Thesis for the degree of Doctor of Philosophy

Enzyme assays for lignin-carbohydrate bond hydrolysis

Travellable roads and *cul-de-sacs*

Hampus Sunner

Department of Biology and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2016

Enzyme assays for lignin-carbohydrate bond hydrolysis
Travellable roads and *cul-de-sacs*
Hampus Sunner
ISBN 978-91-7597-389-0

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 4070
ISSN 0346-718X

Department of Biology and Biological Engineering
Chalmers University of Technology
SE–412 96 Göteborg, Sweden
Telephone + 46 (0) 31 – 772 1000

Cover: The three components of an enzyme assay, interacting within a chemical context, see Figure 2.1 on page 7.

Typeset by the author using LaTeX.

Printed by Chalmers Reproservice
Göteborg, Sweden 2016

Strunt är strunt och snus är snus
om ock i gyllne dosor
Men rosor i ett sprucket krus
är ändå alltid rosor

Gustaf Fröding
Nya dikter

Abstract

To identify, produce, and use enzymes, analytical methods known as enzyme assays are employed. Enzyme assays are based on analysing the changes brought about on a substrate by an enzyme under defined conditions. Assays are often based on simplified reactions acting as substitutes for the reaction of interest. In practice, this means that the ability to discover and use enzymes as biocatalysts is determined by the availability and applicability of such simplified reaction systems.

The first part of biomass conversion is the degradation of lignified plant matter and the main bottleneck of this step is the non-destructive disassociation of polymeric biomass components. Some of the degradation recalcitrance is believed to be due to covalent bonds between the lignin and sugar components of the material (LC-bonds). Thus, enzymatic hydrolysis of these bonds can potentially improve component separation. So far, only a few enzymes capable of degrading LC-bonds have been identified. The low number may be due to the lack of enzyme assays for discovery and characterization.

The purpose of this research effort has been to design assays for enzymes capable of breaking LC-bonds. The published works associated with this thesis describe various assay methods relevant to this goal: Paper I defines procedures for generating LC-bond-rich substrates from natural sources (lignin-carbohydrate complexes), with the aim of demonstrating their presence and detection by size-exclusion chromatography. Papers II and III describe synthetic-substrate assays for glucuronoyl esterases (GEs), the enzyme class with the best evidence of LC-bond hydrolysis. Paper II includes the synthesis of a β-diaryl ether for use as a GE assay substrate and Paper III presents and discusses several assays with different detection methods based on a commercially available GE substrate. Paper IV presents assays for enzyme synergy and shows how mass spectrometry can be used as an auxiliary detection method to better understand enzyme activities.

This thesis places the enclosed articles into the overall context of LC-bond assays and describes possibilities for the combination of substrates, enzyme activities, and detection methods for the construction of novel LC-bond assays. As such, this work should offer background and a starting point for anyone wishing to do practical work on enzymatic LC-bond hydrolysis.

Keywords: Lignin-carbohydrate bonds, Enzyme assays, Enzyme kinetic parameter estimation, Synergistic enzyme assay, Glucuronoyl esterase

Preface

This PhD dissertation partially fulfils the requirements for a PhD degree at the Department of Biology and Biological Engineering, Chalmers University of Technology, Sweden. The PhD project ran through 2009–2016 as part of the Wallenberg Wood Science Center (WWSC), a joint research centre of Chalmers University of Technology and KTH Royal Institute of Technology, financed by Knut and Alice Wallenberg Foundation. The main part of the work on enzyme assays for glucuronoyl esterases was carried out at Chalmers University of Technology, in the Industrial Biotechnology division of the Department of Biology and Biological Engineering, under the supervision of Professor Lisbeth Olsson. Parts of the research concerning the enzymatic degradation of xylan, was performed at KTH Royal Institute och Technology, in collaboration with Dr Lauren McKee and Assistant Professor Francisco Vilaplana.

List of publications

This thesis is based on the following papers:

Paper I. N Westerberg, H Sunner, M Helander, G Henriksson, M Lawoko, and A Rasmuson (2012). Separation of galactoglucomannans, lignin, and lignin-carbohydrate complexes from hot-water-extracted Norway spruce by cross-flow filtration and adsorption chromatography. *BioResources* 7.4, pp. 4501–4516

Paper II. F Nylander, H Sunner, L Olsson, P Christakopoulos, and G Westman (2016). Synthesis and enzymatic hydrolysis of a diaryl benzyl ester model of a lignin-carbohydrate complex (LCC). *Holzforschung* 70.5, pp. 385–391

Paper III. H Sunner, MD Charavgi, L Olsson, E Topakas, and P Christakopoulos (2015). Glucuronoyl Esterase Screening and Characterization Assays Utilizing Commercially Available Benzyl Glucuronic Acid Ester. *Molecules* 20.10, pp. 17807–17817

Paper IV. LS McKee, H Sunner, GE Anasontzis, G Toriz, P Gatenholm, V Bulone, F Vilaplana, and L Olsson (2016). A GH115 α-glucuronidase from Schizophyllum commune contributes to the synergistic enzymatic deconstruction of softwood glucuronoarabinoxylan. *Biotechnology for biofuels* 9.2

Other relevant publications co-authored by Hampus Sunner

R Olivares-Hernández, H Sunner, JC Frisvad, L Olsson, J Nielsen, and G Panagiotou (2010). Combining substrate specificity analysis with support vector classifiers reveals feruloyl esterase as a phylogenetically informative protein group. *PLOS ONE* 5.9, e12781

Author's contributions

Paper I. Second author. Planned and carried out the enzymatic treatment and subsequent analysis together with the main author. Discussed and interpreted the data and contributed to the writing of the manuscript.

Paper II. First author (shared). Conceived the study. Planned, developed methods for and performed the biochemical work, as well as HPLC and FTIR analyses. Analysed, plotted and interpreted the biochemical data. Discussed all results and wrote the manuscript with the other main author.

Paper III. First author. Designed, validated, planned and performed the spectrophotometric assay and its implementation as a screening assay. Developed the TLC assay. Did the alkaline hydrolysis rate experiments. Constructed and programmed the analysis methods and analysed all data. Discussed the results. Wrote the manuscript and the accompanying methods.

Paper IV. Second author. Planned and performed end-point and rate experiments. Planned analysis and data analysis. Discussed and interpreted the data and contributed to the writing of the manuscript.

Contents

Part I

Introductory chapters

Chapter 1

Introduction

This chapter introduces the concept of the enzyme assay, justifies the need for lignin-carbohydrate (LC)-bond assays, gives an overview of the thesis, and explains its structure.

The central concept of the enzyme assay

The fundamental operation of any practical enzyme work is the *enzyme assay*. The simplest enzyme assay is based on a single catalysed reaction and consists of three components: an *enzyme*, a *substrate*, and a *detection method* (Figure 1.1 and Chapter 2). The role of the enzyme is to catalyse the conversion of the substrate and that of the detection method is to study this conversion (Bisswanger, 2012b). This basic system can be modified in endless ways to suit different requirements (Bisswanger, 2012a).

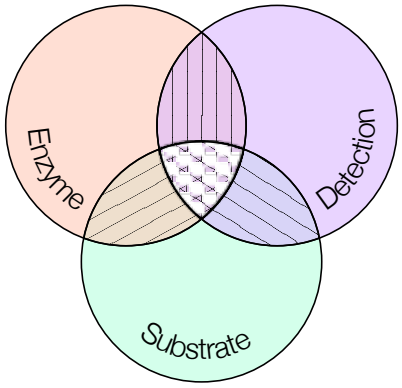


Figure 1.1: The principal components of an enzyme assay are depicted as a Venn diagram that illustrates their interactions. The enzyme activity targeted by the assay should fall inside the central overlap region.

Lignin-carbohydrate bond degrading enzymes

Lignin is the second most abundant biopolymer. Together with carbohydrate polymers, it makes up most of the dry mass of plant cell walls (see Chapter 3 and Henriksson et al. (2009)).

Lignin consists of aromatic monomers that are polymerized by radical coupling into what may appear as a stochastic combination of monomers (Henriksson, 2009; Brunow and Lundquist, 2010; Dimmel, 2010; Leisola et al., 2012). Lignin formation is nonetheless controlled by the organism into purposeful structures with a level of organization that is only superficially understood (Henriksson, 2009). In contrast, the main cell wall polysaccharides – the homopolymer cellulose and various heteropolymers referred to as hemicelluloses – have a more ordered and better understood structure (Teleman, 2009). The lignin in the cell wall is covalently linked to the hemicelluloses by ether, ester or hemiacetal bonds (Jeffries, 1990; Brunow and

Lundquist, 2010). These bonds are collectively referred to as lignin-carbohydrate bonds (LC-bonds).

To enable the industrial use of plant cell walls (*lignocellulosics*; see Section 3.1) as raw materials for the production of biomaterials and biofuels, the cell wall components must be separated from each other, *i.e.* the lignin must be separated from the cellulose and from the different kinds of hemicelluloses. In nature, the organisms degrading lignocellulose and its residual products possess enzymes to aid this process (Cragg et al., 2015) and it is thought that some of these enzymes are capable of breaking LC-bonds (Jeffries, 1990). Some may be used to improve industrial processes that separate lignocellulosic components. To *discover*, *identify*, *characterize* and *produce* such enzymes (the steps required from *discovery to application*) appropriate enzyme assays are required (Bisswanger, 2012b). The assays must suit each enzyme–substrate combination and match the requirements of each step (an overview of the steps and a discussion of their requirements is found in Chapter 2). However, such assays have been described only for the two types of LC-bonds for which hydrolysing enzymes are known (Chapter 4).

Aims

The scientific aim of this thesis is to present and discuss the design of the different enzyme assays that could be used to discover, identify, characterize, and produce LC-bond-degrading enzymes. A second aim is to document the work I have done during my PhD studies and to discuss it in the context of my scientific aim. While Part I of this thesis is concerned with these aims in general, the published articles of Part II address the following specific aspects:

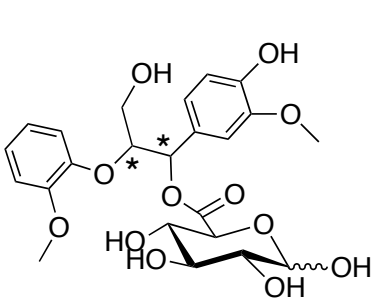
Paper I. Production of a LC-bond-rich assay substrate from a natural source. Evaluation of a previously used assay for the detection of LC-bonds in a complex substrate (Lawoko et al., 2006). Assessment of size-exclusion chromatography as a detection method for LC-bond assays.

Paper II. Synthesis of and construction of an enzyme assay based on the synthetic guaiacylglycerol-β-guaiacyl glucuronate LC-bond model (β-O-4′ glucuronate; compound **1**; Figure 1.2). Demonstration of the model's hydrolysability using glucuronoyl esterases (GEs).

