes are widespread, but not necessarily present in fungal genomes, which indicates that they are not crucial for all fungi (De Santi et al., 2016; Dilokpimol et al., 2018; Ďuranová, Hirsch, et al., 2009; Hüttner, Klaubauf, et al., 2017; Lin et al., 2018; Monrad et al., 2018). In a recent phylogenetic study by Monrad et al. (2018), fungal CE15 sequences clustered into two clades. A larger and more diverse clade, with exclusively bacterial targets (and the archaeal enzyme) was also identified, as well as a few small clades containing a mixture of bacterial and fungal sequences. The close relation between bacterial and fungal sequences in the mixed clades has been suggested to be the result of horizontal gene transfer (Monrad et al., 2018).

The sequence diversity of CE15 was further investigated in the present work (**Paper IV**). All CE15 sequences available in CAZy at that time (July 2018: 239 entries, in Jan. 2019: 265 entries) were aligned, and a phylogenetic tree was constructed. This phylogenetic tree displayed two clear fungal clades, in agreement with previous reports (Figure 27). The fungal GEs characterised were present in the major clade, distant in sequence from many of the bacterial targets. The bacterial CE15 sequences, which constitute the majority of all CE15 entries in CAZy, grouped into several distinct clades across the tree, with as low as 15 % sequence identity. Two previously reported bacterial GEs (Biely et al., 2015; De Santi et al., 2016) were found in the middle section of the tree, but the major part was unmapped. The large sequence divergence between the bacterial CE15 members and their fungal counterparts raised questions on potential broader substrate specificities and novel activities among the bacterial targets.

The likeliness of functional diversity within CE15 applies not only to enzymes from different sources. Several microorganisms encode multiple CE15 enzymes (1-4 genes in a single genome; (Lombard et al., 2014)), which indicates differences in the biological function of CE15 enzymes within these organisms. The fact that multiple and non-redundant genes from a range of CAZy families are present in the genomes of microorganisms, could also be related to the highly complex and diverse structure of lignocellulose. The 11 bacterial CE15 enzymes described in this thesis originate from four bacteria with multiple CE15 genes: *Opitutus terrae*, *Spirosoma linguale*, *Solibacter usitatus (***Paper IV***)* and *Teredinibacter turnerae* (**Paper V**). The enzymes are spread across the phylogenetic tree, and the sequence identities were as low as 25 % (Figure 27).

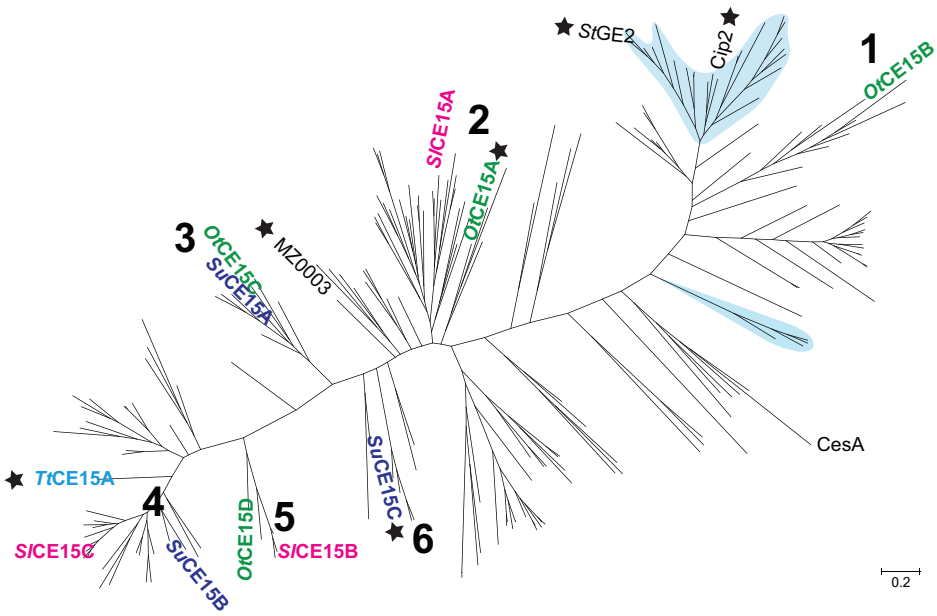


Figure 27. Phylogenetic tree including the CE15 catalytic domains found in CAZy. Clades shaded in blue represent fungal CE15 members. Stars indicate structurally determined enzymes. The bacterial enzymes characterised (**Papers IV, V**) are labelled with their protein names and colour coded: green (*O.terrae*), magenta (*S. linguale*), blue (*S. usitatus*) and turquoise (*T. turnerae*). These CE15 members were further grouped into several clades, labelled 1-6. The two previously characterised bacterial GEs (MZ0003 and CesA) are found in the middle of the tree. All characterised fungal GEs are located in the larger fungal clade (including the structurally determined *St*GE2 and Cip2). (Figure courtesy of Johan Larsbrink.)

7.3 Chemical properties and modularity

The GEs characterised to date have shown pH optima in the acidic to slightly alkali pH range (pH 5-8.5), and temperature optima between 35 and 60 °C. The typical molecular weight of the catalytic domain of GEs is 40-50 kDa (Monrad et al., 2018). Several CE15 enzymes are multi-modular, linked to CBMs or to enzymes with other functionalities. All CBM-containing GEs characterised to date possess CBM1 (probably cellulose binding) at the N-terminus. CE15 sequences that are connected to CE3 or CE6 acetyl xylan esterases and various xylanases have been reported for a few bacterial species, such as the shipworm gut bacterium *T. turnerae* (Yang et al., 2009) and the anaerobe *Ruminococcus flavefaciens 17* (Aurilia et al., 2000). CE15 sequences from *Bacteroides* species can be found as part of predicted polysaccharide utilisation loci (PULs) in the PUL data base (PULDB) (Terrapon et al., 2018). Some of these PULs contain xylanases, but more interestingly, activities for the degradation of rhamnogalacturonan are common, which indicates that CE15 enzymes might be involved in the hydrolysis of ester linkages present in pectin.

7.4 Assaying GE activity

GE research has been hampered by the lack of relevant substrates and appropriate analytical methods. However, some chromatographic and spectrophotometric assays are suitable for the detection of GE activity on different ester-linked substrates. TLC has been used to study GE activity qualitatively, while high-performance liquid chromatography (HPLC) has been used for quantitative activity assessments on simple alkyl and aryl alkyl esters of GlcA (Monrad et al., 2018). Four spectrophotometric enzyme-coupled assays have been described, which have facilitated high-throughput screening of GEs ((Fraňová et al., 2016; Sunner et

al., 2015), **Paper IV**). The robust, coupled assay established by Sunner et al. (2015) was improved and developed for operation in continuous mode in the present work (**Paper IV**, described in Chapter 5). The continuous monitoring of reactions allowed for rapid screening and collection of kinetic data. The assay has the advantage of being compatible with commercially available esters of GlcA, but with the limitation that the coupling enzyme (UDH) is only able to detect the β-anomer of GE-generated GlcA. The two assays described by Fraňová et al. (2016) allow for the detection of both α- and β-anomers, but require substrate synthesis and in neither case has operation in continuous mode been reported.

Assays for the detection of GE activity on complex and native ester substrates are less well developed, but various NMR techniques have been regarded as promising (as described in Chapter 5). 1D NMR has been used to monitor the disappearance of methyl ester protons from esterified glucuronoxylan upon GE treatment (Biely et al., 2015). In the present work, ^{31}P NMR complemented with SEC, was used to demonstrate GE-catalysed cleavage of LC esters extracted from spruce and birch (**Paper III**). HSQC was also used to detect GE hydrolysis of benzyl esters in LCCs isolated from Japanese beech (unpublished data, described in Section 7.8).

Commercial and synthetic model substrates

Model substrates relevant for assaying GE activity are based on GlcA compounds esterified with alcohols of different complexity. Benzyl-, allyl-, and methyl-substituted GlcA esters (BnzGlcA, AllylGlcA and MeGlcA; Figure 28A-C) are commercially available, and have been used as model substrates in several studies, including the present work (**Papers IV, V**). BnzGlcA is regarded as the most relevant among the commercially available substrates, due to its aromatic alcohol moiety, representing the lignin portion of LC esters. The commercial substrates are simple α/β-anomer mixtures of GlcA (α-linked) and they lack the 4-*O*-methylation present in native LC esters. The lack of substituent at the anomeric position makes it possible for these substrates to undergo isomerisation, which has complicated TLC and HPLC analysis by appearing as multiple spots/peaks. GE activity with BnzGlcA, AllylGlcA, MeGlcA and also MeGalA (Figure 28D) was, however, successfully demonstrated with the continuous coupled spectrophotometric assay (**Papers IV**, **V**).

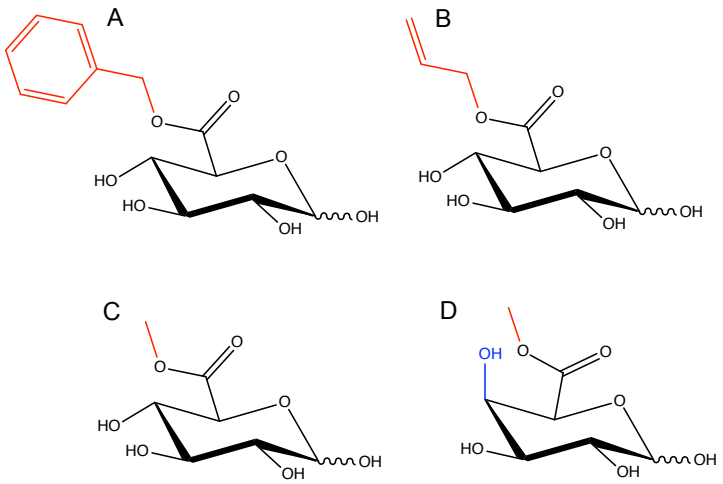


Figure 28. Commercially available model substrates used for assaying GE activity: BnzGlcA (**A**), AllylGlcA (**B**) and MeGlcA (**C**). MeGalA (**D**) was included as a substrate to investigate the substrate promiscuity among the putative GEs. The configuration of the C4 hydroxyl moiety is oriented differently in MeGalA (indicated in blue). Red indicates the alcohol portion of the substrates.

Chemical synthesis allows the use of more appropriate GE substrates, but often with the drawback of several laborious synthesis steps. Chemically synthesised substrates have been relevant in GE research through the addition of more complex aromatic alcohol moieties (Nylander et al., 2016), the presence of 4-*O*-methylation (Špániková et al., 2007) and by γ-linked esters (D'Errico et al., 2016; D'Errico et al., 2015) (Figure 29). However, problems associated with substrate instability and limited solubility have been reported, and these substrates have not been widely applied.

Plant biomass-derived substrates

The most relevant substrates for GE research are native LC esters present in plant biomass. However, the estimated low abundance of LC ester bonds prevents their detection, and enrichment through LCC isolation is required. Chemical degradation of the labile LC esters has been reported to be problematic in many solvent-based extraction procedures (Balakshin et al., 2014), but in later years milder fractionation protocols have been developed for various types of plant material, including woody biomass (Giummarella et al., 2016). For enzymatic reactions, it is important that the isolated LCC material is soluble in salt buffers as well as in the solvent used in the subsequent analysis. In the present work (**Paper III**), water-soluble LCCs from spruce and birch were isolated using the mild and universal protocol developed by Giummarella et al. (2016). Fractionation was performed at near-neutral pH and low temperature (80 °C) with the aim of preserving the native LCC structure. The chemical composition of these LCC fractions is presented in Table 6.

Table 6. Mass balance and chemical composition of spruce and birch LCCs used as GE substrates (**Paper III**)

LCC fraction	**Mass balance**	**Lignin content** (% of fraction)		**Sugar content** (%)*					**4-*O*-MeGlcA**
	(% of wood)	Dioxane 96 % soluble	Total**	Ara	Gal	Glc	Xyl	Man	(% on xylan)
Spruce	10	1	15	3.4	5.5	16.5	17.3	57.3	3.5
Birch	10	4	13	1.7	3.1	12.3	75.3	7.6	1

*According to HPAEC-PAD
**Obtained after acetobromination

Obviously, an isolated LCC fraction is only part of the whole LCC present in native lignocellulose. Intact LCCs would therefore be the most natural substrate to use. However, no methods have been reported that are sensitive enough for the detection of GE action on non-isolated LCCs. Ball-milled plant biomass has been used for indirect measurements of GE activity in boosting assays, where *endo*-xylanases or lignocellulolytic enzyme cocktails were supplemented with GEs, and increased saccharification was observed (**Paper IV**; (D'Errico et al., 2016; Mosbech et al., 2018)).

7.5 Biochemical characterisation and substrate specificity

GE activity has been reported on alkyl and aryl alkyl esters of D-glucuronic acid in a handful of studies. The following factors (structurally significant in native LCCs) have been discussed in relation to substrate specificity: the length of the anomeric carbohydrate substitution on the GlcA, the alcohol portion of the ester, 4-*O*-methylation of the GlcA and

the type of ester linkage (α [benzyl] or γ) (Figure 29) (D'Errico et al., 2016; D'Errico et al., 2015; Huynh et al., 2018; Monrad et al., 2018).