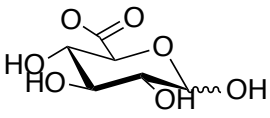
Paper III. Design and evaluation of several GE assays based on the commercially available substrate benzyl D-glucuronate (BnGlcA; compound **2**; Figure 1.2). Enzyme assays with different detection methods and applications are presented.

Paper IV. Design of various enzyme assays of the synergistic type (Section 7.4.1) using a xylan substrate (Section 3.7.1). Evaluation of the combination of ion chromatography and mass spectrometry (IC and MS, in sections 5.2.2 and 5.3, respectively) for quantitative and structural analysis of the assay products.

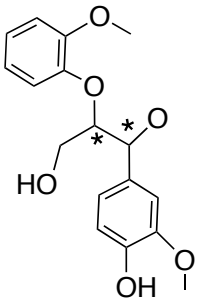
The study focuses on enzymes that would be specific to LC-bonds and not affect other parts of the lignocellulosic structure. Thus, non-specific lignin-modifying



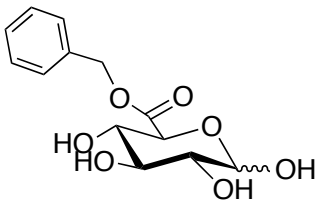
1 β-O-4' glucuronate guaiacylglycerol-β-guaiacyl ether D-glucuronate



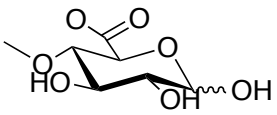
3 GlcA D-glucuronic acid



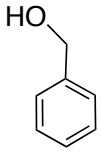
5 β-O-4' guaiacylglycerol-β-guaiacyl ether



2 BnGlcA benzyl D-glucuronate



4 mGlcA 4-O-methyl-D-glucuronic acid



6 BnOH benzyl alcohol

Figure 1.2: The key compounds of this thesis. The synthetic substrates **1** and **2** are esters of GlcA (**3**), which occurs as as mGlcA (**4**) in the natural structure that **1** and **2** are models of. The aryl part of the model is constituted either of β-O-4′ or BnOH (**5** and **6**, respectively). The table presents the numbers and abbreviated names that will be used to identify these compounds throughout this thesis, as well as common names for the compounds.

enzymes have not been considered in my work. They are not included in this thesis, except for a brief summary in Chapter 4, where the enzymes relevant to LC-bond scission are discussed.

Content and structure of this thesis

The next chapter reviews different enzyme assays, providing a conceptual framework for the discussion. Each assay component will be introduced separately, detailing the substrates (Chapter 3), enzymes (Chapter 4), and detection methods (Chapter 5) that have been contemplated, investigated or experimented upon during the course of my work for this thesis. Chapter 5 also summarizes some unsuccessful attempts at GE assay development that is not included in the appended papers in order to exemplify applications of the detection methods and to provide opportunities for

later discussion.

The following chapters offer an analysis of the work that was done as part of this thesis, with a focus on how the intricacies of the assay components and the interaction between them affect the properties and applicability of the assay. To this end, Chapter 6 summarizes my work in the development of GE assays, (including some dead ends – *cul-de-sacs* – not included in the papers of Part II) and critically discusses the results in terms of challenges and opportunities.

Chapter 7 starts by commenting on the current situation for LC-bond assays and how these interact to create a complex inter-relationship that needs to be taken into consideration when designing the assays. This is followed by a discussion about LC-bond assay substrates. The remainder of the chapter is devoted to delivering both general and concrete suggestions for the design of assays for LC-bond degrading enzymes.

The Conclusion, Chapter 8, summarizes my thesis work and the final chapter of Part I puts the work in perspective, suggesting potentially profitable paths for immediate or long-term exploitation.

In the text, chemical structures are mainly referred to as numbers. The corresponding structures are found in the Appendix (*pp.* 149) and Part II lists the peer-reviewed publications included in this thesis.

Chapter 2

Enzyme assays

In order to facilitate a high-level discussion in the coming chapters, here I define the relevant terminology and the conceptual framework for the abstraction and practical classification of enzyme assays.

The *enzyme assay* is based on i) a reactant that undergoes an ii) enzyme-catalysed reaction, which can be iii) qualitatively or quantitatively detected. In practice, this means that three components are required: i) a *substrate*, ii) an *enzyme* and iii) a *detection method*. In addition, the assay takes place in iv) a *chemical context*, usually in solution and in the presence of additional molecules. Figure 2.1 shows these four components and illustrates their interactions.(Bisswanger, 2012b)

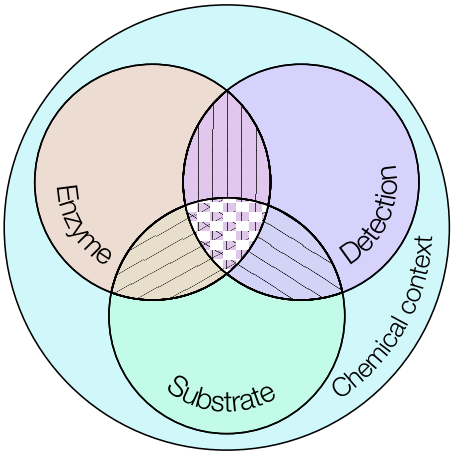


Figure 2.1: The principal components of an enzyme assay depicted as a Venn diagram to illustrates their interactions. The enzyme activity targeted by the assay should fall inside the central overlap region.

In the context of this thesis, an enzyme assay will be regarded as a chemical *system* of potentially interacting elements, which includes at least one enzyme-catalysed reaction and all four components listed above. Furthermore, the purpose of the (enzyme assay) system would be to study one or more *system properties*, such as the amount of enzyme of the interactions between system components (illustrated in Figure 2.2; Marangoni (2002b)).

The property we want to measure in an enzyme system (the *target property*) is almost always some property of the enzyme-catalysed reaction, usually the *enzyme activity* (Bisswanger, 2012b). The target property of an enzyme assay can rarely be quantified directly, but has to be inferred by applying experimental values (or *metrics*) to the *model* we have of the system. To this end, enzyme activities are usually determined by measuring the change in concentration of one or more reactants over time (the metrics) and correlating this with the chemical equation of the enzyme-catalysed reaction (the model) (Marangoni, 2002b).

As target properties have to be inferred, their *validity* depends on that of the model on which they are based. In turn, the validity of the model, and the conclusions it entails, depends on that of the model's *assumptions* (Parkin, 2002). For simple systems with few assumptions or for trivial hypotheses (such as "pure enzyme E can break bond A in a defined substrate S"), a positive result in an assay that employs routine controls may implicitly validate the model. However, in the less trivial case, conscious validation of the model is necessary in order to ascertain the validity of the inferred target properties (or, in everyday terms, validate the assays results). Practically, validation is done by performing experiments (or relating to previous ones) that explicitly test the model's assumptions and by critically examining data while considering alternative hypotheses (Bisswanger, 2014).

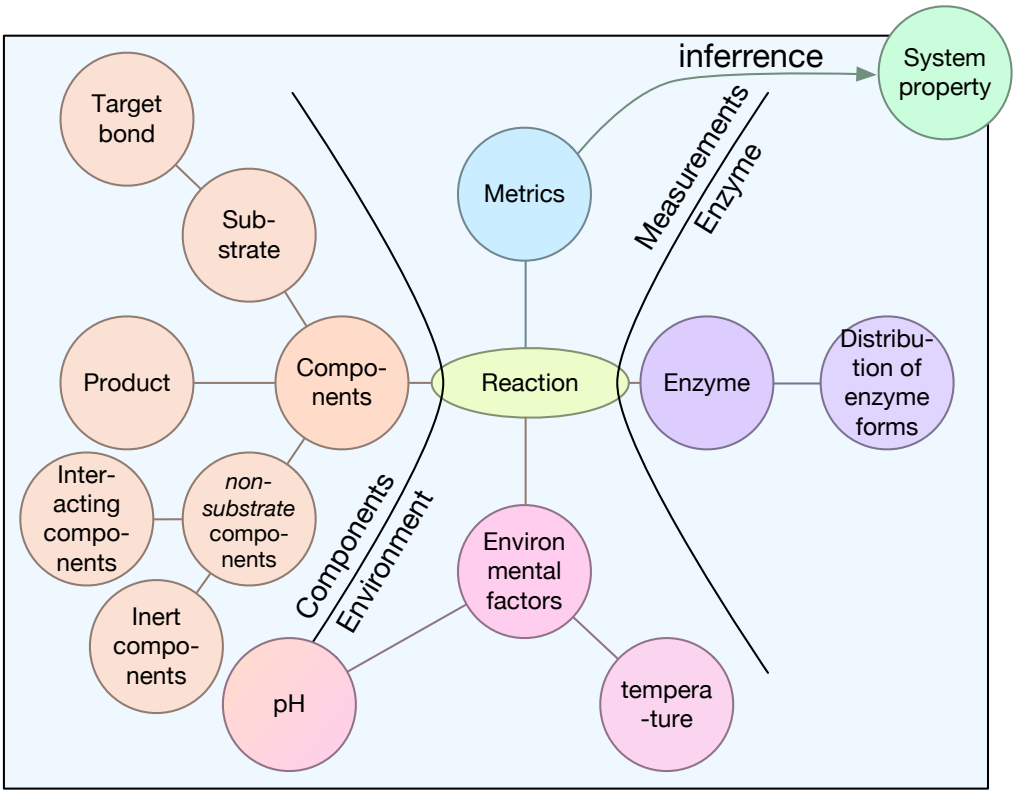


Figure 2.2: Model of an enzyme assay system. The enzyme assay consists of a number of components that may and may not interact with each other, with the enzyme and with the environment. The enzyme catalyses the reaction that is the basis of the assay, and measurements are made to quantify the reaction. From these, some property of the system can be inferred, based on the model we have of that system.

2.1 Enzyme assay simplifications

In many situations, the system for which we want to have information (here referred to as the *prototype*, as in the sense of ancestral) is not available or is impractical to measure on. In these cases, we can construct a concrete model of the prototype system (here referred to as a *proxy*), and subject it to our measurements. The proxy

is always a concrete and assayable system, while the prototype is not necessarily either of the two. Usually, the term *enzyme assay* is used to refer to this type proxy system.

The key simplifications offered by the proxy system are i) the usage of idealized substrates, ii) substrates allowing for simplified detection methods, and iii) simplified detection methods. However, with each simplification, it becomes less certain that prototype properties can be validly inferred from measurements on the proxy system. Therefore the prototype-proxy equivalence should be considered when designing enzyme assays and evaluating their results and relevance.

2.2 Target properties

The target properties for an enzyme assay can be broadly classified into one of the following categories, corresponding to measurements of the components and interactions of Figure 2.1:

Amount of enzyme This follows the principle that, generally, the amount of detected enzyme activity is proportional to the amount of enzyme. It includes determination of the relative amount of enzyme in each sample in a series and rate of enzyme inactivation due to environment factors, such as temperature, pH, species and solvents present in the assay system (Bisswanger, 2012b).

Enzyme–Substrate interactions These define the effect of substrate properties on enzyme activity (Parkin, 2002).

Enzyme–Environment interactions These determine the general impact of temperature and chemical context on the rate of catalysis. Additionally, they help determine enzymatic mechanisms upon addition of specific agents (*eg.*, chelators, selective inhibitors) or compatibility with agents that are relevant for practical applications (Marangoni, 2002a).

Substrate properties These include qualitative or quantitative analysis of the substrate itself.

Detection properties Investigating detection limits and linear ranges as well as selectivity and sensitivity and a main part of the development of natural-substrate assays (Section 6.1.1)

Complex interaction effects These arise when multiple or complex substrates are used or when multiple enzymes are applied. It also includes assays modelled after real-world applications and other situations where compound effects are determinable but individual interactions are not. This includes the synergy assays that are commonly used to measure hydrolysis of a complex carbohydrate substrate by a combination of enzymes to investigate the *eg.* the contribution of individual enzymes. Relevant examples of assays targeting complex interaction effects include Paper IV and d'Errico et al. (2016).

Chapter 3

Substrates

This chapter outlines the structure of wood, focusing on how enzyme assay substrates can be made to mimic the LC-bond structures found in the material

The substrate defines the chemo-, stereo- and regioselectivities of an assay's enzyme (Larsson, 2012). The choice of substrate also dictates which detection methods are feasible and, to a varying degree, it defines the chemical context of the assay (Bisswanger, 2014). All of this impacts on the conclusions that can be drawn. Consequently, the choice of substrate is the central and most significant one in the design of an LC-bond.

To create enzyme assays that are valid representations of (*i.e.* proxies for) enzymatic action on lignocellulosic substrates, we must first try to understand the structural features of those materials. This chapter starts by outlining the components of lignocellulose and their interactions. Subsequently, the suggested covalent LC-bonds, their relevance to our treatise, and the synthetic substrates that could represent them are discussed in separate sections. Finally, the concept of naturally-derived substrates is discussed and the plant preparations that were used in the course of this work are presented.

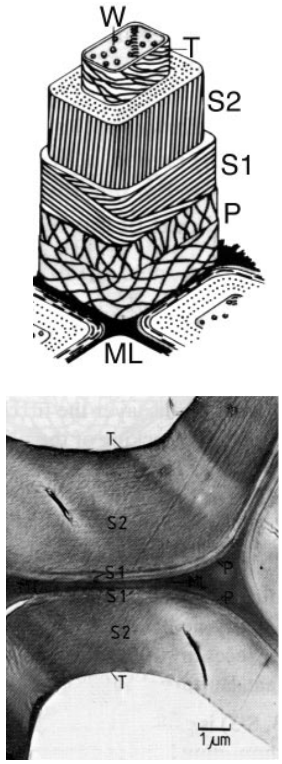


Figure 3.1: (Top) Model of plant fibre (plant cell) showing the layers of the cell wall (P–T) and the middle lamella (ML). (Bottom) Transmission electron micrograph of ultra thin cell wall section, showing the same cell wall layers (Ralph et al., 1996).

3.1 Overview of lignocellulosic structure

A generalized cell wall of a lignified plant tissue comprises three types of components: *cellulose*, *hemicelluloses* and *lignin*[1]. During cell growth and maturation, the cell wall

[1]Pectic substances and many other components, reside in the cell wall. The three mentioned here are the most abundant and the ones relevant to our discourse.

components are deposited, in distinct layers, from the outside of the cell and inwards (Figure 3.1; Henriksson et al. (2009) and Gibson (2012)).

Material containing these three types of components is called "lignocellulose", or "lignocellulosic material". This type of structure is complex and varies with genome, tissue, environmental factors, developmental stage, and cellular location (*i.e.* specific cell wall layers may show different organization; Henriksson et al. (2009)). Here, I will focus on what is relevant for LC-bond degradation of wood, particularly from coniferous trees (*softwood*).

Cellulose consists of long parallel chains of β-1,4-linked glucose, tightly packed into highly crystalline *microfibrils* of a few dozen parallel chains. Cellulose microfibrills form bundles called *fibrils*, that polymerize around the cell, similarly to the threads of a butterfly's cocoon, but with defined orientations in the cell wall secondary layers (S1 and S2 in Figure 3.1; Thibaut et al. (2001) and Henriksson and Lennholm (2009)).

Hemicellulose is a loosely defined group of branched hetero– and homo-polysaccharides that form a loose network between microfibrils and cellulose fibrils (Timell, 1967; Teleman, 2009). Two hemicelluloses relevant to LC-bond degradation are arabinoglucuronoxylan and galactoglucomannan (AGX and GGM, respectively; see Section 4.6; Timell (1967), Jeffries (1990), Lawoko et al. (2006), Moreira and Filho (2008), and Teleman (2009)).

Lignin, is the second most abundant component. It consists of a large number of closely related aromatic monomers that are polymerized *in situ*, creating a semi-random network(Henriksson, 2009; Dimmel, 2010; Leisola et al., 2012). Lignin confers hydrophobicity, rigidity, and resistance to microbial degradation (Henriksson, 2009; Dimmel, 2010; KE Eriksson, 2010; Gibson, 2012).

Lignin fills the space between the cell walls of adjacent cells, known as the middle lamella (ML in Figure 3.1), "glueing" the cells together (Henriksson, 2009; Gibson, 2012). In addition, the middle lamella is rich in *pectin* (Yapo, 2011), another heterogeneous group of generally water-soluble polysaccharides that are important for plants, but of lesser interest for industrial lignocellulosic utilization and therefore disregarded in the present thesis work.

Figure 3.2 shows an artist's rendition of the arrangement of cell wall components. The same network of wood components in a plane between two cellulose fibrils is depicted in Figure 3.3 in the form of structural formulae.

3.2 Interactions between components

Lignin and hemicellulose are cross-linked by *lignin-carbohydrate bonds* (*LC-bonds*; reviewed in *eg.* Jeffries (1990), Koshijima and Watanabe (2003), and Brunow and Lundquist (2010)). As the cell wall lignifies, most of the water is expelled and replaced by lignin, creating a solid matrix of hydrophobic interactions (Henriksson, 2009; Brunow and Lundquist, 2010). LC-bonds are formed during lignin polymerization (Adler, 1977; Brunow and Lundquist, 2010). Given that the cell dies during lignification, the interactions between cell wall components – covalent inter-molecular cross-links and extensive hydrophobic bonds – must maintain the integrity of the cell wall for the remainder of the plant's life (Thibaut et al., 2001; Gibson, 2012).

While cell wall interactions may be crucial for the growing plant, as well as for human exploitation (Thibaut et al., 2001; Hubbe and Lucia, 2007), the hydrophobic and covalent interactions (particularly intermolecular cross-links) complicate the industrial separation of single polymers (Otero et al., 2007; Leisola et al., 2012).

The next sections will deal with the different types of LC-bonds. These bonds are depicted in Figure 3.4 as fragments (show in context in Figure 3.3) that in a size that would be suitable for use as assay substrates. These could, in principle be

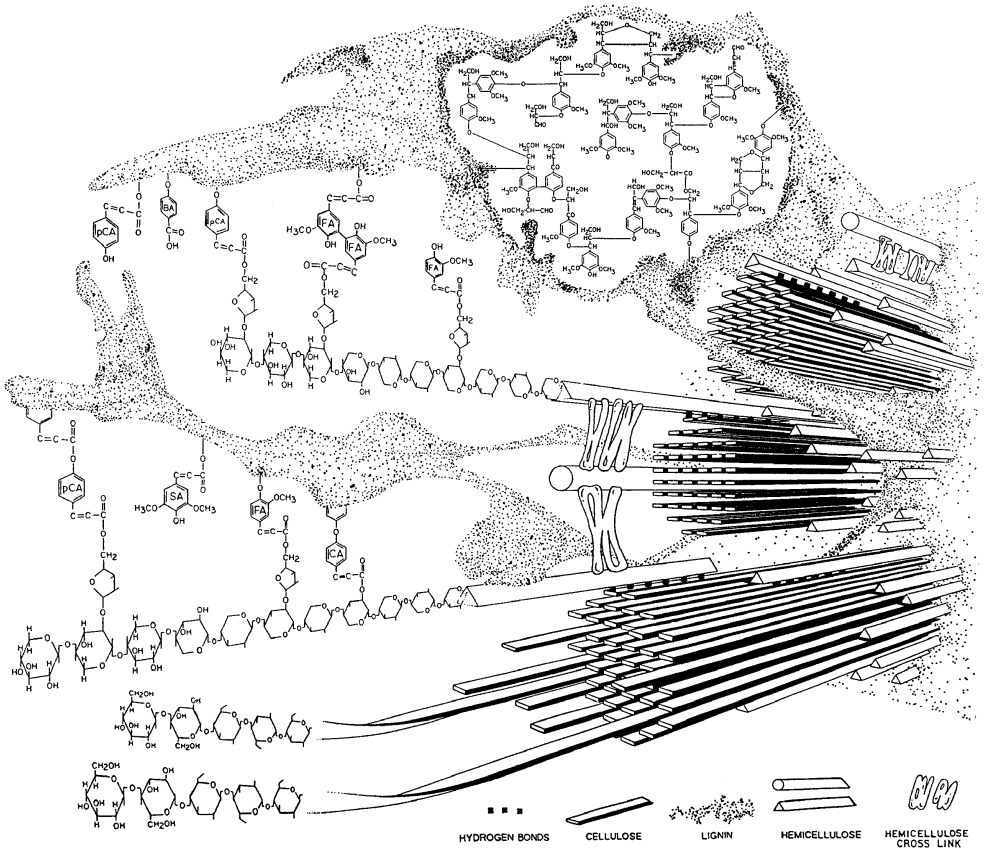


Figure 3.2: Illustration of the arrangement of cell wall components around three cellulose microfibrils in a generalized secondary cell wall of a grass (Bidlack et al., 1992).