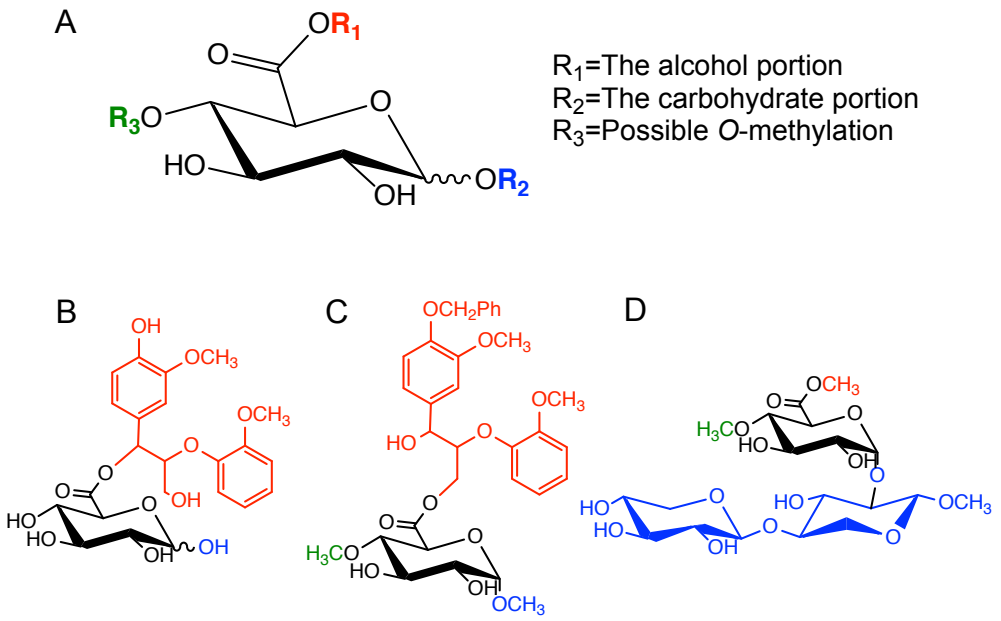


Figure 29. Esters of GlcA used as GE substrates. Ester of GlcA with highlighted key structures discussed in relation to substrate specificity, with R_1 (red): the alcohol ("lignin") portion (either α- or γ-linked), R_2 (blue): the anomeric carbohydrate substitution ("xylan") and R_3 (green): the position of 4-*O*-methylation (**A**). Examples of synthesised esters of GlcA that have been used as GE model substrates, demonstrating different types of alcohol portions, carbohydrate portions and ester linkages, as well as 4-*O*-methylation (**B-D**) (D'Errico et al., 2015, Nylander et al., 2016, Špániková et al., 2007).

Fungal GEs

In the early years of GE research, activity was assayed on a variety of simple methyl esters of GlcA (Ďuranová, Hirsch, et al., 2009; Li et al., 2007; Špániková & Biely, 2006; Špániková et al., 2007; Topakas et al., 2010; Vafiadi et al., 2009; Wong et al., 2012). Based on the results of these studies, it was suggested that GlcA recognition was critical for GE activity, demonstrated by the lack of activity on GalA esters (Ďuranová, Hirsch, et al., 2009; Wong et al., 2012).

The importance of the character of the alcohol portion has been studied by assaying fungal GEs on GlcA substrates esterified with bulky aromatic alcohols (to mimic the nature of native LC esters). Increased affinity of GEs to substrates with larger alcohol portions has been reported in several cases (D'Errico et al., 2015; Hüttner, Klaubauf, et al., 2017; Nylander et al., 2016). 4-*O*-methylation has been reported as another important structural feature for increased substrate specificity (D'Errico et al., 2015; Ďuranová, Hirsch, et al., 2009; Huynh et al., 2018; Špániková et al., 2007), while the length of the xylan portion linked to the anomeric position of the GlcA seems to be less important for fungal GE activity (Biely et al., 2015; Špániková et al., 2007). Only limited kinetic information is available when γ-linked esters are used as GE substrates, and it has still not been confirmed whether the type of ester linkage is important for GE activity (D'Errico et al., 2015; Huynh & Arioka, 2016; Huynh et al., 2018). One case of feruloyl esterase side activity has been described for *Cu*GE from *Cerrena unicolor* (Mosbech et al., 2018), but fungal GEs with other side activities have not been reported.

The reported kinetic parameters for fungal GEs on a range of model substrate vary: K_M=0.25-80 mM and k_{cat}=0.2-285 s^{-1}; the highest catalytic efficiencies being observed on 4-*O*-methylated esters of GlcA (Monrad et al., 2018). The trends in substrate specificity should, however, be interpreted with caution, due to the lack of studies using directly comparable model substrates and experimental setups. A handful of GEs have been assayed on the substrate BnzGlcA, allowing for kinetic comparison between candidates. The kinetic parameters for fungal GEs on BnzGlcA are summarised in Table 7, together with values obtained for bacterial GEs in this work (**Papers IV, V**). The characterised fungal GEs in general show low values of k_{cat}/K_M on BnzGlcA, which might indicate that these enzymes are more efficient on larger LCC fractions found in nature, but is also likely to be an effect of the lack of 4-*O*-methylation in BnzGlcA.

Table 7. Kinetic parameters for selected bacterial and fungal GEs on BnzGlcA, listed in order of increasing catalytic efficiency (k_{cat}/K_M). The bacterial GEs (indicated by an asterisk) in general have higher k_{cat}/K_M values than their fungal counterparts, and many of them also have exceptionally low K_M values.

Enzyme	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}mM^{-1}$)	Reference
*SlCE15C**	7.6	0.74	0.097	Paper IV
*Pc*GE1	2.9	0.37	0.13	Hüttner et al., 2017
*Nc*GE	35	8.4	0.24	Huynh et al., 2018
*Aa*GE1	1.7	0.79	0.46	Hüttner et al., 2017
*Pa*GE1	12	7.8	0.64	Sunner et al., 2015
*Af*GE	16	24	1.5	Huynh et al., 2018
*Sl*CE15A*	6.0	11	1.9	Paper IV
*Sl*CE15B*	0.62	1.6	2.6	Paper IV
*Ot*CE15A*	4.2	19	4.6	Paper IV
*Ot*CE15D*	0.61	6.2	11	Paper IV
*Ot*CE15C*	0.45	5.2	12	Paper IV
*Su*CE15A*	0.42	9.3	22	Paper IV
*Su*CE15C*	2.6	58	23	Paper IV
*Tt*CE15A*	3.5	130	38	Paper V

*Bacterial GE

Bacterial GEs

Two bacterial CE15 enzymes had been studied previously to this work, CesA from *R. flavefaciens* and MZ0003, isolated from a marine Arctic metagenome (Biely et al., 2015; De Santi et al., 2016). Both of them have been characterised as GEs, but with reported acetyl esterase activity, not observed for any fungal GEs. The broader substrate specificity of CesA was suggested to be an effect of its multimodular nature, with a C-terminal domain with GE activity (Biely et al., 2015) and an N-terminal domain with acetyl xylan esterase activity (Aurilia et al., 2000). The single domain MZ0003, has been shown to be more active on *p*NP-acetate than on esters of GlcA. However, the GE activity was only assessed qualitatively and more detailed kinetic data would be required to draw any conclusions on the substrate specificity of MZ0003 (De Santi et al., 2016). In the work described in this thesis, eleven novel bacterial GEs were biochemically characterised (**Papers IV, V**), which contributed significantly to our knowledge on GEs from this kingdom and to the diversity of GEs across CE15.

Novel and diverse substrate specificities of GEs across CE15

So far, GE activity is the only enzyme activity that has been found in CE15. However, this has been based on biased sampling of closely related CE15 members of almost exclusively fungal origin. The purpose of the investigations reported in **Papers IV** and **V**, was to gain knowledge on CE15 by characterising bacterial enzymes distributed throughout the family. The studies were initiated with the assumption of discovering broad substrate specificities

and possible new activities among the candidates. Eleven CE15 enzymes encoded by four bacteria (*Ot*CE15A-D, *Sl*CE15A-C, *Su*CE15A-C (**Paper IV**) and *Tt*CE15A (**Paper V**)) were selected for cloning and expression, and were assayed on the substrates BnzGlcA, AllylGlcA, MeGlcA and MeGalA. Assays were also performed with *p*NP-Ac, to test for acetyl esterase activity reported for previously characterised bacterial GEs.

The eleven enzymes studied were all confirmed to be GEs, and several of them exhibited extremely high catalytic efficiencies on the model substrates. A few of them demonstrated strict substrate preferences, whereas others were found to be more promiscuous. However, only trace amounts of acetyl esterase activity were observed, confirming their specificity towards substrates of uronic acids. The kinetic parameters for the GEs encoded by *O. terrae*, *S. linguale* and *S. usitatus* are presented in Table 1 and Table S2, in **Paper IV**, illustrating the diversity in substrate affinity and turnover rate of GEs within one genome.

It was clearly shown that GEs belonging to similar/closely related clades in the phylogenetic tree (Figure 27) shared substrate preferences (e.g. a preference for benzyl substituents, able to hydrolyse MeGalA; Table 8). Five of the enzymes (*Ot*CE15C & D, *Su*CE15A & C and *Tt*CE15A) showed exceptionally high catalytic efficiencies on BnzGlcA (10-40 $s^{-1}mM^{-1}$), i.e., 10-100 times higher than for fungal GEs assayed on the same model substrate (Tables 7 & 8). Fungal GEs with k_{cat}/K_M values in this range have only been reported for 4-*O*-methylated substrates (D'Errico et al., 2016; D'Errico et al., 2015; Ďuranová, Hirsch, et al., 2009; Wong et al., 2012). A large proportion of the bacterial enzymes were also highly active on allyl- and methyl-substituted GlcA substrates (especially *Su*CE15C), and did not discriminate between the type of ester substitution (benzyl, allyl or methyl) (Table 8).

The benzyl moiety has been described to be crucial for fungal GEs, but a broad range of K_M values has been reported (~2-80 mM). The bacterial GEs exhibited much lower and a more narrow distribution of K_M values on BnzGlcA (down to 0.4-0.6 mM). These low K_M values suggest that 4-*O*-methylation might be less crucial for the substrate specificity of these GEs. This observation is interesting in relation to the glucuronoxylan composition in different plant sources. In xylan from poplar (*Populus)*, all the GlcA substituents are 4-*O*-methylated, in contrast to *Arabidopsis* xylan, where only 60-70 % of the GlcA substituents are 4-*O*-methylated (Urbanowicz et al., 2012; Yuan et al., 2014). GEs with different preferences for 4-*O*-methylation might therefore be of biological relevance for the hydrolysis of LC esters from different plant sources with different characteristics.

Some of the investigated bacterial GEs were only active on esters of GlcA, as expected, but several of them were highly efficient on MeGalA (Table 8). This substrate promiscuity, not reported for fungal GEs, indicates that the configuration of the C4 hydroxyl moiety is of minor importance for some GEs. Significant diversity was observed among the bacterial GEs in terms of substrate specificity, but characterisation on complex and complementary model substrates is required to give a more complete picture of their individual substrate preferences and possible functional differences.

Table 8. Catalytic efficiencies of characterised bacterial GEs on esters of GlcA and GalA (**Papers IV, V**). The shading indicates the level of efficiency, from dark green (highest) to dark orange (no activity). GEs from clades 3 and 5 are indicated by an asterisk, and exhibited exceptionally low K_M values on BnzGlcA (< 1mM).

Enzyme	Clade	BnzGlcA	AllylGlcA	MeGlcA	MeGalA
*Ot*CE15B	1	+	+	+	+
*Ot*CE15A	2	+++	+++	+++	+++
*Sl*CE15A	2	+++	+++	+++	+
*Ot*CE15C*	3	++++	+++	++	+++
*Su*CE15A*	3	++++	+++	+++	+++
*Sl*CE15C	4	++	++	++	-
*Su*CE15B	4	+++	++	(-)	(-)
*Tt*CE15A	4	++++	+++	+++	+
*Ot*CE15D*	5	++++	+++	++	-
*Sl*CE15B*	5	+++	+++	++	-
*Su*CE15C	6	++++	++++	++++	+++

7.6 Structural determination

To date, two fungal GEs and four bacterial GEs have been structurally solved using X-ray crystallography. The fungal GEs are: *Tr*GE (Cip2) from *T. reesei* (*Hypocrea jecorina*) (PDB: 3pic) (Pokkuluri et al., 2011) and *St*GE2 from *Thermothelomyces thermophila* (previously *Sporotrichum thermophile*) (PDB: 4g4g, 4g4i and 4g4j) (Charavgi et al., 2013), and the bacterial GEs are MZ0003 (PDB: 6ehn; cloned from a marine metagenome) (De Santi et al., 2017), *Ot*CE15A (PDB: 6grw and 6gs0), *Su*CE15C (PDB: 6gry and 6gu8) and *Tt*CE15A (PDB: 6hsw). The last three bacterial enzymes were crystallised and their structures determined as part of a fruitful collaboration with Dr Scott Mazurkewich and Professor Leila Lo Leggio (**Papers IV, V**), thus doubling the number of solved GE structures.

Overall structures

All structurally determined GEs share the α/β-hydrolase fold, consisting of a three-layer αβα sandwich with the active site, comprising the catalytic triad, present in a solvent-exposed cleft. The solved bacterial structures differ significantly from the two fungal ones, namely, in the insertion of 3-4 regions (Reg1-3 and RegN). Reg1-3 are present in all bacterial structures (**Papers IV, V**; (De Santi et al., 2017)), whereas RegN was found only in the structure of *Tt*CE15A (**Paper V**) (Figure 30A-B).

The flat surface exposed catalytic site of the fungal GEs (*Tr*GE and *St*GE2), suggests the ability of these enzymes to act on large and/or relatively intact LCC polymers (Figure 30C). The substrate pocket of the bacterial GEs is deeper, due to the presence of Reg2, which might indicate that smaller fragmented LC esters are the preferred substrates. Reg1 and Reg3 have analogous shapes and high sequence similarity in all the bacterial GEs structurally determined so far, but Reg2 contains a much longer helix in *Tt*CE15A and forms an even deeper active site pocket. The side of Reg2 facing the pocket contains several hydrophobic residues, potentially involved in the interaction with lignin. The other side, facing the solvent, has a more hydrophilic character. RegN, present in *Tt*CE15A, is located opposite Reg2 and contributes to additional deepening of the pocket (Figure 30A-B). RegN consists of several hydrophilic residues with the suggested potential to participate in the interaction with carbohydrate structures.