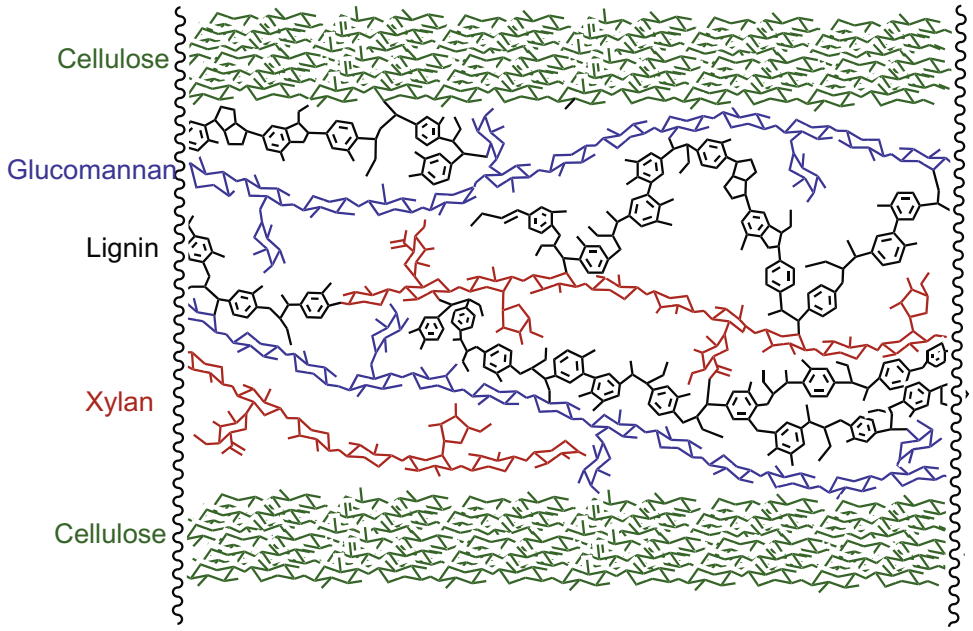


Figure 3.3: A representation of wood composition based on chemical structures (Oinonen et al., 2015) showing the cellulose in crystalline form – as rigid parallel chains – with a hemicellulose matrix that is embedded in cross-linking lignin (see Section 3.2). The structure shows several of the lignin-carbohydrate bonds that are discussed in this chapter and depicted at greater detail in Figure 3.4. *Kindly provided by Prof Gunnar Henriksson and reprinted with permission from Elsevier Ltd.*

prepared either by i) gradually degrading the polymer towards the LC-bond (natural substrate) or ii) synthesizing a selected part of the structure. The preparation of assay substrates by one of these methods is a central theme to this thesis.

A molecule with inseparable lignin and carbohydrate components is often referred to as a *lignin-carbohydrate complex*, or *LCC* (Björkman, 1954). The covalent bonds joining the components on an LCC are in some literature called LCC-bonds, which is generally synonymous to LC-bonds (Henriksson, 2009).

3.3 Ether-type LC-bonds

The most interesting and general of the proposed LC-bonds is the ether bond between the α-carbon of the arylpropane (or the γ-carbon of arylpropene; Watanabe et al. (1989)) lignin unit and the hydroxyl of a carbohydrate residue (Figure 3.4; Košíková et al. (1979), Ö Eriksson et al. (1980), Obst (1982), Lundquist et al. (1983), Xie et al. (2000), Karlsson et al. (2004), Balakshin et al. (2007), and Balakshin et al. (2011); reviewed in Watanabe (1989), Jeffries (1990), and Brunow and Lundquist (2010). Ether bonds are stable and likely sources of residual lignin in paper and pulp production (Gierer and Wännström, 1986; Taneda et al., 1987). Few specific etherases are known (see section 4.3), why the discovery of LC-etherases would be a substantial novelty.

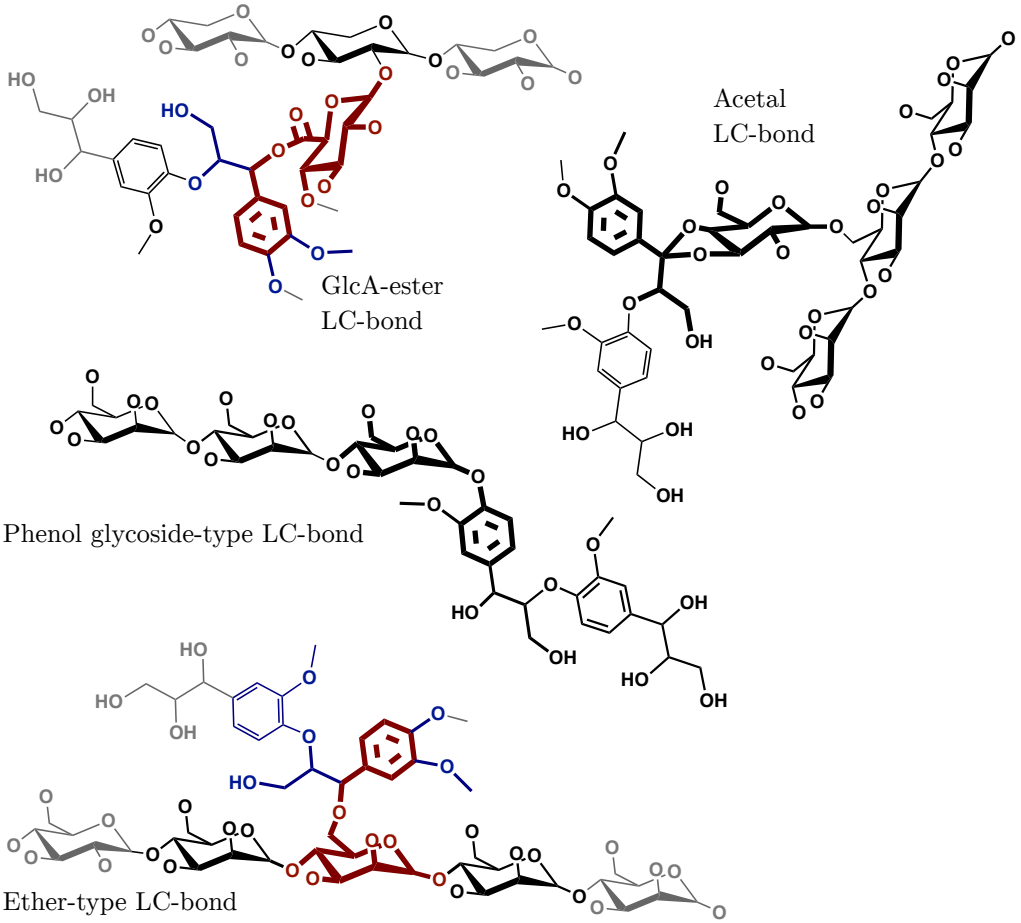


Figure 3.4: Illustration of the core fragments of natural lignin-carbohydrate complexes: the GlcA-ester LC-bond of an AGX side chain, as well as the acetal, phenolic glycoside and ether LC-bond types exemplified as LC-bonds to GGM. Regarding the selection for structures to prepare for use in synthetic substrates, the bolded inner part, consisting of a carbohydrate residue and an aryl, could be considered the minimal structure for chemospecific assaying. Though, use of progressively larger fragments, illustrated by increasingly lighter shading, may be required for enzyme activity, for enzyme selectivity or for substrate stability (Section 6.2.2). In the GlcA LC-ester fragment the inner red part corresponds to substrate **2** of Paper III and the red and blue parts together correspond to substrate **1** of Paper II. These are paralleled by similar structures in the ether-type LC-bond fragment that are discussed in the text as possible synthetic-substrate assays (Section 3.3.1).

The generality of the LC-ether bond stems from the fact that all exposed hydroxyls of hemicellulose monomers participate in such bonds. In an extensive experiment Watanabe et al. (1989) acetylated, *DDQ*-oxidized[2] and methylated two LCC preparations, from normal and compressed wood of Japanese Red Pine (*Pinus densiflora*). Analysis of the hydrolysis products has shown the presence of most lignin ether bonds at the C-6 hydroxyl of mannose, galactose, and glucose, with some linkages at the C-2 and C-3 positions. For xylose, the only available hydroxyls in xylan (C-2 and C-3), have also been detected. In addition, evidence of LCC ether bonds at C-2

[2]2,3-dichloro-5,6-dicyano-l,4-benzoquinone, see Section 5.5

and C-3 of arabinose was found in a study on Black Spruce (*Picea mariana*) LCC, using a different analytical method (Ö Eriksson et al., 1980).

Many studies support the existence of LC-ether bonds, and lignin–pectin ether bonds have also been suggested (see Jeffries, 1990, for an overview). In the last decade, LC-ether bonds have also been observed in high resolution 2D NMR (Section 5.4; Xie et al. (2000), Balakshin et al. (2007), Balakshin et al. (2011), Yuan et al. (2011), Du et al. (2014), and Ando et al. (2015)). The prevalent view is that cellulose does not directly bind to lignin but that cellulose-lignin interaction are mediated by intermediate hemicellulose (Brunow and Lundquist, 2010), though lignin–cellulose bonds may be formed during pulping (Gierer and Wännström, 1986; Lawoko et al., 2003).

3.3.1 Synthetic LC-ether substrates

For use as a synthetic substrate in an assay for LC-ether bond hydrolysing enzymes, ethers of guaiacylglycerol β-guaiacyl to the any of the relevant hydroxyls of sugar residues (*eg.* **7** and **8** drawn in the Appendix, *pp.* 149; see also Figure 3.4) could be seen as the minimal substrates identifiable as LCC ethers (Karlsson et al., 2004). However, even further simplified substrates, using a single benzyl (*eg.* **9** and **10**) may be sufficient for enzyme recognition, paralleling the substrate simplification made for GE substrates (Section 3.4.1; papers II and III).

LC-ether substrates are not commercially available, but many syntheses have been performed (Ralph and Young, 1983; Taneda et al., 1987; Sipilä and Brunow, 1991a; K Li and Helm, 1995; Karlsson et al., 2004). Given the difficulty of direct observation, the stereo- and regiostructure of intact LC-ether can only be inferred, and given this racemic nature of lignin (Henriksson, 2009), and the stereospecificity of lignin-degrading etherases (see section 4.3), the different stereoisomers may require different enzymes. If the selectivity for aryl structures is as relaxed in LC-etherases as it is for GEs (sections 4.1 and 6.4), substrates such as **9** and **10** could be suitable for enzyme recognition. In addition, when selecting an LCC fragment to synthesize, the stability (Paper II and Section 6.2.2) and water solubility (Section 6.3) of the substrate are important factors (Chapter 6). All in all, many structures for LCC ether substrates are theoretically possible, yet need to be synthesised prior to evaluation as enzyme substrates.

3.4 Ester-type LC-bonds

The best supported of the LCC-bond types is the ester bond linking the α- or γ-carbon of the β-diaryl lignin unit (Figure 3.4) to the carboxyl of a 4-*O*-methyl glucuronic acid (mGlcA) side chain of a xylan (*i.e.* glucuronoxylan and AGX; see Section 4.6 and Figure 4.4; Ö Eriksson et al. (1980), Das et al. (1981), Obst (1982), Lundquist et al. (1983), Das et al. (1984a), Das et al. (1984b), and Watanabe and Koshijima (1988), for an overview, see Jeffries (1990)).

Ester bonds have been central LCC bond candidates since the early studies of lignin-carbohydrate association (see Ö Eriksson et al., 1980, for an overview). Evi-

dence of LCC esters includes: i) the implication of carboxyls in LCC binding (Bolker, 1963), ii) the degree of esterification (Wang et al., 1967) of mGlcA substitutions of xylan (see Figure 4.4), iii) the disassociation of lignin and carbohydrates upon alkaline (Yaku et al., 1976; Ö Eriksson et al., 1980) and borohydride treatments (Das et al., 1981), as well as more recent iv) 2D NMR experiments (Xie et al., 2000; Balakshin et al., 2007; Balakshin et al., 2011; Yuan et al., 2011).

The early observations supporting LC-ester bonds (Bolker, 1963; Wang et al., 1967) were substantiated by a series of experiments during the 1980's using borohydride reduction, alkali treatment, or DDQ oxidation (Ö Eriksson et al., 1980; Obst, 1982; Lundquist et al., 1983; Das et al., 1984a; Das et al., 1984b; Watanabe and Koshijima, 1988). While DDQ oxidation has been regarded as specific to the α-carbon ester linkage (Watanabe, 1989), LC-ester model compounds in solution appear to be able of migrating to the arylpropane γ-position (K Li and Helm, 1995) and in recent studies, only 2D NMR signals attributable to GlcA γ-esters were detectable (Balakshin et al., 2007; Balakshin et al., 2011).

In addition to the GlcA ester bond, *p*-coumaryl and ferulic acid esters occur in plants, as inter- and intra-polymeric cross-links (Lam et al., 2001). However, to the best of my knowledge, they are not prevalent in wood fibres (see Section 4.2).

3.4.1 Synthetic LC-ester substrates

Synthesis and enzymatic hydrolysis of β-O-4′ glucuronate

In Paper II we followed the method used for LC-ethers (Sipilä and Brunow, 1991a) and supposedly used for LC-esters by Sipilä and Brunow (1991b) to synthesize the guaiacylglycerol-β-guaiacyl glucuronate LC-ester model **1** (β-O-4′ glucuronate). Model **1** was chosen as the synthesis target, as it constitutes the smallest structure specifically representing an LC-bond GlcA ester. Variants of this compound have been synthesized several times by multiple routes (Enoki et al., 1983; Ralph and Young, 1983; Joniak, 1995; K Li and Helm, 1995; Toikka et al., 1998). While **1** proved to be a very interesting LC-bond assay substrate (Section 6.1.2), its instability, even at acidic pH, prevented any further work (see Section 6.2.2).

GE assays using commercial BnGlcA

In Paper III, I developed several GE assays based on the commercially available substrate benzyl glucuronate **2** (BnGlcA; d'Errico et al., 2015). **2** represents an even smaller structure than **1**, but is not identifiable as a lignin analogue. Also, the ester bond in **2** is chemically different from the one in **1**, which has two electron-donating ring substituents and a vicinal alkoxy group. **2** is more stable than **1** (Section 6.2.2), and it was proven to be useful GE assay substrate if the assay conditions and sample handling are chosen to accommodate for substrate instability.

3.5 Phenyl glycoside LC-bonds

Phenol glycosidic LC-bonds are interesting, due to their *in vitro* formation and potential hydrolysability by glycosidases. The existence of phenyl glycoside LC-bonds has been suggested by several studies (Yaku et al., 1976; Joseleau and Kesraoui, 1986; Lawoko, 2005; Balakshin et al., 2007). Phenyl glycosides are formed by enzymes excreted by wood-rotting fungi (Kondo et al., 1993) and in the laboratory they can be prepared *in vitro* through transglycosylation by β-glucosidas (Kondo et al., 1988; Biely and Puchart, 2006).

Studies of phenyl glucosidic bonds and complicated by the ease with which these bonds are formed, possibly even during enzymatic hydrolysis (which typically includes β-glucosidases). As a consequence, the origin of phenyl glycosides that are found in preparations of natural substrates could generally be questioned. On the other hand, preparation of synthetic substrates for phenyl glycoside assays may be greatly facilitated by enzymatic conjugation. Furthermore, many types of phenyl glycosides are commercially available, and some are routinely used as substrates for glycosidic assays. However, it would not be trivial to design an enzyme assay that is specific to LC-bond phenyl glycosides, distinguishing them from other glycosidic activities (see Section 7.4.8 for a discussion on the possibilities for phenyl glucoside assays).

3.6 Acetal-type LC-bonds

The acetal LCC-bonds were heralded as "the" LCC-bonds in a brief report in 1963, based on literature review and FTIR analysis (Section 5.1.2) of 18 wood and pulp samples (Bolker, 1963). In a ^{13}C tracer NMR experiment published in 2000 (Xie et al., 2000), an NMR peak was assigned to the acetal-type LCC-bond based on an early study on model compounds of phenolic proanthocyanidins (Jacques et al., 1974). Whereas the 1963 study assigned the bond to the arylpropane β-carbon, the 2000 experiment assigned it to the α-carbon. Given the weak evidence, scarcity of studies, and the limited stability of acetals in acid solutions, acetal LCC-bonds appear to be the least relevant for enzymatic LCC degradation.

3.7 Naturally-derived substrates

A preparation of lignocellulosic material that can be used as a substrate in an LC-bond assay (*natural substrate*, *LCC substrate* or *LCCs* depending on emphasis) should have the following characteristics:

- There are LC bonds of a certain type (the *target* LC-bonds)
- We can detect, and ideally quantify, either the presence of the target LC bonds *or* the loss of them
- The target bond would be accessible to a potential LC-bond-acting enzyme

Lignocellulosic material cannot be used for LC-bond assays without preparation: Native lignocellulosic materials are generally impermeable to enzymes. Also, given the low frequency of LC-bonds (Obst, 1982; Balakshin et al., 2011), these are not identifiable in intact lignocellulosic materials even by high resolution methods (Kim et al., 2012). Consequently, to use a natural substrate in LC-bond assays it is necessary to i) modify a material to expose its LC-bonds to enzymatic action, and ii) enrich the number of LC-bonds sufficiently to be able to detect and quantify their degradation.

As LC-bonds cannot be quantified in intact substrates, preparation of LCC substrates must proceed without analytic confirmation, and rely instead on published information and personal experience. In addition, the procedure may introduce changes, including losing and gaining LC-bonds (*eg.* Kleinert, 1970; Gierer and Wännström, 1986; Kondo et al., 1988). Furthermore, the inability to determine LC-bonds in the precursor material means that the representativeness of the derived preparation cannot be objectively assessed. While the wealth and variation in source material and preparation techniques that have been employed make direct comparisons difficult.

LCC-rich lignocellulosic materials have been prepared for a multitude of materials by different combinations of techniques. A few examples are given here for reference: Materials have been prepared from softwood (Lawoko et al., 2006), hardwood (Takahashi and Koshijima, 1988), and grasses (Singh et al., 2005), and include isotope-labeled material (Xie et al., 2000) and genetically modified plants (Min et al., 2014). LCCs are commonly prepared by a combination of steps: Initial physical disruption of the material structure include various industrially relevant treatments, such as steam explosion (J Li et al., 2007), chemical pulping (Košíková and Ebringerova, 1994; Lawoko et al., 2003), mechanical pulping (Oinonen et al., 2013), as well as various forms of laboratory grinding (Lawoko et al., 2006) and ball-milling procedures (Björkman, 1954). For the separation of the material components, and for the concentration of LCCs, solubilization (Imamura et al., 1994) and precipitation (Koshijima et al., 1972), dialysis (Yaku et al., 1976), chromatography (Paper I; Takahashi et al. (1982)) or some form of extraction (Zikeli et al., 2014) is generally used. To enrich the LCC structures, enzymatic treatment by glycosidases (Section 4.6) is commonly used as a way to selectively degrade carbohydrates of the preparation (Pew and Weyna, 1962; Iversen, 1985; Watanabe et al., 1989; Lawoko, 2005; Du et al., 2014). For a comprehensive description of LCC preparation methods, see Balakshin et al. (2014) or Koshijima and Watanabe (2003).

3.7.1 Natural LCC substrates used in the thesis work

The following preparations of natural substrates were used in the present work. All the procedures were based on a plant meal (mainly Norway spruce, *Picea abies*), obtained through grinding or milling. Formation of new LCC-bonds during this step is possible, as radical formation on mechanical chain rupture can be expected (Ikeda et al., 2002) and may lead to LC-bond formation (Kleinert, 1970; Iversen, 1985).

Hot-water extract of spruce In Paper I, wood meal from Norway spruce was extracted according to Song et al. (2008). Size-exclusion chromatography and enzymatic carbohydrate degradation was then applied to the intermediate fraction showing that the size of lignin-containing components (as detected by their UV absorbance, see Section 5.1.1) was reduced upon carbohydrate hydrolysis, thus indicating the presence of LCCs. Hot-water extraction is a mild procedure and could in itself not be expected to give rise to significant amounts of LC-bonds.

Various hot-water extracts and thermo-mechanical pulping fluids Various hot-water extracts from ball-milled or ground plant material and freeze-dried 1–5 kDa fractions of thermo-mechanical pulping (TMP) fluids (Oinonen et al., 2013) were used as substrates in an attempt to develop a GE assay (Section 5.1.2), but no LCC esters could be confirmed. However, a 5–10 kDa hot water extract from spruce sawdust, was subsequently shown to contain LC-bonds (Oinonen et al., 2015), probably LC-ether bonds. These materials are comparable to the spruce hot-water extracts above.

Azuma LCCs In an effort to create an LCC substrate for GE assays, wood chips from Norway spruce were ground, jet milled and extracted according to Azuma and Koshijima (1988). The freeze-dried product of the solvent extraction, without fractionation, was then subjected to mass spectrometry (see Section 5.3) and NMR (see Section 5.4). As in ball milling, LCCs could potentially form during jet milling.

Sodium chlorite Xylan extracts The spruce AGX extract used in Paper IV was de-lignified by acidic sodium chlorite and extracted by potassium hydroxide (Escalante et al., 2012). While the substrate had a lignin content of 5.7%, the near-complete enzymatic hydrolysis indicate that very little lignin was bound to soluble AGX in the form of LC-bonds. Given the harsh de-lignification treatment, it is unlikely that remaining lignin residues retained their native structure (Gierer and Huber, 1964).