The two fungal structures exhibit three disulphide bridges close to the catalytic site, which contribute to the rigidity of the architecture. No disulphide bridges were found in the bacterial structures, which has been interpreted as the explanation of their flexibility and broader substrate specificity (De Santi et al., 2017). However, multiple sequence alignments indicated that disulphide bridges were present in several bacterial CE15 enzymes, e.g. in *Tt*CE15B and *Tt*CE15C (encoded by the same species as *Tt*CE15A) (**Paper V**), suggesting that some bacterial CE15 enzymes might be more similar to their fungal counterparts in terms of structure and also substrate specificity. Future structural GE determinations may elucidate these structural elements, providing a better understanding of the differences and similarities between fungal and bacterial GEs.

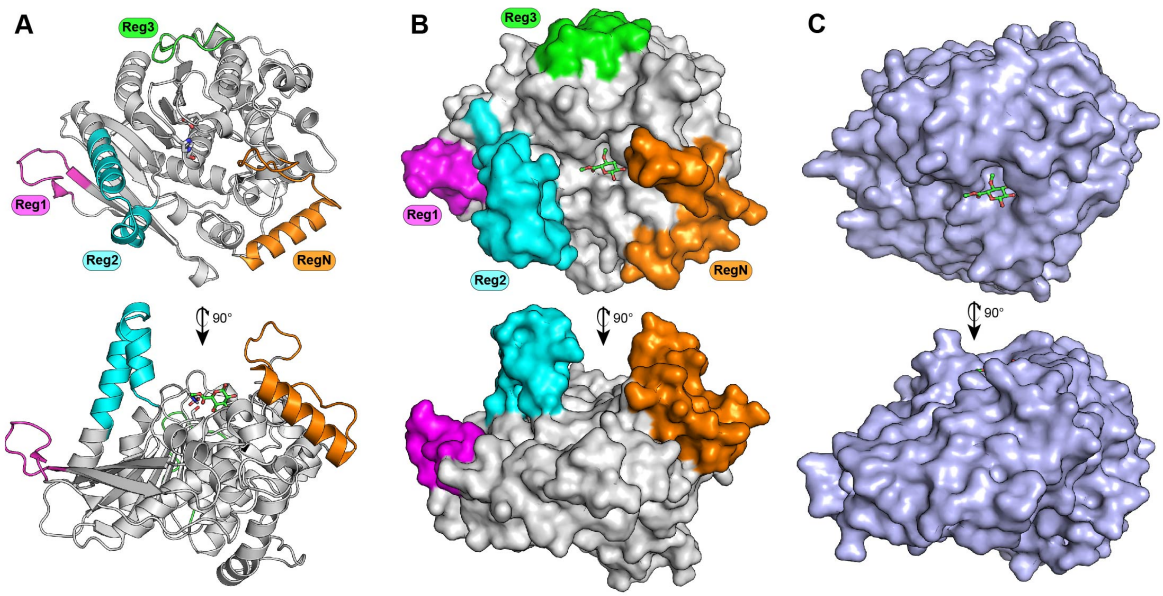


Figure 30. Overall structure (**A**) and space-filling representation (**B**) of *Tt*CE15A (**Paper V**). Inserted regions relative to the fungal GEs (N, 1, 2 and 3) are coloured in orange, magenta, cyan and green. The space-filling representation of *St*GE2 (**C**) displays the flat surface exposed catalytic site compared to *Tt*CE15A. The 4-*O*-methyl glucuronoate substrate, shown as green sticks, was generated from structural alignment with the co-crystallised structure of *St*GE2 (PDB: 4g4j). The overall structures of *Ot*CE15A and *Su*CE15C are similar to that of *Tt*CE15B, but lack region N (Figure 2 in **Paper IV**). (Figures courtesy of Scott Mazurkewich.)

Active sites

All the solved structures confirm that GEs are serine-type hydrolases, with the conserved nucleophilic serine in a novel consensus sequence: G-X-S-R-X-G-K, compared to the classical G-X-S-X-G serine motif commonly found in carboxyl esterases (Bornscheuer, 2002). The basic active site residue histidine is also highly conserved in all GE structures. Based on studies of fungal GEs, it has previously been believed that a glutamate (in this work called the canonical glutamate) was conserved, probably acting as the acidic regulator of the basic character of the histidine (Figure 31). However, the position of the glutamate in MZ0003 is occupied by a cysteine, and the acidic residue of the triad has been confirmed by mutagenesis to be an aspartate (a position occupied by isoleucine in both fungal GEs) (De Santi et al., 2017).

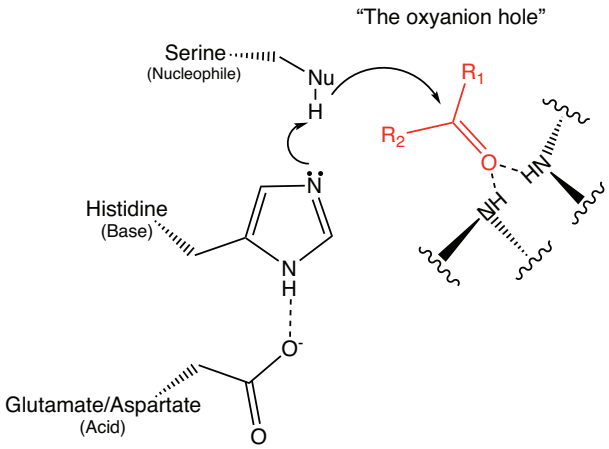


Figure 31. Schematic illustration of the catalytic triad in serine-type hydrolases. In GEs, the acid residue is represented by aspartate or glutamate, the base by histidine, and the nucleophilic attack is performed by a serine. The so-called "oxyanion hole" of the ester was found to be stabilised by an arginine from the ligand complex of *St*GE2.

The active sites of the bacterial structures solved within this work (**Papers IV, V**) have some similarities, but also differ from each other, and from previously solved structures (Figure 32). The canonical glutamate was observed present in both *Ot*CE15A and *Su*CE15C (**Paper IV)**, but not in *Tt*CE15A. Instead, a glutamate was found in the same position as the aspartate in MZ0003, confirmed by mutagenesis to be catalytically involved (**Paper V**). Interestingly, it was observed that, in addition to their canonical glutamate residues, *Ot*CE15A and *Su*CE15C contained aspartate residues in the equivalent position to the *Tt*CE15A glutamate. However, it has not yet been determined whether the aspartate or the canonical glutamate of these GEs is involved in the catalytic triad.

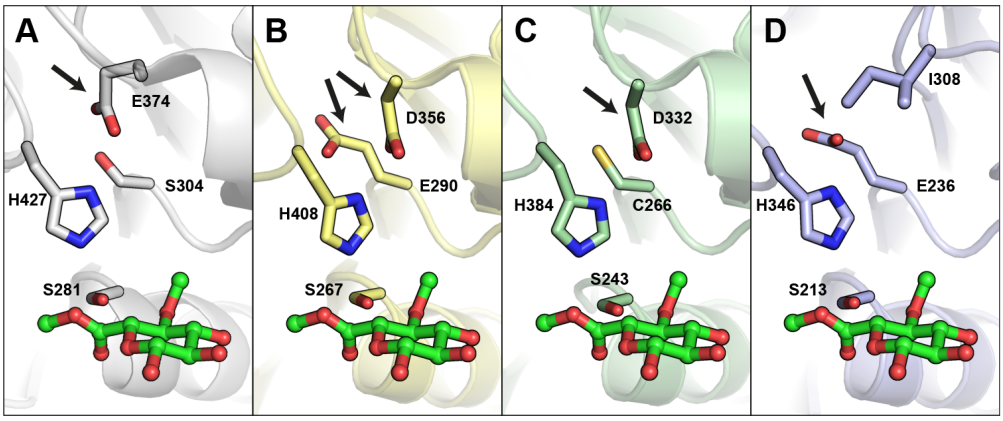


Figure 32. Catalytic residues in selected structurally determined GEs: *Tt*CE15A (**Paper V**) (**A**), *Ot*CE15A (**Paper IV**) (**B**), MZ0003 (**C**) and *St*GE2 (**D**). A methyl ester of 4-*O*-methyl glucuronoate is generated from structural alignment with the ligand complex of *St*GE2 (PDB: 4g4j). The acidic residues are indicated by black arrows. (Figures courtesy of Scott Mazurkewich.)

The only structure of a CE15 enzyme that comprises a ligand in its active site is represented by the S213A variant of *St*GE2 in a complex with a methyl ester of 4-*O*-methyl glucuronoate (PDB: 4g4j) (Figure 33A) (Charavgi et al., 2013). Several hydrogen bonds, from the substrate analogue to the active-site residues, were found in the complex, and the catalytic triad is positioned in an orientation described as "*ready for nucleophilic attack*". The 4-*O*-methoxy group exhibits several van der Waals interactions with residues in the active site in *St*GE2, suggesting the importance of 4-*O*-methylation in substrate recognition.

There are variations in the active site residues between all GEs, but it is interesting to bring up important residues in the *St*GE2 ligand complex that are conserved across the solved structures (Figure 33A-B). These include an arginine positioned next to the catalytic serine in the consensus sequence, suggested to stabilise the so-called oxyanion hole during hydrolysis (Figure 31), and a lysine that is hydrogen bonded with the 3-OH and the 4-*O*-methoxy oxygen. In addition to the lysine, a closely positioned hydrophobic patch has been suggested to be important for the correct positioning of the 4-*O*-methoxy moiety. These hydrophobic residues are conserved in all the fungal GEs characterised to date and were also observed in *Tt*CE15A (**Paper V**). The equivalent space in *Ot*CE15A and *Su*CE15C was occupied by smaller residues and the non-canonical aspartate. It was suggested in **Paper IV** that the acidic aspartate could facilitate binding to non-methylated substrates, which strengthens the hypothesis that 4-*O*-methylation may not be relevant for these GEs.

The structures of *Ot*CE15A and *Su*CE15C revealed a phenylalanine (Phe141 and Phe135, respectively) within the pocket formed by Reg2 (Figure 2 in **Paper IV**). This residue was shown to be conserved in MZ0003 (Phe117), but not in the fungal GE structures. However, the orientation of the co-crystallised ester substrate in the *St*GE2 ligand complex suggested a potential binding interaction with the phenylalanine in the bacterial GEs and an aromatic lignin substituent. To investigate the potential lignin-binding role of the phenylalanine residue and to look for other putative lignin- and carbohydrate-binding sites, docking simulations were performed with *Ot*CE15A and *Su*CE15C with a benzyl ester of 4-*O*-methyl-glucuronoxylotriose (Figure 33C) (**Paper IV**).

The simulations indeed confirmed stacking of the aromatic benzyl moiety on top of the Phe141 and Phe135 in the two enzymes. The carbohydrate portion of the substrate analogue (xylotriose) was more variable in its binding poses, but consistently stacked against a tryptophan residue (Trp358 in *Ot*CE15A and Trp348 in *Su*CE15C). The tryptophan is conserved across both fungal and bacterial GE structures, and has also been reported to participate in hydrogen bonding to the glucuronoate moiety in the *St*GE2 ligand complex (Charavgi et al., 2013). The docking results therefore suggested important interactions with LC esters by two key residues in the bacterial GEs, the "lignin-stacking" phenylalanine and the "glucuronoxylan-binding" tryptophan.

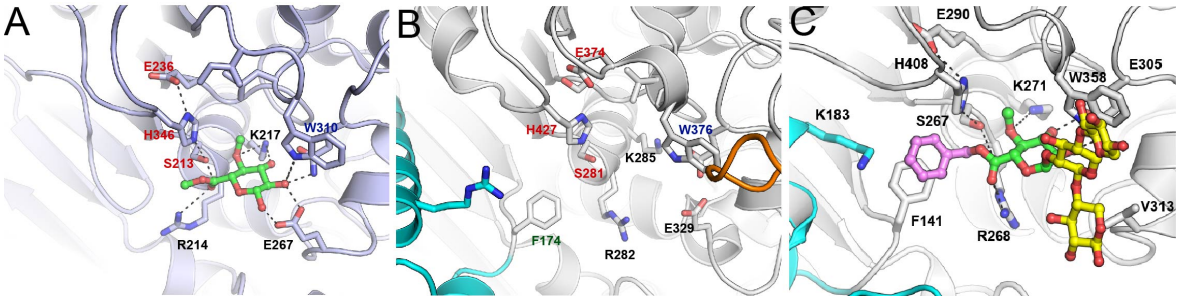


Figure 33. The active site of *St*GE2, co-crystallised with a methyl ester of 4-*O*-methyl glucuronoate (Charavgi et al., 2013) (**A**), compared with the active site of *Tt*CE15A (**Paper V**) (**B**). Residues belonging to the catalytic triad are labelled in red, the "glucuronoxylan-binding" tryptophan in blue, the "lignin-binding" phenylalanine in green, and selected conserved residues that bind to the glucuronate moiety in black. Docking simulation of *Ot*CE15A (**Paper IV**) with a benzyl ester of 4-*O*-methyl-glucuronoxylotriose (**C**) demonstrated the stacking of the benzyl moiety against a phenylalanine and the xylotriose portion against a tryptophan, residues that are conserved in all the solved bacterial GE structures. (Figures courtesy of Scott Mazurkewich.)

Investigation of key residues for substrate interaction
The docking study described above (**Paper IV**), raised questions and suggested hypotheses on key interactions between enzyme residues and the substrate. The characterisation of the novel GE from *T. turnerae* (*Tt*CE15A) was initiated at that time. Its structure revealed the presence of the conserved phenylalanine and tryptophan residues (Phe174 and Trp376), similar to *Ot*CE15A and *Su*CE15C, observed to interact with the aromatics and carbohydrates in the docking study. *Tt*CE15A was therefore subjected to mutagenesis to investigate the potential involvement of Phe174 and Trp376 in substrate interaction (**Paper V**).