Chapter 4

Enzymes

Four enzyme groups, relevant to LC-bond assays, are discussed in this chapter: i) The glucuronoyl and feruloyl esterases, the two enzyme classes that have been attributed the capacity to hydrolyse LC-bonds, (sections 4.1 and 4.2, respectively). ii) Etherases, among which there are enzymes known to break lignin–lignin ether bonds, enabling lignin depolymerization (Section 4.3). iii) Metalloproteins that act on lignin or carbohydrates (sections 4.4 and 4.5, respectively). iv) Glycosidases (Section 4.6) which have the potential to degrade complex natural materials into suitable substrate precursors (see Section 7.3.1).

4.1 Glucuronoyl esterases

Glucuronoyl esterases (GE) were discovered serendipitously during a study on α-glucuronidase (Biely et al., 2000). It was noted that a crude *Schizophyllum commune* enzyme preparation liberated mGlcA (**2**) from methyl ester of 4-*O*-methyl–D-glucuronic acid α-1,2-linked to 4-nitrophenyl-β-D-xylopyranoside (**12**; see the appendix, *pp.* 149, for structural formulae of the compounds referred to in this chapter.) without prior demethylation, suggesting α-glucuronidase activity (Špániková and Biely, 2006). The enzyme, *Sc*GE1, was subsequently purified from cellulose-spent *Schizophyllum commune* culture fluid and characterized on **12** (Table 4.1). The enzyme was found to have no activity on citrus pectin, methyl, and ethyl ferulate esters, or on paranitrophenol acetate. Conversely, enzymes that are active on one or more of these

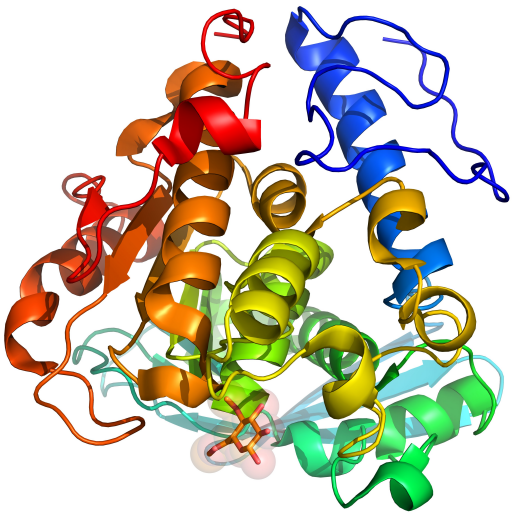


Figure 4.1: Model of the structure of the catalytic domain of *St*GE2 in complex with the a methyl mGlcA substrate (**11**), for which both C-1 and carboxy methyl substituents protrude away from the enzyme (PDB ID 4g4j; Charavgi et al., 2013). Prepared by PyMOL (Schrodinger, LLC, 2010).

substrates (four feruloyl esterases, three acetyl xylan esterases, and two pectinase preparations) did not have any activity on any of the two methyl glucuronates used in the study (**11** and **12**; Špániková and Biely (2006)).

The discovery of *Sc*GE1 in the culture filtrate of a wood-rotting fungus, indicated that GEs may hydrolyse the LC-ester bonds of GlcA side chains to AGX (Section 3.4). While this hypothesis is yet to be confirmed, a number of reports have shown GE activity on various substrates. For instance, *Sc*GE1 was shown to have a broad specificity for the carboxyl component of the substrate, being active on a number of methyl esters of glucopyransyluronate xylo-oligomers (Table 4.1; Špániková et al. (2007)). Activity on methyl esters of polymeric glucuronoxylan has since been demonstrated for *Sc*GE1 and two other GEs (Biely et al., 2015). However, when tested for activity on methyl esters of galacturonic acid – the 4-epimer of GlcA, – no activity was found for *Sc*GE1 on **13** and **14**, or for *Pc*GE1 and *Pc*GE2 on **13** (Duranová et al., 2009a).

In addition, the two first discovered GEs, *Sc*GE1 and Cip2 (here referred to as *Hj*Cip2; XL Li et al. (2007)) were shown to be active on esters of aryl alcohols of GlcA XL Li et al. (2007) and Špániková et al. (2007). Since then, understanding of the chemo- and regio-selectivities of GEs has increased notably (Table 4.1). Katsimpouras et al. (2014) showed that *Pa*GE1 was active on two aryl propane γ-esters, Paper II demonstrated GE activity on the β-O-4′ glucuronate (**1**), Paper III and d'Errico et al. (2015) found activity of several GEs on various phenyl and phenyl alkyl alcohol esters of GlcA, and d'Errico et al. (2016) recently presented GE activity on **15**, a benzoylated γ-bond variant of the β-O-4′ glucuronate.

Based on the sequence of a *Sc*GE1 peptide fragment, the gene sequence of *Hj*Cip2 was selected for expression in the hope that it would prove to have the same GE activity as *Sc*GE1. The positive conclusion of this effort determined the first GE sequence and led to the establishment of a distinct family of enzymes (XL Li et al., 2007). This paved the way for many of the studies mentioned here, as well as the Topakas et al. (2010) establishment of the consensus sequence (G-C-S-R-X-G) around the nucleophilic serine.

Two crystal structures of GEs have been determined to date. First, the catalytic domain of *Hj*Cip2 (Pokkuluri et al., 2011), followed by *St*GE2 (Charavgi et al., 2013), the latter also in complex with **11**. The structures revealed that the catalytic site was situated on the enzyme surface (Figure 4.1), thus explaining the low regio-selectivity for the aryl moiety (Table 4.1).

Despite efforts to show LC-bond-breaking activity on natural substrates, including considerable ones made as part of this work, the natural role of GEs has not been elucidated. However, there are good indication that there is a role to be found: i) there is a high degree of esterification of mGlcA in wood xylan (Lundquist et al., 1983), ii) the heterologous expression of a *Phanerochaete carnosa* GE over-expression in *Arabidopsis thaliana* caused increases in lignin content and significant decreases in the cell wall thickness of lignified cells (Tsai et al., 2012), iii) the presence of GE genes in plant-degrading fungi (Duranová et al., 2009b) and iv) the increase in hydrolysis seen in the enzymatic degradation of corn fibre on the addition of GEs (d'Errico et al., 2016).

Table 4.1: Summary of quantitative data from GE assays presented in the literature. For the structure of the substrates, refer to the Appendix (*pp.* 149). To fit the table format, published data has been unit-converted and k_{cat}/K_M calculated as required.

Enzyme	Substrate	K_M	k_{cat}	k_{cat}/K_M	±	Reference
*St*GE2	**1**	11±5 mM	7.7±1.4 s^{-1}	0.7±0.3 $mM^{-1}s^{-1}$	e	Paper II
*St*GE2	**2**	8.9 mM	7.8 s^{-1}	0.4 $mM^{-1}s^{-1}$		Paper III
*Pa*GE1	**2**	12.1 mM	3.5 s^{-1}	0.6 $mM^{-1}s^{-1}$		Paper III
*Sc*GE1	**2**	51±8.0 mM	64±5.0 s^{-1}	1.2 $mM^{-1}s^{-1}$	e	(d'Errico et al., 2015)
*Cu*GE	**2**	80±24 mM	48±7.5 s^{-1}	0.6 $mM^{-1}s^{-1}$	e	(d'Errico et al., 2015)
*Sc*GE1	**11**		*25±2 s^{-1}		s	(Špániková et al., 2007)
*Sc*GE1	**11**		*10 s^{-1}			(Duranová et al., 2009a)
*Hj*Cip2	**11**	0.5 mM	4.5 s^{-1}	9.7 $mM^{-1}s^{-1}$		(XL Li et al., 2007)
*Hj*Cip2	**11**		*8.9 s^{-1}			(Duranová et al., 2009a)
*Pc*GE1	**11**		*15 s^{-1}			(Duranová et al., 2009a)
*Pc*GE2	**11**		*15 s^{-1}			(Duranová et al., 2009a)
*Sc*GE1	**12**	0.31 mM	3.2 s^{-1}	10.32 $mM^{-1}s^{-1}$		(Špániková and Biely, 2006)
*Sc*GE1	**12**	0.25 mM	9.27 s^{-1}	37 $mM^{-1}s^{-1}$		(Wong et al., 2012)
*Hj*Cip2	**12**	0.5 mM	4.5 s^{-1}	9.7 $mM^{-1}s^{-1}$		(XL Li et al., 2007)
*Pc*GE1	**12**	0.83 mM	11.2 s^{-1}	13.5 $mM^{-1}s^{-1}$		(Duranová et al., 2009a)
*Pc*GE2	**12**	1.82 mM	62.3 s^{-1}	34.2 $mM^{-1}s^{-1}$		(Duranová et al., 2009a)
*St*GE1	**12**	1.3±0.1 mM	0.81±0.04 s^{-1}	0.6±0.1 $mM^{-1}s^{-1}$		(Vafiadi et al., 2009)
*Pa*GE1	**12**	7.6 mM	0.28 s^{-1}	0.04 mM^{-1} s^{-1}		(Katsimpouras et al., 2014)
*Sc*GE1	**16**		*13 s^{-1}			(Duranová et al., 2009a)
*Hj*Cip2	**16**		*5.3 s^{-1}			(Duranová et al., 2009a)
*Pc*GE1	**16**		*12 s^{-1}			(Duranová et al., 2009a)
*Pc*GE2	**16**		*13 s^{-1}			(Duranová et al., 2009a)
*Sc*GE1	**17**	11±5.8 mM			e	(d'Errico et al., 2015)
*Cu*GE	**17**	8.9±3.2 mM			e	(d'Errico et al., 2015)
*Sc*GE1	**18**	3.7±1.2 mM	118±9.4 s^{-1}	32 $mM^{-1}s^{-1}$	e/s	(d'Errico et al., 2015)
*Cu*GE	**18**	4.6±1.0 mM	129±7.6s^{-1}	28 $mM^{-1}s^{-1}$	e/s	(d'Errico et al., 2015)
*Sc*GE1	**19**	4.31 mM	0.2 s^{-1}	0.05 $mM^{-1}s^{-1}$		(Špániková et al., 2007)
*Sc*GE1	**20**	66±22 mM	15±2.7 s^{-1}	0.2 $mM^{-1}s^{-1}$	e	(d'Errico et al., 2015)
*Cu*GE	**20**	55±14 mM	17±1.8 s^{-1}	0.3 $mM^{-1}s^{-1}$	e	(d'Errico et al., 2015)
*Sc*GE1	**21**	1.78 mM	7.8 s^{-1}	4.38 $mM^{-1}s^{-1}$		(Špániková et al., 2007)
*Hj*Cip2	**21**			1.1 $mM^{-1}s^{-1}$		(XL Li et al., 2007)
*Sc*GE1	**15**	1.4±0.3 mM	125±6.9 s^{-1}	89 $mM^{-1}s^{-1}$	s	(d'Errico et al., 2016)
*Cu*GE	**15**	3.4±0.7 mM	285±22 s^{-1}	83 $mM^{-1}s^{-1}$	s	(d'Errico et al., 2016)
*Pa*GE1	**22**	1.34±0.4 mM	0.19±0.03 s^{-1}	0.14±0.04 $mM^{-1}s^{-1}$	e	(Katsimpouras et al., 2014)
*Pa*GE1	**23**	0.4±0.1 mM	0.78±0.05 s^{-1}	0.82±0.13 $mM^{-1}s^{-1}$	e	(Katsimpouras et al., 2014)
*St*GE2	**23**	7.24±3.3 mM	2.8±1.0 s^{-1}	0.4±0.2 $mM^{-1}s^{-1}$	e	(Katsimpouras et al., 2014)
*St*GE2	**24**	3.63±0.6 mM	1.9±0.1 s^{-1}	0.5±0.1 $mM^{-1}s^{-1}$	e	(Katsimpouras et al., 2014)
*Pa*GE1	**24**	2.66±0.5 mM	5.3±0.7 s^{-1}	2.0±0.5 $mM^{-1}s^{-1}$	e	(Katsimpouras et al., 2014)
*Sc*GE1	**25**		*30±1 s^{-1}		s	(Špániková et al., 2007)
*Sc*GE1	**26**		*31±1 s^{-1}		s	(Špániková et al., 2007)
*Sc*GE1	**27**		*34±2 s^{-1}		s	(Špániková et al., 2007)
*Sc*GE1	**28**		*28±2 s^{-1}		s	(Špániková et al., 2007)

The "±" column indicates whether the given uncertainty is a standard error (e) or a standard deviation (s). Values marked with an asterisk (*) in the k_{cat} column are enzyme activities, measured at 20 mM substrate concentration and included for reference.