Four enzyme variants were created, by substituting "the lignin-stacking" Phe174 or "the glucuronoxylan-binding" Trp376 with alanine or aspartate. The wild-type enzyme and the mutated variants were compared in terms of their activity on model substrates and the inhibition of their action on BnzGlcA by aromatic and carbohydrate compounds. All mutations resulted in *Tt*CE15A being catalytically impaired, indicating the importance of Phe174 and Trp376 for the catalytic function of the enzyme. Inhibition studies on the wild-type enzyme, revealed the interaction of *Tt*CE15A with various aromatic compounds and a GlcA-appended xylotriose ($XUXX_R$), drastically reducing the GE activity. Moreover, comparison of the wild-type enzyme and the mutated W376A variant showed that the presence of Trp376 was required for inhibition by $XUXX_R$ (Table 2 and Figure 4D in **Paper V**). Xylotriose alone inhibited neither the mutated or the wild-type version of *Tt*CE15A, and GlcA was therefore suggested to be crucial for carbohydrate recognition by Trp376.

The results of inhibition studies with aromatic compounds and Phe174 variants of the enzyme were ambiguous, and the involvement of the phenylalanine in binding to these aromatic structures could not be elucidated (Table 2 and Figure 4A-C in **Paper V**). The mutated Phe174 variants were catalytically inhibited by several of the aromatic compounds (similar to, or even more than the wild-type enzyme), possibly due to interactions between other active site residues and, for example, methoxy groups present in some of the aromatic compounds. This work did not provide a complete answer regarding the key residues involved in carbohydrate and lignin recognition, but constitutes a first and important step towards revealing detailed enzyme-substrate interactions. Further mutational studies are required to investigate the key residues in GEs, and as more enzyme structures become available, new target residues suitable for examination will be exposed.

7.7 The biological role of GEs

Since the discovery of GEs, researchers have strived to elucidate the biological role of these enzymes. Understanding their function and diversity in nature would provide fundamental insight into this class of enzymes.

GEs from diverse habitats
The function and the enzymatic mode of action during biomass degradation is likely to be related to the environments in which GE-encoding organisms are found. Bacterial CE15 genes, for example, are encoded by organisms found in diverse habitats (soil, forests, farmland, marine environments, the gastrointestinal tract of animals, etc.), which suggests a considerable variation in biological function and substrate specificity. A phylogenetic study including over 1000 CE15 gene sequences has shown that CE15 genes often originate from coprophilic fungal species (Agger et al., 2017). The ecological habitat of coprophilic fungi

is the droppings of herbivores, containing biomass already highly processed in the gut of the animal. The GE-mediated release of 4-*O*-MeGlcA to increase the yield of carbohydrates from lignin may therefore be important in these species. Coprophilic fungi could serve as attractive sources of GEs, to decouple intertwined parts of lignocellulosic material. CE15 sequences are also frequently found in white-rot fungi and some brown-rot fungi. Brown-rot species produce relatively few glucuronoxylan-degrading enzymes and are poor lignin degraders. Their inability to degrade lignin and hemicellulose to the same extent as other fungal species suggests a different mode of action of the GEs from these fungi. Additionally, brown-rot fungi often attack conifers, and they may, therefore, be interesting in the discovery of GEs active on LC esters from softwood (Agger et al., 2017; Singh & Singh, 2014).

Factors regulating the production of GEs

The fact that several organisms (both fungal and bacterial) encode multiple, non-redundant CE15 enzymes raises questions: firstly, biological differences across CE15 and secondly, different roles within one organism. Upregulation of CE15 genes has been observed in microorganisms when grown on lignocellulosic material (including softwood and hardwood) (Couturier et al., 2015; Hüttner, Nguyen, et al., 2017), but a detailed understanding of the factors that regulate the transcription of these genes is still lacking. An initial transcriptomic study was performed in the present work, including three CE15 enzymes originating from *S. linguale*. *S. linguale* was grown on glucose, xylose, corn cob xylan and ball-milled corn cob biomass, and the expression levels of *slce15a-c* were quantified (**Paper IV**).

The results indeed revealed differences in transcription levels for the three genes compared to growth on glucose. The expression of *slce15a* increased twofold when grown on ball-milled biomass, while *slce15c* showed an increased response with all the xylose-containing carbon sources. The expression of *slce15b* was similar under all growth conditions (Figure 34). The differences in gene regulation in combination with the diverse substrate specificity of the three *Sl*CE15 enzymes support the hypothesis of their various functions (e.g. activity on LCCs of different size or from specific biomass sources).

The transcription responses of *slce15a* and *slce15c* led to the hypothesis that *Sl*CE15A (triggered by growth on complex biomass) acts more readily on large intact LCCs, whereas *Sl*CE15C (also triggered by xylose residues) may instead act mainly on smaller LCC fragments. It is still unclear what upregulates the expression of *slce15b*. This gene might be constitutively expressed, but experiments on other types of carbon sources could possibly provide more information on the function of *Sl*CE15B. Extensive transcriptomic studies and the investigation of a larger number of CE15-encoding organisms could provide comprehensive information on the biology behind the expression of various GEs.

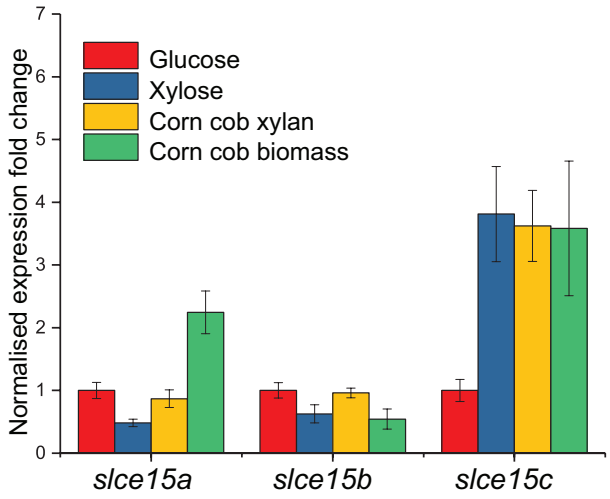


Figure 34. Change in CE15 gene expression (*slce15a*, *slce15b* and *slce15c*) in *S. linguale* cells grown on xylose (blue), corn cob xylan (yellow), and corn cob biomass (green), normalised to growth on glucose (red). The RNA polymerase sigma factor rpoD served as reference gene for the quantitative polymerase chain reaction. Error bars represent the SEM of triplicate measurements.

7.8 GE activity on native lignocellulose

Investigations of sequence diversity, gene expression and substrate specificities on model substrates provide indications concerning diverse functions that may reflect differences in the preferences of GEs for biological substrates. However, the main question of whether GEs are active on native LC esters can only be answered by demonstrating their activity on LCCs in plant biomass. A large part of the present work was, therefore, focused on the detection of GE-catalysed LC ester hydrolysis (**Papers III, IV**).

Indications of GE activity

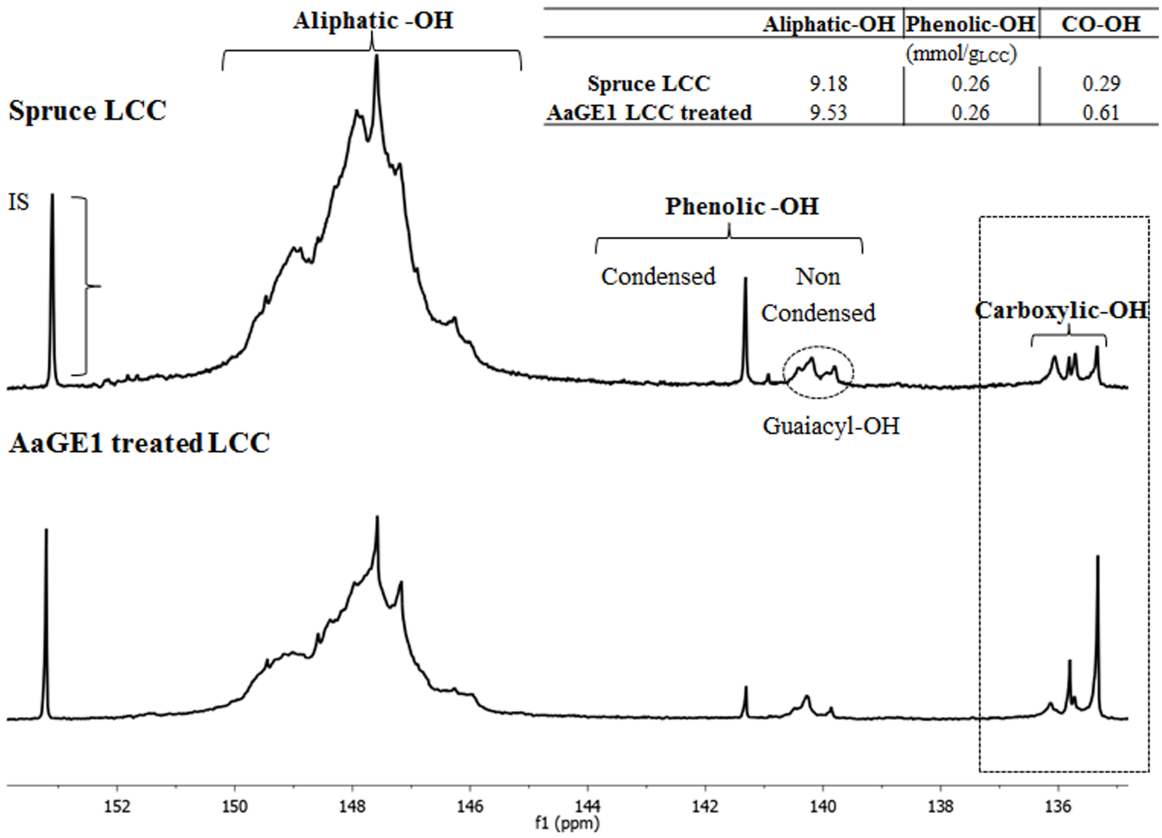
A few indirect methods have been used to verify the activity of GEs on plant biomass. One approach has been to study the overexpression of the fungal GE *Pca*GE from *Phanerochaete carnosa* in plants (*Arabidopsis* and poplar). The results showed altered cell-wall properties and facilitated separation of lignin from the carbohydrate fractions, explained by reduced lignin-carbohydrate crosslinking (Latha Gandla et al., 2015; Tsai et al., 2012). Another way of demonstrating the biological role of GEs has been to confirm the contribution of these enzymes in plant biomass saccharification. *Tr*GE (Cip2) from *T. reesei* has been reported to be determinative for the efficient hydrolysis of corn fibre (Lehmann et al., 2016), and when cellulolytic and hemicellulolytic enzymes were supplemented with GEs in reactions with plant biomass, an increased release of monosaccharides has been observed (discussed in detail in Section 7.9) (**Paper IV**; (D'Errico et al., 2016; Mosbech et al., 2018)). Furthermore, the release of aldouronic acids from lignin-enriched birch wood by a GE from *Cerrena unicolor*, has recently been reported, verifying LC ester cleavage by product release from a native substrate (Mosbech et al., 2018).

Detection of LC ester cleavage

Direct detection of LC ester cleavage would obviously be the best way to confirm the proposed biological role of GEs, but it is necessary to use combined approaches and complementary analytical tools due to the lack of appropriate detection methods. Furthermore, the enrichment of LCCs is key to overcoming the problems associated with the low abundancy of LC esters in plant biomass. These strategies were applied in this research

in order to demonstrate GE activity on biologically important LC esters. The combination of NMR techniques and SEC was successful in demonstrating the cleavage of LC ester bonds extracted from spruce and birch (**Paper III**). A recombinantly produced GE from the fungus *Acremonium alcalophilum* (*Aa*GE1) was incubated with the LCC material, and the reaction products analysed. The release of lignin from the carbohydrate fraction was confirmed by SEC, and the increase in COOH was detected upon GE treatment using ^{31}P NMR (Figure 35). The increase in COOH indeed demonstrated the esterase activity of *Aa*GE1, and HSQC verified that the increase in COOH was not an effect of acetyl esterase activity.

Recently obtained data also showed that the combined action of bacterial GEs (*Su*CE15C or *Tt*CE15A) and a GH115 α-glucuronidase released 4-*O*-MeGlcA from birch-derived LCCs. The release of 4-*O*-MeGlcA was detected and confirmed by TLC and LC-MS (unpublished work from collaboration with Professor Master's group at Toronto University). Although it can be questioned whether this constitutes direct proof, the use of new and combined detection techniques makes the findings of this research the most reliable evidence of GE activity on native LC esters to date.



	Aliphatic-OH	Phenolic-OH	CO-OH
		(mmol/g_{LCC})	
Spruce LCC	9.18	0.26	0.29
AaGE1 LCC treated	9.53	0.26	0.61

Figure 35. ^{31}P NMR spectra of phosphitylated LCCs from spruce before and after treatment with *Aa*GE1. The dotted box (at 136 ppm) indicates the region in which carboxylic-OH (COOH) is identified. The concentrations of hydroxyl groups before and after enzymatic treatment presented in the table were calculated by normalisation to the internal standard (N-hydroxy-5-norbornene-2,3-dicar-boximide). The concentration of COOH increased twofold upon treatment with *Aa*GE1. Similar results were obtained for birch-derived LCCs (Figure 3B in **Paper III**).