4.2 Feruloyl esterases

Feruloyl esterases (FAEs) are named for their ability to hydrolyse esters of trans-ferulic acid (FA) but they display a considerable variation in substrate specificity, generally hydrolysing the ester bonds of various hydroxycinnamates (Crepin et al., 2004; Levasseur et al., 2006; Wong, 2006). The benzyl position of FA can be esterified to lignin (Lam et al., 2001). It is also often esterified at the 5-*O* position of arabinose on AGX side chains, though esterification to pectin and other targets is also prevalent (Fry, 1982; Ralph et al., 1996; Lam et al., 2001). However, the relevance of FAEs for softwood and hardwood lignocellulose is unclear given that most studies have been carried out in grasses, where FA is much more abundant (Ralph et al., 1996; Oinonen et al., 2013). In one study (Reiter et al., 2013) FA was produced in the depolymerization of softwood kraft lignin (Reiter et al., 2013), but it is difficult to find concrete published evidence for the presence of FA in wood fibre cells, though there are several reports of esters of ferulic acid in softwood bark (Rowe et al., 1969; Laver and Fang, 1989; Virgili et al., 2000).

In grass tissues, FA dimers commonly occur as cross-links of hemicellulose chains (Williamson et al., 1998) and some FAEs can hydrolyse the FA dimer esters, increasing hydrolysis yields (Topakas et al., 2007). It is not clear, however, if there exists FAEs with a regioselectivity that allows hydrolysis of bonds between carbohydrate and the lignin polymer, and therefore they will not be further discussed.

4.3 Etherases

At present, no LC-ether-breaking enzymes are known, even though this is the most abundant group of LC-bonds (Section 3.3). Interestingly, a few relevant etherases are known. Given that the discovery of LC-bond etherases would be a significant achievement it is relevant to construct some conceptual models for how such an enzyme would work: As ether bonds are relatively stable it is likely that enzymatic LC-ether scission requires energy input. Thus, anticipating possible energy sources and including them in screening assays for LC-ether-breaking enzymes might increase the chance of success.

The bacterial β-etherase system is capable of breaking α-ethers of lignin β-diaryl (Figure 4.2). It was discovered in *Sphingobium* sp. SYK-6 (Masai et al., 1993; Masai et al., 2003) and later found in other species (Picart et al., 2014). According to the reaction model for (refined by Gall et al., 2014), the α-hydroxyl of a β-aryl ether is first reduced to a ketone by a NAD^+-dependant dehydrogenase. A glutathione-*S*-transferase then transfers the aryl propanol to a reduced glutathione (GSH), releasing the guaiacol and a thioetherase for removal of the glutathionyl. Recent studies have shown that several homologous enzymes with different stereospecificities exist (Masai et al., 2007; Gall et al., 2014; Helmich et al., 2016) and that the process can be run *in vitro*, for de depolymerization of natural as well as and synthetic lignin models if that NAD^+ and GSH are recycled (Reiter et al., 2013; Rosini et al., 2016).

In the context of LC-etherases, the β-etherase degradation system can be viewed as a model of i) how etherase activity can be driven, ii) ether degradation, and

iii) scission of an ether bonds to a β-aryl ether (the bond type of the stipulated LC-ether; Section 3.3). In addition, the ability to depolymerize lignin *in vitro* may be exploited for the enrichment of LC-bonds in naturally-derived LCC substates for use in LC-bond assays (Section 7.3.1) and for direct observations of LC-bonds (Section 7.2).

Another known intracellular etherase is the *N*-acetylmuramic acid 6-phosphate lyase (MurNAc etherase) (reviewed in Jaeger and Mayer, 2008). This enzyme removes the etherically-bound lactic acid from *N*-acetylmuramic acid (MurNAc), a bacterial cell wall sugar. Apart from its etherase activity, the proposed mode of action for this enzyme represents an interesting model (Figure 4.3 ; Jaeger and Mayer, 2008). Like LPMO enzymes (Secion 4.5), MurNAc etherase removes an ether-bonded carbohydrate substituent, though it should be noted that in MurNAc the eliminated alkoxy is vicinal to an amination, which may be crucial to the enzyme action.

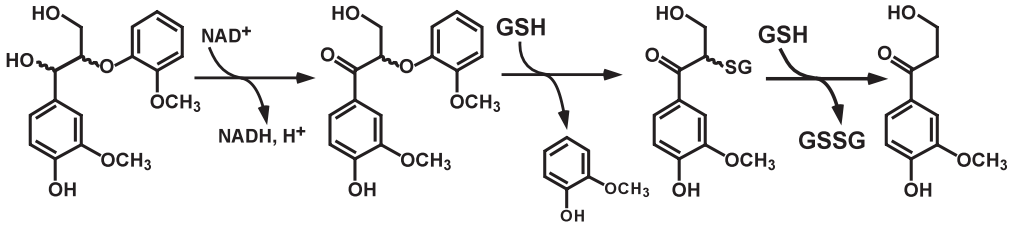


Figure 4.2: Enzymatic degradation of a β-aryl ether. The first step is catalysed by a Cα-dehydrogenase, the second by a glutathione-*S*-transferase and the third by a glutathion lyase of the bacterial β-aryl ether degradation pathway. Adapted from Masai et al. (2007). GSH: glutathione; GSSG glutathione disulfide.

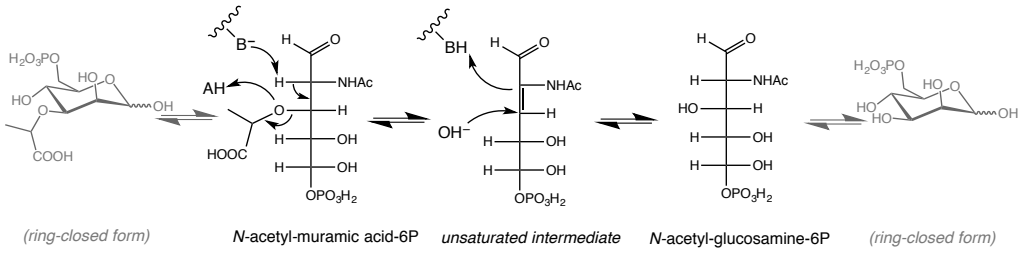


Figure 4.3: Proposed reaction mechanism of the MurNAc etherase (Jaeger and Mayer, 2008). The α-carbon of MurNAc is detached by enzymatic base catalysis and, probably facilitated by a protonation, lactic acid is eliminated, creating an unsaturated intermediate. The enzyme catalyses the hydration of the intermediate, yielding an *N*-acetyl-glucosamine 6-phosphate. For clarity, the reverse reaction is not drawn.

4.4 Lignin-modifying enzymes

4.4.1 Lignin-, manganese-, and versatile peroxidases

Lignin peroxidase (LiP), manganese peroxidase (MnP), and versatile peroxidase (VP) are secreted, heme-containing enzymes, that scavenge electrons from their substrates to form H_2O using H_2O_2 as electron acceptor (comprehensively reviewed in Wong, 2009). LiPs can oxidize various low-molecular-weight aromatic substrates, forming radicals that can cause the substrate structure to lyse, but may result in secondary polymerization, depending on the fate of the formed radical (Widsten and Kandelbauer, 2008; Henriksson, 2009). MnPs reduce a bound Mn^{II} to Mn^{III} and allow the ion to diffuse into a macromolecular substrate into which the enzyme itself is too large to penetrate (KE Eriksson, 2010). VPs appear to be hybrids of LiPs and MnP but have a wider substrate range and use a Mn^{III} for catalysis (Pérez-Boada et al., 2005).

4.4.2 Multi-copper oxidases and laccases

Multi-copper oxidases (MCO) are a widespread group of oxidoreductases that include laccases (Solomon et al., 1996). These oxidize low-molecular-weight aromatic compounds in an unspecific manner using oxygen as electron acceptor to form phenoxy radicals, which may lead to depolymerization or polymerization, depending on the chemical context of the system (Widsten and Kandelbauer, 2008). Laccases have found a widespread industrial use (KE Eriksson, 2010). While their activity may be relevant to LC-bond hydrolysis assays, their lack of specificity and control of the radical reactions means they fall outside the scope of this thesis.

4.5 Lytic polysaccharide mono-oxygenases

Lytic polysaccharide mono-oxygenases (LPMOs), (recently and comprehensively reviewed as cellulose degrading enzymes in Beeson et al., 2015), are a large and newly described group of enzymes. They oxidatively cleave glycosidic bonds through a reaction catalysed by a Cu(II) ion held in the flat active site (Quinlan et al., 2011). The substrate is hydroxylated using oxygen and an electron donor: ascorbic acid (Vaaje-Kolstad et al., 2010), gallic acid (Quinlan et al., 2011), cellobiose (catalysed by cellobiose dehydrogenase; CDH) (Langston et al., 2011; Vu et al., 2014; Beeson et al., 2015; Bennati-Granier et al., 2015; Hemsworth et al., 2015), or various small-molecular lignin-like compounds (Westereng et al., 2015). The cellulolytic LPMOs, on which most studies have been focused, oxidize glycosidic bonds of crystalline cellulose at C-1 or C-4 (see table in Hemsworth et al., 2015), leaving an oxidized carbohydrate residue at one of the newly formed chain ends.