Detection of LC ester cleavage by 2D NMR

2D NMR techniques have been regarded as promising methods (and by some, as possible direct methods) of assaying GE activity on native LC esters. However, the heavily overlapped γ-ester signal has been a problem. The α-linked benzyl ester is located in a less noisy region, but has only rarely been observed in 2D NMR spectra from LCCs. However, the research group of Professor Watanabe at Kyoto University has achieved HSQC spectra with a clear benzyl ester signal when studying LCCs extracted from Japanese beech (Figure 36A). The same LCC fractions were used for reactions with *Su*CE15C. *Su*CE15C was applied to the LCC material at increasing concentrations and incubated for 48 hours, after which the reactions were analysed with HSQC. The HSQC spectra obtained clearly showed a signal that could be assigned to the benzyl ester, which decreased with increasing concentration of *Su*CE15C (Figure 36B-D). These results clearly show the action of *Su*CE15C on native benzyl esters, and the potential of HSQC as a method of detecting GE activity. However, replicate experiments on LCC materials containing benzyl esters, and preferably also γ-ester (frequently detected with 2D NMR and possibly biologically more relevant) must be performed to confirm these findings.

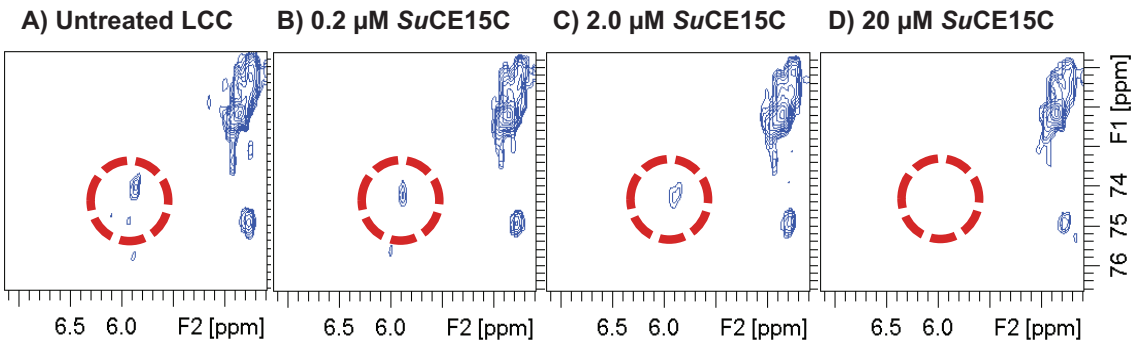


Figure 36. HSQC spectra from GE-treated LCCs isolated from Japanese beech (48-h incubations). The red circle indicates the region identified as the benzyl ester. The signal assigned to benzyl esters in native LCCs (**A**), can be seen to decrease with increasing concentration of the *Su*CE15C (**B-D**). (Figures courtesy of Hiroshi Nishimura.)

7.9 Biotechnical potential of GEs

The demonstration of the action of GEs on LC esters indicates their industrial potential as delignifying agents in a wood-based biorefinery. However, the industrial benefit of these enzymes has been less well-explored. In this work, the potential boosting effect of GEs was investigated, and improved enzymatic saccharification of biomass was observed when adding GEs as auxiliary enzymes. The hemicellulolytic cocktail Ultraflo®, rich in xylanases and cellulases, was used in reactions with ball-milled corn cob (a plant biomass rich in xylan), with or without the supplementation of GEs (**Paper IV**). An increase in xylose and arabinose release was observed in the GE-enhanced reactions (Figure 37), likely due to enhanced enzymatic access to the xylan polymer after separation from the lignin. However, the GEs were not only capable of boosting the xylan-acting enzymes, but also acted synergistically with the cellulolytic enzymes, as evidenced by a significant increase in glucose yield (Figure 37). This suggests that there is a close relation between chemical structures and the ultra-structure in lignocellulose, and that changes at one level affects the other. However, it is not known whether GEs act prior to, in concert with, or after, xylanases in separating lignin from xylan. Temporally resolved assays, in which the enzymes are added

in different orders and in different combinations could provide the answer to this question. Using GEs to boost enzymatic saccharification on an industrial scale must be economically feasible, and the identification of, or engineering to achieve, more efficient GEs might also be required.

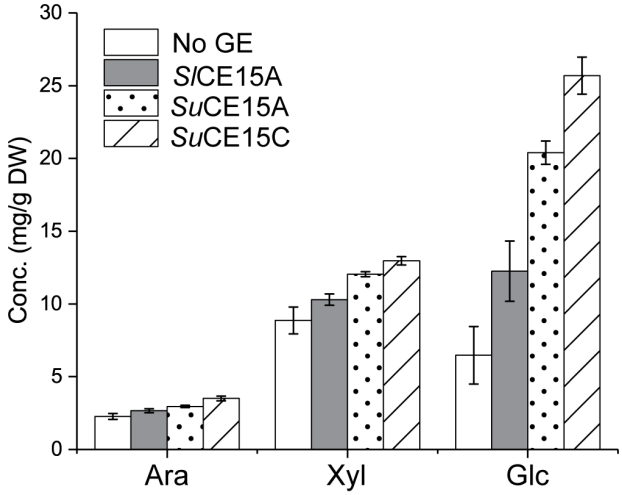


Figure 37. Release of monosaccharides (arabinose, xylose and glucose) after 24-h enzymatic hydrolysis of ball-milled corn cob using Ultraflo®, without the addition of GE (white), and with the addition of *Sl*CE15A (grey), *Su*CE15A (dotted), or *Su*CE15C (hatched). Error bars represent the SEM of triplicate reactions.

The industrial use of GEs as delignifying agents can be questioned due to the labile nature of LC esters, which are easily hydrolysed during thermo-chemical pre-treatment. However, replacing traditional forms of pre-treatment with GEs could be beneficial as this would reduce the demand for energy and chemicals in the processing of woody biomass. Combining GEs with less severe pre-treatment methods might also be advantageous. Mild STEX pre-treatment followed by hydrolysis using GEs could be an efficient means of loosening the dense wood structure and removing lignin, while keeping the hemicelluloses intact. Pre-treatment at extreme pH values and temperatures further causes lignin condensation, resulting in lignin with a lower quality. GEs might therefore be useful as an alternative method of pre-treatment for selective lignin recovery for value-added applications (Mosbech et al., 2018).

GEs could theoretically be used in combination with other lignocellulolytic enzymes for the production of specific oligosaccharides and chemicals. One example of this is to combine the action of GEs and α-glucuronidases, resulting in the release of 4-*O*-MeGlcA from lignocellulose. The 4-*O*-MeGlcA could then be further converted into glucaric acid, which is regarded as one of the top value-added platform chemicals for a range of applications (Chen et al., 2018). Another possible use of GEs could be as a bleaching aid in the production of pulp, by the removal of residual xylans. GEs clearly have industrial potential, but transforming theoretical ideas into practical applications will require the evaluation of these enzymes on a wide range of lignocellulosic materials in combination with different kinds of pre-treatment, and together with other CAZymes.

8. Conclusions

Lignocellulolytic enzymes are promising tools for the selective decoupling and deconstruction of woody biomass, but we still lack a detailed understanding of how these enzymes interact with, and act on, complex lignocellulosic structures, which is required to exploit their full potential. The scientific research summarised in this thesis has contributed important knowledge on specific enzymes that could be used in a wood-based biorefinery for the production of value-added polymeric and oligomeric components.

The importance of both the physical and chemical effects of STEX pre-treatment on enhanced enzyme hydrolysis of wood chips was demonstrated **(Paper I).** The structural changes resulting from the explosion step played a vital role in the improved enzyme performance, which can be explained by **the enhanced accessibility of the substrate to the enzymes**. Furthermore, it was proposed that mild STEX, with a lower chemical impact on hemicellulose hydrolysis, could be suitable for pre-treatment in a materials biorefinery, in order to retain more of the polymeric wood structure.

Detailed information on **mannanase hydrolysis of the chemically complex and industrially important SpGGM** was obtained (**Paper II**) using a range of advanced detection methods including HPAEC-PAD and ESI-MS. These studies revealed significant differences in the mode of action of two mannanases on SpGGM; *Cj*Man5A generating M3 and M4 as the smallest oligosaccharides, and *Cj*Man26A producing a large amount of M1 and M2. The low DSac of native SpGGM had no or little effect on mannanase hydrolysis, suggesting that galactose decorations and the composition of the mannan backbone had a greater impact on the mannanase reaction rate and product formation. However, chemical acetylation of mannan polymers was shown to significantly hinder enzymatic action, and was suggested as a means of limiting the biodegradation of hemicellulolytic material.

The studies presented in **Papers III, IV** and **V** provided novel information, contributing to **our understanding of GEs and their potential in hydrolysing recalcitrant LC ester linkages** in woody biomass. This work demonstrated, for the first time, GE-catalysed LC ester cleavage in wood-derived LCC fractions (**Paper III**), constituting an important advancement from previous studies on characterisation using small model substrates. The potential of GEs in facilitating the action of other hydrolytic enzymes by making the recalcitrant lignocellulose more accessible was also demonstrated (**Paper IV**). Furthermore, a fundamental understanding of the relatively unexplored carbohydrate esterase family 15 was achieved. GE diversity and substrate specificity were investigated by the systematic characterisation of 10 putative GEs across CE15 (**Paper IV**), which constitutes the most comprehensive study of GEs to date. Results from complementary methods of analysis:

biochemical characterisation, structural determination, transcriptomic analysis, mutagenesis and inhibition studies, demonstrated distinct differences in the structure and catalytic function of these enzymes. This work also brought bacterial GEs to light and addressed their high catalytic efficiencies and considerable variation in substrate preferences (e.g. alcohol portion and uronic acid recognition) compared to previously characterised fungal GEs (**Papers IV, V**).

9. Future perspectives

Several questions remain to be answered before we can fully understand how specific enzymes decouple and deconstruct the polymers in woody biomass, both in nature and in industrial settings. In this final chapter, suggestions for future research, essential for continuing the work presented in this thesis, will be discussed.

Considerable effort should be devoted to developing and evaluating mild pre-treatment methods that could improve enzymatic performance on wood structures, where the aim is to upgrade the biomass components. Here, it is important to make woody biomass more accessible to the enzymes, while avoiding lignin condensation and the decomposition of the hemicelluloses. The development of methods that are able to detect enzyme interactions in complex substrates will be of key importance if we are to fully understand the contributions of the various parameters that influence enzyme accessibility in lignocellulosic structures. Attempts to use FRAP for this purpose with GFP-CBM probes were not completely successful, but the technique has been used successfully to study the mobility of dextran probes in wood structures (Paës et al., 2017). The next step is to adapt FRAP for enzyme analysis in wood, to obtain insight into the correlation between structural accessibility and catalytic efficiency. Investigations of the substrate interaction of specific enzymes with diverse activities will further improve our understanding of how different lignocellulosic structures and components are affected by pre-treatment.

We are still far from comprehending the mechanisms behind the enzymatic hydrolysis of heterogeneous and decorated hemicelluloses in woody biomass, and future research will involve continued work on elucidating how the chemical structure affects the action of enzymes. This will require research on complex, lignocellulose-derived substrates from various wood sources. Continued research on mannanases from different families and microorganisms will be particularly interesting regarding the selective degradation of softwood GGM. Knowledge on the synergism between mannanases and/or other mannanolytic enzymes would provide further insight into enzymatic modes of action and which key structures in the substrate affect the activity of individual enzymes.

Elucidating the influence of native acetylation on enzymatic action is especially relevant in regard to the development of mild pre-treatment and fractionation methods, leading to reduced deacetylation of wood hemicelluloses. The findings presented in this thesis showed that a low DSac had little effect on the enzymatic cleavage pattern of the mannanases investigated, but complementary studies on SpGGM containing higher DSac are required. A high degree of acetylation has been reported to have a negative influence on enzymatic hydrolysis (Von Freiesleben et al., 2018; Tenkanen et al., 1995), but the importance of the

position of the acetyl group in reducing substrate hydrolysability is still unknown. Tandem MS (MS^2), involving multiple steps of fragmentation, is a promising tool for detecting the exact positions of acetyl groups on oligosaccharides (Kailemia et al., 2014). The application of MS^2 to mannanase-treated SpGGM could provide detailed information on acetylation patterns, possibly providing novel information on enzyme tolerance to the acetyl position, and contributing to the elucidation of the exact chemical structure of SpGGM. Determining the positions of acetyl groups in various hemicelluloses would give further indications of how plants may use acetylation as a form of protection against microbial degradation in nature.

The work presented here contributed significantly to knowledge on GEs and the enzymatic decoupling of lignin from carbohydrates, but the number of GEs characterised is still low compared to other CAZymes. Future research should continue investigating substrate specificity, and the characterisation of additional CE15 members may possibly reveal new enzyme activities. The fact that several bacterial GE candidates exhibited activity on esters of GalA is noteworthy (**Paper IV**), and indicates the existence of pectin-active CE15 enzymes. Feruloyl and acetyl esterase activities previously observed in a few GEs (De Santi et al., 2016; Mosbech et al., 2018) may also be found in a larger number of enzymes. Assaying CE15 members for these esterase activities would contribute to our knowledge on substrate promiscuity across the family. Detailed knowledge on important substrate features, e.g. 4-*O*-methylation and preferences to α- or γ-linked esters is still lacking in order to fully comprehend the biology and diversity of GEs. The absence of relevant and stable model substrates has always been problematic, and new efforts in substrate synthesis are therefore necessary. Knowledge on enzyme-substrate interactions could also be obtained by structural studies of a larger number of CE15 enzymes and GEs in complex with various ligands.

In order to understand the nature of GEs, it is necessary to use biological substrates. These include substrates derived from a variety of plant and wood species, in native and pre-treated form. Non-destructive extraction processes that can isolate LCCs enriched in LC esters, at sufficient levels for detection, will be important in this context. The development of efficient mild extraction protocols has already made LCCs more suitable for use as enzyme substrates (**Paper III**). However, the detection of LC esters using 2D NMR suffers from overlapping signals from carbohydrate signals. Polishing the LCC fractions by adding hemicellulases, e.g. xylanases, could therefore be a promising approach in the extraction procedure.

The detection of LC ester cleavage is dependent on the availability of suitable substrates, but also requires more sensitive detection methods. The analytical toolbox is continuously expanding with improved NMR techniques that will provide insight on the molecular level of woody biomass. Long-range correlated 2D NMR techniques, such as HSQC-total correlation spectroscopy (TOCSY) and heteronuclear multiple bond correlation spectroscopy (HMBC), correlating multiple neighbouring atoms, have been reported to be promising in detecting key structures and linkages in lignocellulosic biomass with higher accuracy than HSQC (Nishimura et al., 2018). These NMR techniques should be evaluated for assaying GE activity on native LCCs. As demonstrated in this work, it is essential to confirm native GE activity by using multiple complementary analytical tools. Methods that directly detect changes in the lignocellulosic structure upon GE treatment (e.g. NMR and imaging techniques) should preferably be combined with techniques that analyse enzymatically released products (e.g. HPAEC-PAD and MS). Improved and more sensitive analytical methods are also of wider interest for evaluating the activity of other CAZymes on complex lignocellulosic material.

Several as yet unidentified enzyme activities have been proposed to exist in nature. In order to identify these activities, and to develop suitable substrates for screening procedures, it is necessary to gain a deeper understanding of lignocellulosic structures and potential enzyme targets. An example related to LCCs, is the LC ether linkage between lignin and xylan, on which no known enzymes are able to act. Finding hypothetical "LC etherases", will be an important future research task, crucial in elucidating the enzymatic machineries of microorganisms and for the efficient extraction and polishing of the polymers in woody biomass.

Research on specific lignocellulosic enzymes is highly dependent on the availability of suitable detection methods and interdisciplinary research involving enzymologists, wood chemists and organic chemists. The understanding of lignocellulosic structures and the use of relevant substrates are key in obtaining deeper knowledge on the behaviour of enzymes in nature, and important in evaluating their suitability in a biorefinery context. Fundamental knowledge and the industrial potential of enzymes are, after all, closely connected.

Acknowledgements

Life as a PhD student would have been very lonely without the many people who endured, supported, challenged and believed in me during the whole or part of this journey. I want to express my gratitude to all of you. Some have been deeply involved in my research projects, some have been important in other parts of my life, and some have been part of both.

Lisbeth Olsson, my main supervisor. Thank you for responding to my initial e-mail, for guiding me, believing in me, and challenging me along the way. You have helped me achieve what at first seemed impossible, and for that I am deeply grateful.

Johan Larsbrink, my co-supervisor, who is a true enzymologist, and introduced me to the world of CAZymes. Thank you for your honesty, and for always being available, whether it was for long scientific discussions or trivial lab issues.

Scott Mazurkewich: what would "Team GE" have been without you? You are a great colleague, office mate, tutor and friend. It has truly been a pleasure to share ideas, successes, ups, downs, disappointments, publications, phosphate buffers and cheesy Christmas playlists with you. From now on, someone else will have to steal your pens.

Sylvia Klaubauf: thank you for mentoring me during the first part of my PhD studies, and for introducing me to lab work.

Amanda Sörensen Ristinmaa: you were the most diligent and curious Master's student, and gave me a chance to practice supervision. It was a challenging but exciting learning process for both of us. Thank you!

I would also like to acknowledge my collaborators: **Francisco Vilaplana, Jennie Berglund, Antonio Martínez-Abad, Martin Lawoko, Nicola Giummarella, Anders Rasmuson, Muhammed Muzamal, Patric Kvist** and **Leila Lo Leggio.** Our work has resulted in publications, novel findings and valuable knowledge exchange.

I wish to thank all the present and former members of the Industrial Biotechnology Division for their contributions to the open and accepting lab environment. It has been fun to hang out with all the IndBio PhD students, and share our struggles and concerns. Keep up the good work and continue to fight for your rights! **Joakim** and **Joshua**, it was not always a pleasure to operate the IC, but I couldn't have wished for better co-superusers than you, and we managed to keep our heads above the water (most of the times). **Julia L**, you have not only been a great help in the lab, but also my daily therapist. Thank you for your constant

reminders about work-life balance. **Silfa**, it has been great fun to spend time with you in the lab, on American road-trips and in Gothenburg. **Helén**, I have to thank you as both a research engineer and as a fellow PhD student. I really appreciated our time in Delft. **Lina**, you have been the experienced and wiser big sister, exercising patience with my incessant questions. Thanks for all the help and motivation you have given over the years. **Cecilia T**, even though you are not part of IndBio, you have been extremely valuable in my daily life at Chalmers. Uninspired lunch boxes and lukewarm coffee taste better in your company. **Ausra**, **Emma** and **Vera**, you must be collectively mentioned here, because you are my true rock stars.

I am also glad to have been part of the Wallenberg Wood Science Center, and want to acknowledge **Lars Berglund** for managing the WWSC and **Paul Gatenholm** for directing the WWSC Academy. WWSC PhD students: we have definitely built up a valuable social and professional network during these years.

I am grateful for all the opportunities I had to attend conferences and courses that took me abroad, and I am especially grateful for my research visits to Japan and Canada. I would like to thank **Takashi Watanabe** and **Hiroshi Nishimura** for inviting me to collaborate with them at Kyoto University, and **Emma Master** for welcoming me to her group at Toronto University. **Olan** and **Thu**, it was a pleasure working with you, combining our enzymes, knowledge and humour.

Finally, I want to thank my family and friends for all their love and support, especially my parents **Kerstin** and **Erland**. You have always supported and encouraged me in everything I have undertaken.

Bibliography

Agger, J. W., Busk, P. K., Pilgaard, B., Meyer, A. S., & Lange, L. (2017). A new functional classification of glucuronoyl esterases by peptide pattern recognition. *Frontiers in Microbiology, 8*, 309.

Alvira, P., Tomas-Pejo, E., Ballesteros, M., & Negro, M. J. (2010). Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. *Bioresource Technology, 101*(13), 4851-4861.

Anderson, I., Tindall, B. J., Pomrenke, H., Goker, M., Lapidus, A., Nolan, M., et al. (2009). Complete genome sequence of *Halorhabdus utahensis* type strain (AX-2^T). *Standards in Genomic Sciences, 1*(3), 218-225.

Argyropoulos, D. S. (2010). Heteronuclear NMR Spectroscopy of Lignins. In C. Heitner, D. Dimmel, & J. A. Schmidt (Eds.), *Lignins and Lignans: Advances in Chemistry* (pp. 245-263): CRC Press.

Armenta, S., Moreno-Mendieta, S., Sanchez-Cuapio, Z., Sanchez, S., & Rodriguez-Sanoja, R. (2017). Advances in molecular engineering of carbohydrate-binding modules. *Proteins, 85*(9), 1602-1617.

Aurilia, V., Martin, J. C., McCrae, S. I., Scott, K. P., Rincon, M. T., & Flint, H. J. (2000). Three multidomain esterases from the cellulolytic rumen anaerobe *Ruminococcus flavefaciens* 17 that carry divergent dockerin sequences. *Microbiology, 146*, 1391-1397.

Auxenfans, T., Cronier, D., Chabbert, B., & Paës, G. (2017). Understanding the structural and chemical changes of plant biomass following steam explosion pretreatment. *Biotechnology for Biofuels, 10*, 36.

Balakshin, M., Capanema, E., & Berlin, A. (2014). Isolation an Analysis of Lignin-Carbohydrate Complexes Preparations with Traditional and Advanced Methods: A Review. In Atta-ur-Rahman (Ed.), *Studies in Natural Products Chemistry* (Vol. 42, pp. 83-115): Elsvier.

Balakshin, M., Capanema, E., Gracz, H., Chang, H. M., & Jameel, H. (2011). Quantification of lignin-carbohydrate linkages with high-resolution NMR spectroscopy. *Planta, 233*(6), 1097-1110.

Balakshin, M. Y., Capanema, E. A., & Chang, H. M. (2007). MWL fraction with a high concentration of lignin-carbohydrate linkages: Isolation and 2D NMR spectroscopic analysis. *Holzforschung, 61*(1), 1-7.

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

Baumberger, S., Abaecherli, A., Fasching, M., Gellerstedt, G., Gosselink, R., Hortling, B., et al. (2007). Molar mass determination of lignins by size-exclusion chromatography: towards standardisation of the method. *Holzforschung, 61*(4), 1-10.

Berntsson, T., Sandén, B., Olsson, L., & Åsblad, A. (2013). What is a biorefinery? In *Systems perspectives on biorefineries 2013* (pp. 16-25). Chalmers University of Technology: Göteborg, Sweden.

Bi, R., Berglund, J., Vilaplana, F., McKee, L. S., & Henriksson, G. (2016). The degree of acetylation affects the microbial degradability of mannans. *Polymer degradation and stability, 133*, 36-46.

Biely, P. (2012). Microbial carbohydrate esterases deacetylating plant polysaccharides. *Biotechnology advances, 30*(6), 1575-1588.

Biely, P., De Vries, R. P., Vrsanska, M., & Visser, J. (2000). Inverting character of α-glucuronidase A from *Aspergillus tubingensis*. *Biochimica et Biophysica Acta, 1474*(3), 360-364.

Biely, P., Malovíková, A., Uhliariková, I., Li, X. L., & Wong, D. W. (2015). Glucuronoyl esterases are active on the polymeric substrate methyl esterified glucuronoxylan. *Febs Letters, 589*(18), 2334-2339.

Bisswanger, H. (2014). Enzyme assays. *Perspectives in Science, 1*(1-6), 41-55.

Björkman, A. (1957). Studies on finely divided wood. II. Extraction of lignin-carbohydrate complexes with neutral solvents. *Svensk Papperstidning, 60*, 243-251.

Bornscheuer, U. T. (2002). Microbial carboxyl esterases: classification, properties and application in biocatalysis. *FEMS Microbiology Reviews, 26*(1), 73-81.

Brändström, J. (2001). Micro- and ultrastructural aspects of Norway spruce tracheids: a review. *IAWA Journal, 22*(4), 333-353.

Brunow, G., & Lundquist, K. (2010). Functional Groups and Bonding Patterns in Lignin (Including the Lignin-Carbohydrate Complexes). In C. Heitner, D. Dimmel, & J. A. Schmidt (Eds.), *Lignin and Lignans*. Boca Raton, FL: CRC Press.

Capek, P., Alfoldi, J., & Liskova, D. (2002). An acetylated galactoglucomannan from *Picea abies* L. Karst. *Carbohydrate Research, 337*(11), 1033-1037.

Cataldi, T. R., Campa, C., & De Benedetto, G. E. (2000). Carbohydrate analysis by high-performance anion-exchange chromatography with pulsed amperometric detection: the potential is still growing. *Fresenius Journal of Analytical Chemistry, 368*(8), 739-758.

Chandra, R. P., Bura, R., Mabee, W. E., Berlin, A., Pan, X., & Saddler, J. N. (2007). Substrate pretreatment: The key to effective enzymatic hydrolysis of lignocellulosics? *Biofuels, 108*, 67-93.

Chandra, R. P., Ewanick, S. M., Chung, P. A., Au-Yeung, K., Del Rio, L., Mabee, W., et al. (2009). Comparison of methods to assess the enzyme accessibility and hydrolysis of pretreated lignocellulosic substrates. *Biotechnology Letters, 31*(8), 1217-1222.

Charavgi, M. D., Dimarogona, M., Topakas, E., Christakopoulos, P., & Chrysina, E. D. (2013). The structure of a novel glucuronoyl esterase from *Myceliophthora thermophila* gives new insights into its role as a potential biocatalyst. *Acta crystallographica. Section D, Biological crystallography, 69*, 63-73.

Chen, N., Wang, J. Y., Zhao, Y. Y., & Deng, Y. (2018). Metabolic engineering of *Saccharomyces cerevisiae* for efficient production of glucaric acid at high titer. *Microbial Cell Factories, 17*, 67.

Couturier, M., Navarro, D., Chevret, D., Henrissat, B., Piumi, F., Ruiz-Duenas, F. J., et al. (2015). Enhanced degradation of softwood versus hardwood by the white-rot fungus *Pycnoporus coccineus*. *Biotechnology for Biofuels, 8*, 216.

Couturier, M., Roussel, A., Rosengren, A., Leone, P., Stålbrand, H., & Berrin, J. G. (2013). Structural and biochemical analyses of glycoside hydrolase families 5 and 26 β-(1,4)-mannanases from *Podospora anserina* reveal differences upon manno-oligosaccharide catalysis. *Journal of Biological Chemistry, 288*(20), 14624-14635.

Cuyvers, S., Hendrix, J., Dornez, E., Engelborghs, Y., Delcour, J. A., & Courtin, C. M. (2011). Both substrate hydrolysis and secondary substrate binding determine xylanase mobility as assessed by FRAP. *Journal of Physical Chemistry B, 115*(16), 4810-4817.

D'Errico, C., Börjesson, J., Ding, H. S., Krogh, K. B. R. M., Spodsberg, N., Madsen, R., et al. (2016). Improved biomass degradation using fungal glucuronoyl-esterases-hydrolysis of natural corn fiber substrate. *Journal of Biotechnology, 219*, 117-123.

D'Errico, C., Jorgensen, J. O., Krogh, K. B., Spodsberg, N., Madsen, R., & Monrad, R. N. (2015). Enzymatic degradation of lignin-carbohydrate complexes (LCCs): model studies using a fungal glucuronoyl esterase from *Cerrena unicolor*. *Biotechnology and Bioengineering, 112*(5), 914-922.

Da Silva, J. C., De Oliveira, R. C., Neto, A. D., Pimentel, V. C., & Dos Santos, A. D. (2015). Extraction, addition and characterization of hemicelluloses from corn cobs to development of paper properties. *Procedia Material Science, 8*, 793-801.

Daniel, G. (2009). Wood and Fibre Morphology. In M. Ek, G. Gellerstedt, & G. Henriksson (Eds.), *Wood Chemistry and Wood Biotechnology*. Berlin: Walter de Gruyter.

Davies, G., & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure, 3*(9), 853-859.

Davies, G. J., Wilson, K. S., & Henrissat, B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochemical Journal, 321*, 557-559.

De Santi, C., Gani, O. A., Helland, R., & Williamson, A. (2017). Structural insight into a CE15 esterase from the marine bacterial metagenome. *Scientific Reports, 7*, 17278.

De Santi, C., Willassen, N. P., & Williamson, A. (2016). Biochemical characterization of a family 15 carbohydrate esterase from a bacterial marine arctic metagenome. *PLoS One, 11*(7), e0159345.

Dey, P. M. (1978). Biochemistry of plant galactomannans. *Advances in Carbohydrate Chemistry and Biochemistry, 35*, 341-376.

Dhawan, S., & Kaur, J. (2007). Microbial mannanases: An overview of production and applications. *Critical Reviews in Biotechnology, 27*(4), 197-216.

Dilokpimol, A., Makela, M. R., Cerullo, G., Zhou, M., Varriale, S., Gidijala, L., et al. (2018). Fungal glucuronoyl esterases: Genome mining based enzyme discovery and biochemical characterization. *New Biotechnology, 40*(Pt B), 282-287.

Duff, S. J. B., & Murray, W. D. (1996). Bioconversion of forest products industry waste cellulosics to fuel ethanol: a review. *Bioresource Technology, 55*, 1-33.

Ďuranová, M., Hirsch, J., Kolenova, K., & Biely, P. (2009). Fungal glucuronoyl esterases and substrate uronic acid recognition. *Bioscience Biotechnology and Biochemistry, 73*(11), 2483-2487.

Ďuranová, M., Špániková, S., Wosten, H. A., Biely, P., & De Vries, R. P. (2009). Two glucuronoyl esterases of Phanerochaete chrysosporium. *Archives of Microbiology, 191*(2), 133-140.

Duval, A., & Lawoko, M. (2014). A review on lignin-based polymeric, micro- and nano-structured materials. *Reactive & Functional Polymers, 85*, 78-96.

Ebringerová, A. (2006). Structural diversity and application potential of hemicelluloses. *Macromolecular Symposia, 232*, 1-12.

Erdmann, J. (1866). Ueber der Concretionen in den Birnen. *Annalen der Chemie und Pharmacie, 138*, 1-19.

Esselin, A. (2017). *Discontinuities and innovation-oriented challanges in the value chain of bioeconomy*. Retrieved from Stockholm: The Swedish foundation for strategic environmental research (MISTRA)

Ethier, N., Talbot, G., & Sygusch, J. (1998). Gene cloning, DNA sequencing, and expression of thermostable β-mannanase from *Bacillus stearothermophilus*. *Applied and Environmental Microbiology, 64*(11), 4428-4432.

Evtuguin, D. V., Goodfellow, C., Pascoal Neto, C., & Terashima, N. (2005). *Characterization of lignin-carbohydrate linkages in Eucalyptus globulus by 2D/3D NMR spectroscopy using specific carbon-13 labelling technique.* Paper presented at the 13th ISWFPC, Auckland, New Zealand.

Fengel, D., & Wegener, G. (1989). *Wood - Chemistry, Ultrastructure, Reactions*. Berlin: Walter de Gruyter.

Foreman, P. K., Brown, D., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N. S., et al. (2003). Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. *Journal of Biological Chemistry, 278*(34), 31988-31997.

Foston, M., & Ragauskas, A. J. (2012). Biomass characterization: Recent progress in understanding biomass recalcitrance. *Industrial Biotechnology, 8*(4), 191-208.

Fraňová, L., Puchart, V., & Biely, P. (2016). β-glucuronidase-coupled assays of glucuronoyl esterases. *Analytical Biochemistry, 510*, 114-119.

Gibson, L. J. (2012). The hierarchical structure and mechanics of plant materials. *Journal of the Royal Society Interface, 9*(76), 2749-2766.

Gilbert, H. J., Stålbrand, H., & Brumer, H. (2008). How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation. *Current Opinion in Plant Biology, 11*(3), 338-348.

Giummarella, N. (2018). *Fundamental aspects of lignin carbohydrate complexes (LCC) - Mechanisms, recalcitrance and material concepts.* (Doctoral thesis), KTH Royal Institute of Technology, Stockholm, Sweden.

Giummarella, N., & Lawoko, M. (2016). Structural basis for the formation and regulation of lignin-xylan bonds in birch. *Acs Sustainable Chemistry & Engineering, 4*(10), 5319-5326.

Giummarella, N., & Lawoko, M. (2017). Structural insights on recalcitrance during hydrothermal hemicellulose extraction from wood. *Acs Sustainable Chemistry & Engineering, 5*(6), 5156-5165.

Giummarella, N., Zhang, L. M., Henriksson, G., & Lawoko, M. (2016). Structural features of mildly fractionated lignin carbohydrate complexes (LCC) from spruce. *Rsc Advances, 6*(48), 42120-42131.

Grishkewich, N., Mohammed, N., Tang, J. T., & Tam, K. C. (2017). Recent advances in the application of cellulose nanocrystals. *Current Opinion in Colloid & Interface Science, 29*, 32-45.

Guillen, D., Sanchez, S., & Rodriguez-Sanoja, R. (2010). Carbohydrate-binding domains: multiplicity of biological roles. *Applied Microbiology and Biotechnology, 85*(5), 1241-1249.

Gusakov, A. V., Kondratyeva, E. G., & Sinitsyn, A. P. (2011). Comparison of two methods for assaying reducing sugars in the determination of carbohydrase activities. *International Jornal of Analytical Chemistry, 2011*, 283658.

Hannuksela, T., & Herve du Penhoat, C. (2004). NMR structural determination of dissolved *O*-acetylated galactoglucomannan isolated from spruce thermomechanical pulp. *Carbohydrate Research, 339*(2), 301-312.

Henriksson, G. (2009). Lignin. In M. Ek, G. Gellerstedt, & G. Henriksson (Eds.), *Wood Chemistry and Wood Biotechnology*. Berlin: Walter de Gruyter.

Henriksson, G. (2017). What are the biological functions of lignin and its complexation with carbohydrates? *Nordic Pulp & Paper Research Journal, 32*(4), 527-541.

Henriksson, G., & Lennholm, H. (2009). Cellulose and Carbohydrate Chemistry. In M. Ek, G. Gellerstedt, & G. Henriksson (Eds.), *Wood Chemistry and Wood Biotechnology*. Berlin: Walter de Gruyter.

Herbaut, M., Zoghlami, A., & Paës, G. (2018). Dynamical assessment of fluorescent probes mobility in poplar cell walls reveals nanopores govern saccharification. *Biotechnology for Biofuels, 11*, 271.

Hogg, D., Pell, G., Dupree, P., Goubet, F., Martin-Orue, S. M., Armand, S., et al. (2003). The modular architecture of *Cellvibrio japonicus* mannanases in glycoside hydrolase families 5 and 26 points to differences in their role in mannan degradation. *Biochemical Journal, 371*, 1027-1043.

Hogg, D., Woo, E. J., Bolam, D. N., McKie, V. A., Gilbert, H. J., & Pickersgill, R. W. (2001). Crystal structure of mannanase 26A from *Pseudomonas cellulosa* and analysis of residues involved in substrate binding. *Journal of Biological Chemistry, 276*(33), 31186-31192.

Horn, S. J., Vaaje-Kolstad, G., Westereng, B., & Eijsink, V. G. H. (2012). Novel enzymes for the degradation of cellulose. *Biotechnology for Biofuels, 5*, 45.

Huang, L., Takahashi, R., Kobayashi, S., Kawase, T., & Nishinari, K. (2002). Gelation behavior of native and acetylated konjac glucomannan. *Biomacromolecules, 3*(6), 1296-1303.

Hüttner, S., Klaubauf, S., de Vries, R. P., & Olsson, L. (2017). Characterisation of three fungal glucuronoyl esterases on glucuronic acid ester model compounds. *Applied Microbiology and Biotechnology, 101*(13), 5301-5311.

Hüttner, S., Nguyen, T. T., Granchi, Z., Chin, A. W. T., Ahren, D., Larsbrink, J., et al. (2017). Combined genome and transcriptome sequencing to investigate the plant cell wall degrading enzyme system in the thermophilic fungus *Malbranchea cinnamomea*. *Biotechnology for Biofuels, 10*, 265.

Huynh, H. H., & Arioka, M. (2016). Functional expression and characterization of a glucuronoyl esterase from the fungus *Neurospora crassa*: identification of novel consensus sequences containing the catalytic triad. *Journal of General and Applied Microbiology, 62*(5), 217-224.

Huynh, H. H., Ishii, N., Matsuo, I., & Arioka, M. (2018). A novel glucuronoyl esterase from *Aspergillus fumigatus*-the role of conserved Lys residue in the preference for 4-*O*-methyl glucuronoyl esters. *Applied Microbiology and Biotechnology, 102*(5), 2191-2201.

Jacquet, N., Maniet, G., Vanderghem, C., Delvigne, F., & Richel, A. (2015). Application of steam explosion as pretreatment on lignocellulosic material: A review. *Industrial & Engineering Chemistry Research, 54*(10), 2593-2598.

James, T. L. (1998). *Fundamentals on NMR*. Department of Pharmaceutical Chemistry: University of California.

Jedvert, K., Saltberg, A., Lindström, M. E., & Theliander, H. (2012). Mild steam explosion and chemical pre-treatment of Norway spruce. *Bioresources, 7*(2), 2051-2074.

Jedvert, K., Wang, Y., Saltberg, A., Henriksson, G., Lindström, M. E., & Theliander, H. (2012). Mild steam explosion: A way to activate wood for enzymatic treatment, chemical pulping and biorefinery processes. *Nordic Pulp & Paper Research Journal, 27*(5), 828-835.

Jönsson, L. J., & Martin, C. (2016). Pretreatment of lignocellulose: Formation of inhibitory by-products and strategies for minimizing their effects. *Bioresourse Technology, 199*, 103-112.

Kailemia, M. J., Ruhaak, L. R., Lebrilla, C. B., & Amster, I. J. (2014). Oligosaccharide analysis by mass spectrometry: A review of recent developments. *Analytical Chemistry, 86*(1), 196-212.

Kato, K., & Matsuda, K. (1969). Studies on the chemical structure of konjac mannan: Part I. Isolation and characterization of oligosaccharides from the partial acid hydrolyzate of the mannan. *Agricultural and Biological Chemistry, 33*, 1446-1453.

Koshijima, T., & Watanabe, T. (2003). *Association between Lignin and Carbohydrates in Wood and Other Plant Tissues*. Berlin: Springer.

Latha Gandla, M., Derba-Maceluch, M., Liu, X., Gerber, L., Master, E. R., Mellerowicz, E. J., et al. (2015). Expression of a fungal glucuronoyl esterase in *Populus*: effects on wood properties and saccharification efficiency. *Phytochemistry, 112*, 210-220.

Lehmann, L., Ronnest, N. P., Jorgensen, C. I., Olsson, L., Stocks, S. M., Jorgensen, H. S., et al. (2016). Linking hydrolysis performance to *Trichoderma reesei* cellulolytic enzyme profile. *Biotechnology and Bioengineering, 113*(5), 1001-1010.

Li, K. C., & Helm, R. F. (1995). Synthesis and rearrangement reactions of ester-linked lignin-carbohydrate model compounds. *Journal of Agricultural and Food Chemistry, 43*(8), 2098-2103.

Li, X. L., Špániková, S., De Vries, R. P., & Biely, P. (2007). Identification of genes encoding microbial glucuronoyl esterases. *Febs Letters, 581*(21), 4029-4035.

Lin, M. I., Hiyama, A., Kondo, K., Nagata, T., & Katahira, M. (2018). Classification of fungal glucuronoyl esterases (FGEs) and characterization of two new FGEs from *Ceriporiopsis subvermispora* and *Pleurotus eryngii*. *Applied Microbiology and Biotechnology, 102*(22), 9635-9645.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research, 42*(Database issue), D490-495.

Lorén, N., Hagman, J., Jonasson, J. K., Deschout, H., Bernin, D., Cella-Zanacchi, F., et al. (2015). Fluorescence recovery after photobleaching in material and life sciences: putting theory into practice. *Quarterly Reviews of Biophysics, 48*(3), 323-387.

Malgas, S., Van Dyk, S. J., & Pletschke, B. I. (2015). β-mannanase (Man26A) and α-galactosidase (Aga27A) synergism - a key factor for the hydrolysis of galactomannan substrates. *Enzyme and Microbial Technology, 70*, 1-8.

Michalak, L., Knutsen, S. H., Aarum, I., & Westereng, B. (2018). Effects of pH on steam explosion extraction of acetylated galactoglucomannan from Norway spruce. *Biotechnology for Biofuels, 11*, 311.

Miller, G. L. (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical Chemistry, 31*(3), 426-428.

Monrad, R. N., Eklöf, J., Krogh, K., & Biely, P. (2018). Glucuronoyl esterases: diversity, properties and biotechnological potential. A review. *Critical Reviews in Biotechnology, 38*(7), 1121-1136.

Moran-Mirabal, J. M., Bolewski, J. C., & Walker, L. P. (2011). Reversibility and binding kinetics of *Thermobifida fusca* cellulases studied through fluorescence recovery after photobleaching microscopy. *Biophysical Chemistry, 155*(1), 20-28.

Mosbech, C., Holck, J., Meyer, A. S., & Agger, J. W. (2018). The natural catalytic function of *Cu*GE glucuronoyl esterase in hydrolysis of genuine lignin-carbohydrate complexes from birch. *Biotechnology for Biofuels, 11*, 71.

Nakagame, S., Chandra, R. P., Kadla, J. F., & Saddler, J. N. (2011). The isolation, characterization and effect of lignin isolated from steam pretreated Douglas-fir on the enzymatic hydrolysis of cellulose. *Bioresource Technology, 102*(6), 4507-4517.

Nakagame, S., Chandra, R. P., & Saddler, J. N. (2010). The effect of isolated lignins, obtained from a range of pretreated lignocellulosic substrates, on enzymatic hydrolysis. *Biotechnology and Bioengineering, 105*(5), 871-879.

Nakamura, A. M., Nascimento, A. S., & Polikarpov, I. (2017). Structural diversity of carbohydrate esterases. *Biotechnology Research and Innovation, 1*, 35-51.

Naumoff, D. G. (2011). Hierarchical classification of glycoside hydrolases. *Biochemistry-Moscow, 76*(6), 622-635.

Nishimura, H., Kamiya, A., Nagata, T., Katahira, M., & Watanabe, T. (2018). Direct evidence for α ether linkage between lignin and carbohydrates in wood cell walls. *Scientific Reports, 8*, 6538.

Norgren, M., & Edlund, H. (2014). Lignin: Recent advances and emerging applications. *Current Opinion in Colloid & Interface Science, 19*(5), 409-416.

Nylander, F., Sunner, H., Olsson, L., Christakopoulos, P., & Westman, G. (2016). Synthesis and enzymatic hydrolysis of a diaryl benzyl ester model of a lignin-carbohydrate complex (LCC). *Holzforschung, 70*(5), 385-391.

Paës, G. (2014). Fluorescent probes for exploring plant cell wall deconstruction: a review. *Molecules, 19*(7), 9380-9402.

Paës, G. (2017). Bioinspired assemblies of plant cell walls for measuring protein-carbohydrate interactions by FRAP. *Methods in Molecular Biology, 1588*, 169-179.

Paës, G., Habrant, A., Ossemond, J., & Chabbert, B. (2017). Exploring accessibility of pretreated poplar cell walls by measuring dynamics of fluorescent probes. *Biotechnology for Biofuels, 10*, 15.

Paës, G., von Schantz, L., & Ohlin, M. (2015). Bioinspired assemblies of plant cell wall polymers unravel the affinity properties of carbohydrate-binding modules. *Soft Matter, 11*(33), 6586-6594.

Pauly, M., & Keegstra, K. (2008). Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant Journal, 54*(4), 559-568.

Persson, T. (2016). *Den svenska bioekonomins utveckling - statistik och analys*. Retrieved from Östersund, Sweden: Myndigheten för tillväxtpolitiska utvärderingar och analyser

Pielhop, T., Amgarten, J., Von Rohr, P. R., & Studer, M. H. (2016). Steam explosion pretreatment of softwood: the effect of the explosive decompression on enzymatic digestibility. *Biotechnology for Biofuels, 9*, 152.

Pitt, J. J. (2009). Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clinical Biochemist Reviews, 30*(1), 19-34.

Pokkuluri, P. R., Duke, N. E., Wood, S. J., Cotta, M. A., Li, X. L., Biely, P., et al. (2011). Structure of the catalytic domain of glucuronoyl esterase Cip2 from *Hypocrea jecorina*. *Proteins, 79*(8), 2588-2592.

Ralph, J., & Landucci, L. (2010). NMR of lignins. In C. Heitner, D. Dimmel, & J. A. Schmidt (Eds.), *Lignin and Lignans: Advances in Chemistry* (pp. 137-243): CRC Press.

Sanderson, K. (2011). Lignocellulose: A chewy problem. *Nature, 474*(7352), S12-14.

Scheller, H. V., & Ulvskov, P. (2010). Hemicelluloses. *Annual Review of Plant Biology, 61*, 263-289.

Seidel, C. M., Pielhop, T., Studer, M. H., & Rudolf von Rohr, P. (2017). The influence of the explosive decompression in steam-explosion pretreatment on the enzymatic digestibility of different biomasses. *Faraday Discussions, 202*, 269-280.

Singh, A. P., & Singh, T. (2014). Biotechnological applications of wood-rotting fungi: A review. *Biomass & Bioenergy, 62*, 198-206.

Sjöström, E. (1993). Wood polysaccharides. In *Wood chemistry- Fundamentals and Applications*. San Diego: Academic Press.

Songsiriritthigul, C., Buranabanyat, B., Haltrich, D., & Yamabhai, M. (2010). Efficient recombinant expression and secretion of a thermostable GH26 mannan endo-1,4-β-mannosidase from *Bacillus licheniformis* in *Escherichia coli*. *Microbial Cell Factories, 9*, 20.

Spangenberg, B., Poole, C. F., & Weins, C. (2010). Theoretical basis of thin layer chromatography (TLC). In *Quantitative Thin-Layer Chromatography*. Heidelberg: Springer.

Špániková, S., & Biely, P. (2006). Glucuronoyl esterase - novel carbohydrate esterase produced by *Schizophyllum commune*. *Febs Letters, 580*(19), 4597-4601.

Špániková, S., Polaková, M., Joniak, D., Hirsch, J., & Biely, P. (2007). Synthetic esters recognized by glucuronoyl esterase from *Schizophyllum commune*. *Archives of Microbiology, 188*(2), 185-189.

Srivastava, P. K., & Kapoor, M. (2017). Production, properties, and applications of endo-β-mannanases. *Biotechnology advances, 35*(1), 1-19.

Sunner, H., Charavgi, M. D., Olsson, L., Topakas, E., & Christakopoulos, P. (2015). Glucuronoyl esterase screening and characterization assays utilizing commercially available benzyl glucuronic acid ester. *Molecules, 20*(10), 17807-17817.

Tailford, L. E., Ducros, V. M. A., Flint, J. E., Roberts, S. M., Morland, C., Zechel, D. L., et al. (2009). Understanding how diverse β-mannanases recognize heterogeneous substrates. *Biochemistry, 48*(29), 7009-7018.

Teeri, T., & Henriksson, G. (2009). Enzymes degrading wood components. In M. Ek, G. Gellerstedt, & G. Henriksson (Eds.), *Wood Chemistry and Wood Biotechnology*. Berlin: Walter de Gruyter.

Telemann, A. (2009). Hemicelluloses and pectin. In M. Ek, G. Gellerstedt, & G. Henriksson (Eds.), *Wood Chemistry and Wood Biotechnology*. Berlin: Walter de Gruyter.

Tenkanen, M., Makkonen, M., Perttula, M., Viikari, L., & Teleman, A. (1997). Action of *Trichoderma reesei* mannanase on galactoglucomannan in pine kraft pulp. *Journal of Biotechnology, 57*(1-3), 191-204.

Tenkanen, M., Thornton, J., & Viikari, L. (1995). An acetylglucomannan esterase of *Aspergillus oryzae;* purification, characterization and role in the hydrolysis of *O*-acetyl-galactoglucomannan. *Journal of Biotechnology, 42*(3), 197-206.

Terrapon, N., Lombard, V., Drula, E., Lapebie, P., Al-Masaudi, S., Gilbert, H. J., et al. (2018). PULDB: the expanded database of Polysaccharide Utilization Loci. *Nucleic Acids Research, 46*(D1), D677-D683.

Topakas, E., Moukouli, M., Dimarogona, M., Vafiadi, C., & Christakopoulos, P. (2010). Functional expression of a thermophilic glucuronoyl esterase from *Sporotrichum thermophile*: identification of the nucleophilic serine. *Applied Microbiology and Biotechnology, 87*(5), 1765-1772.

Tsai, A. Y., Canam, T., Gorzsas, A., Mellerowicz, E. J., Campbell, M. M., & Master, E. R. (2012). Constitutive expression of a fungal glucuronoyl esterase in *Arabidopsis* reveals

altered cell wall composition and structure. *Plant Biotechnology Journal, 10*(9), 1077-1087.

Urbanowicz, B. R., Pena, M. J., Ratnaparkhe, S., Avci, U., Backe, J., Steet, H. F., et al. (2012). 4-*O*-methylation of glucuronic acid in *Arabidopsis* glucuronoxylan is catalyzed by a domain of unknown function family 579 protein. *Proceedings of the National Academy of Sciences of the United States of America, 109*(35), 14253-14258.

Vafiadi, C., Topakas, E., Biely, P., & Christakopoulos, P. (2009). Purification, characterization and mass spectrometric sequencing of a thermophilic glucuronoyl esterase from *Sporotrichum thermophile*. *Fems Microbiology Letters, 296*(2), 178-184.

Van Dyk, J. S., & Pletschke, B. I. (2012). A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes-factors affecting enzymes, conversion and synergy. *Biotechnology advances, 30*(6), 1458-1480.

Van Zyl, W. H., Rose, S. H., Trollope, K., & Gorgens, J. F. (2010). Fungal β-mannanases: Mannan hydrolysis, heterologous production and biotechnological applications. *Process Biochemistry, 45*(8), 1203-1213.

Varnai, A., Siika-Aho, M., & Viikari, L. (2013). Carbohydrate-binding modules (CBMs) revisited: reduced amount of water counterbalances the need for CBMs. *Biotechnology for Biofuels, 6*(1), 30.

Von Freiesleben, P., Spodsberg, N., Blicher, T. H., Anderson, L., Jorgensen, H., Stålbrand, H., et al. (2016). An *Aspergillus nidulans* GH26 endo-β-mannanase with a novel degradation pattern on highly substituted galactomannans. *Enzyme and Microbial Technology, 83*, 68-77.

Von Freiesleben, P., Spodsberg, N., Stenbaek, A., Stålbrand, H., Krogh, K., & Meyer, A. S. (2018). Boosting of enzymatic softwood saccharification by fungal GH5 and GH26 endomannanases. *Biotechnology for Biofuels, 11*, 194.

Voragen, A. G. J., Coenen, G. J., Verhoef, R. P., & Schols, H. A. (2009). Pectin, a versatile polysaccharide present in plant cell walls. *Structural Chemistry, 20*(2), 263-275.

Wang, H., Luo, H., Li, J., Bai, Y., Huang, H., Shi, P., et al. (2010). An α-galactosidase from an acidophilic *Bispora* sp. MEY-1 strain acts synergistically with β-mannanase. *Bioresourse Technology, 101*(21), 8376-8382.

Wang, P., Fan, J. Y., & Xie, Y. M. (2013). Synthesis and characterization of pectin-dehydrogenation polymer complex by isotopic labeling method. *Cellulose Chemistry and Technology, 47*(5-6), 401-408.

Watanabe, T. (1995). Important properties of lignin-carbohydrates complexes (LCCs) in environmentally safe paper making. *Trends in Glycoscience and Glycotechnology, 7*(33), 57-68.

Willför, S., Sundberg, K., Tenkanen, M., & Holmbom, B. (2008). Spruce-derived mannans-A potential raw material for hydrocolloids and novel advanced natural materials. *Carbohydrate Polymers, 72*, 197-210.

Wong, D. W., Chan, V. J., McCormack, A. A., Hirsch, J., & Biely, P. (2012). Functional cloning and expression of the *Schizophyllum commune* glucuronoyl esterase gene and characterization of the recombinant enzyme. *Biotechnology Research International, 2012*, 951267.

Wu, M. M., Chang, K., Gregg, D. J., Boussaid, A., Beatson, R. P., & Saddler, J. N. (1999). Optimization of steam explosion to enhance hemicellulose recovery and enzymatic hydrolysis of cellulose in softwoods. *Applied Biochemistry and Biotechnology, 77*(1-3), 47-54.

Xie, Y. M., Yasuda, S., Wu, H., & Liu, H. B. (2000). Analysis of the structure of lignin - carbohydrate complexes by the specific C-13 tracer method. *Journal of Wood Science, 46*(2), 130-136.

Xu, C. L., Leppanen, A. S., Eklund, P., Holmlund, P., Sjoholm, R., Sundberg, K., et al. (2010). Acetylation and characterization of spruce (*Picea abies*) galactoglucomannans. *Carbohydrate Research, 345*(6), 810-816.

Yang, J. C., Madupu, R., Durkin, A. S., Ekborg, N. A., Pedamallu, C. S., Hostetler, J. B., et al. (2009). The complete genome of *Teredinibacter turnerae* T7901: An intracellular

endosymbiont of marine wood-boring bivalves (shipworms). *PLoS One, 4*(7), e6085.

Yuan, T. Q., Sun, S. N., Xu, F., & Sun, R. C. (2011). Characterization of lignin structures and lignin-carbohydrate complex (LCC) linkages by quantitative C-13 and 2D HSQC NMR spectroscopy. *Journal of Agricultural and Food Chemistry, 59*(19), 10604-10614.

Yuan, Y., Teng, Q., Lee, C., Zhong, R., & Ye, Z. H. (2014). Modification of the degree of 4-*O*-methylation of secondary wall glucuronoxylan. *Plant Science, 219-220*, 42-50.

Zhang, X., Wang, S., Wu, X., Liu, S., Li, D., Xu, H., et al. (2015). Subsite-specific contributions of different aromatic residues in the active site architecture of glycoside hydrolase family 12. *Scientific Reports, 5*, 18357.

Zhang, Y., Li, J., Lindström, M. E., Stepan, A., & Gatenholm, P. (2013). Spruce glucomannan: Preparation, structural characteristics and basic film forming ability. *Nordic Pulp & Paper Research Journal, 28*(3), 323-330.

Zhuang, X. S., Wang, W., Yu, Q., Qi, W., Wang, Q., Tan, X. S., et al. (2016). Liquid hot water pretreatment of lignocellulosic biomass for bioethanol production accompanying with high valuable products. *Bioresource Technology, 199*, 68-75.