LPMOs are found throughout fungal and bacterial kingdoms, and many cellulolytic fungi have a large number of LPMOs (Busk and Lange, 2015). Apart from cellulose and chitin, activity on xyloglucan (Agger et al., 2014), β-1,4-linked polysaccharides in general (Bennati-Granier et al., 2015), and starch (Lo Leggio

et al., 2015) have also been shown (see table in Hemsworth et al., 2015). While no LPMOs have been suggested to act on lignin or LC-bonds, the kind of oxidative regio-specific extracellular activity that they display can serve as a possible model for how LC-ether degrading enzymes could function, and catering for the requirements of LPMOs when screening for LC-bond degrading enzymes should be considered. This is further discussed in the Outlook (Chapter 9).

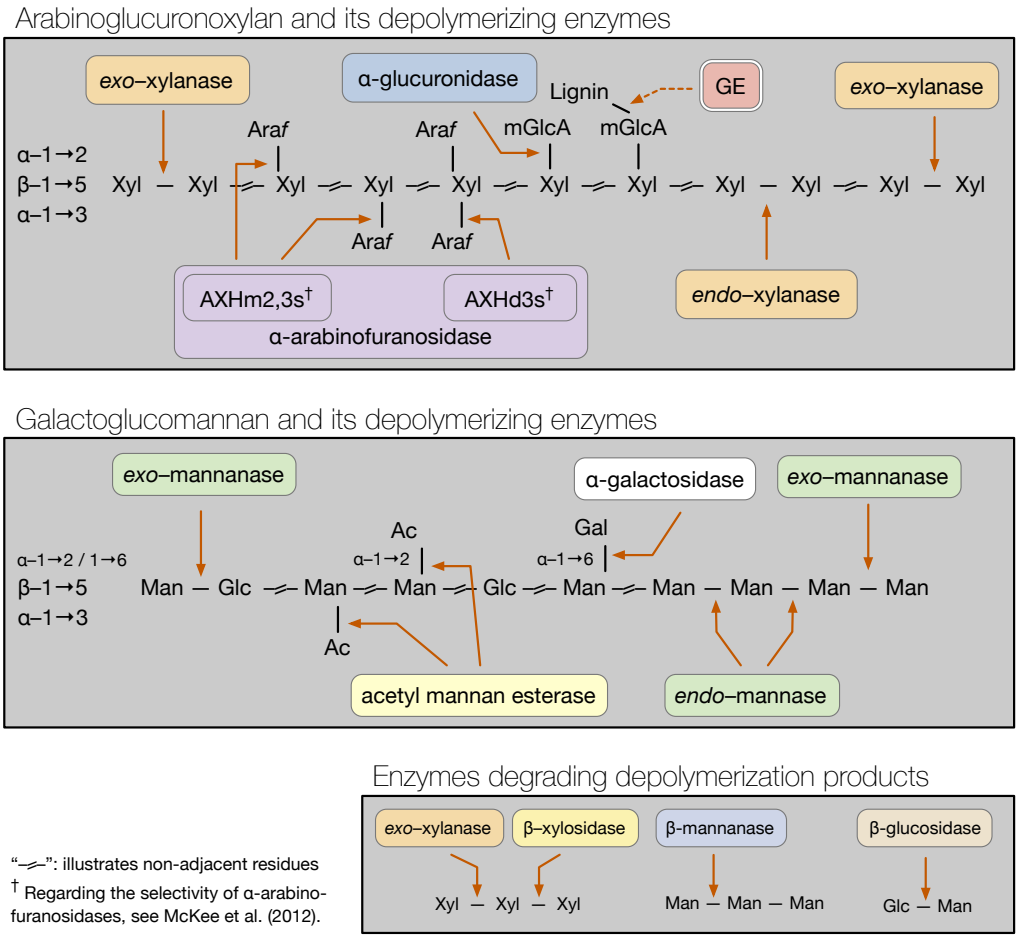
4.6 Glycosidases

Glycosidases hydrolyse the bonds between sugar residues (Davies and Henrissat, 1995). They are important for LC-bond assays for the preparation of naturally derived LCC substrates (Section 3.7) and for complex multi-enzyme and synergy assays (sections 7.4.1).

The two most common hemicelluloses types in softwood, and thus most relevant for softwood LC-bond degradation, are AGX and GGM. The structures of these hemicelluloses are outlined below and depicted in Figure 4.4, which also shows the glycosidic enzymes that act on xylans (the parent hemicellulose of AGX) and on glucomannans (the parent structure of GGM).

Arabinoglucuronoxylans (AGX) have a xylan main chain, consisting of β-1,4–linked xylose (Xyl) units. In addition, AGX is substituted by arabinofuranose (Ara*f*) and 4-*O*-methyl glucuronic acid (mGlcA). Acetylations at *O*-2 and *O*-3 are common in hardwood, but not in softwood (not depicted in Figure 4.4; Teleman (2009)). GlcA substitutions can be esterified to lignin (Section 3.4) or to cell-wall polysaccharides (Lundquist et al., 1983; Thompson and Fry, 2000) while Ara*f* substitutions can be esterified to ferulic acid (sections 3.4 and 4.2; Williamson et al. (1998)). The name glucuronoxylan (GAX; used in Paper IV) refers to a xylan with the same principal structure as AGX, but where mGlcA substitutions are more frequent than the Ara*f* ones.

Galactoglucomannans (GGM) have a glucomannan main chain, consisting of β-1,4–linked (Man) and glucose (Glc) at ratios varying from 1:1 to 4:1, depending on subtype (Timell, 1967; Moreira and Filho, 2008; Teleman, 2009). Softwood glucomannans have galactose (Gal) substituents and acetylations at *O*-2 and *O*-3 are common in both types (Teleman, 2009). Gal substitutions can be etherified to lignin (Section 3.3). In addition to the presence of LC-bonds linking the substituting monosaccharides, it has been suggested that there may be LC-ether bonds between the main chains of xylans and glucomannans (Section 3.3).



"⫽": illustrates non-adjacent residues

† Regarding the selectivity of α-arabinofuranosidases, see McKee et al. (2012).

Figure 4.4: Schematic view of AGX and GGM and the enzymes involved in their hydrolysis. The linkage positions of the main and side chains of the polymers are indicated on the left. The lower box shows how oligomers released from the hydrolysis in the two upper boxes can be further hydrolysed to monomers. (Paper IV, Moreira and Filho, 2008).

Chapter 5

Detection methods

In comparison to synthetic-substrate assays, detecting and quantifying LC-bonds against the background of a complex lignocellulosic matrix is a substantial challenge, and requires a good understanding of the analytical method. During my thesis work I have tested many analytical techniques as detection methods in enzymatic assays or for the evaluation of potential substrates. This chapter outlines the main techniques relevant for LC-bond assay design.

5.1 Spectroscopic detection methods

Spectroscopic analysis is based on the analyte's interaction with electromagnetic radiation. Two types of spectroscopy are relevant to our discourse: absorbance spectrophotometry and infrared (IR) spectroscopy.

5.1.1 Absorbance spectrophotometry

Absorbance spectroscopy in the ultra–violet to visible spectral region (UV–Vis spectrophotometry; ~200–800nm) is based on electronic transitions of non-bonded or π electrons. While the carbohydrates have a limited absorbance ~200 nm, the phenol derivatives found in lignin have a strong absorbance with an peak, usually around 280 nm, depending on structure (Azuma et al., 1981). Absorbance spectra for the two enantiomers of β-O-4′ unit are shown in Figure 5.1

Absorbance bands are comparatively wide why spectrophotometry is sensitive to the sample background and to the composition of analytes, limiting qualitative applicability. In the right conditions, spectrophotometry can be used for quantitative analyses of lignin and LCCs. It is also useful for the detection of chromophores in *eg.* coupled-enzyme assays (Section 7.4.1).

Spectrophotometric detection has been used extensively in this work as detection method in reverse-phase and size exclusion (SEC) reverse-phase high-performance liquid chromatography (HPLC)-based methods (sections 5.2.1 and 5.2.3 and papers II and III, and the GE assay on **2** (Section 6.3.2 and Paper III).

5.1.2 Infrared spectroscopy

IR uses rotational and vibrational excitations in the IR spectral region, which can be studied with a Fourier transform IR (FTIR) spectrophotometer. FTIR has applications similar to those of regular spectrophotometry (Bolker and Terashima, 2009), but its peaks originate from individual bonds and can be resolved with more precision. FTIR was used in Paper II to ascertain the synthesis of **1**. FTIR was also used as a detection method in an effort to develop a natural-substrate GE assay for hot-water extract (HWE) or thermo-mechanical pulping (TMP) (Section 6.1.1).

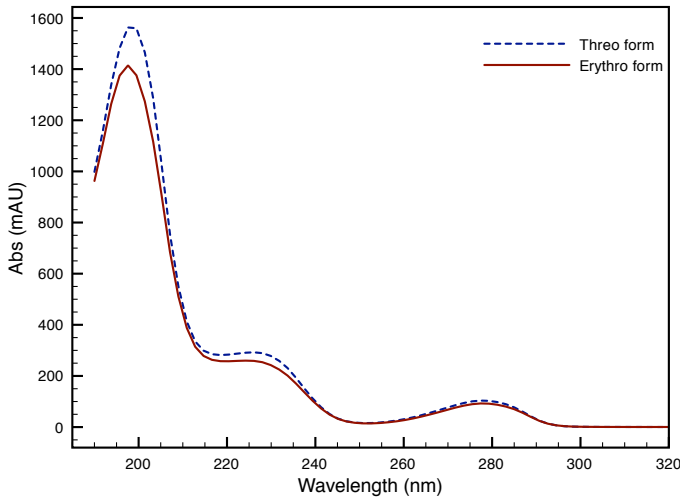


Figure 5.1: Absorbance spectra for the β-O-4′ structure in its threo- and erythro forms. The spectra were used for identification of β-O-4′ and its conjugates in HPLC using the diode array detector. Absorbance peaks are at around 198 nm, 225 nm and 278 nm.

5.2 Chromatographic separation

Chromatographic separation is used as a complement to other analytical methods in order to provide an additional analytical dimension. The separation dimension provides information about the analyte's identity and, implicitly, increases the resolution in samples with multiple analytes. Chromatography is also commonly used for preparative purposes.

5.2.1 Reverse-phase HPLC

In HPLC, the analyte is injected into a liquid flow (the *liquid phase*). The analyte interacts with the liquid as well as with the solid porous material, and the relative strength of these two interactions (the *distribution coefficient*; K_D) determine how much the analyte will be delayed on its journey through the column (*retention factor*). The more hydrophobic the molecule, the more it will interact with the stationary phase and the later it will elute from the column, with potentially large elution time differences between analytes of slightly different K_D.

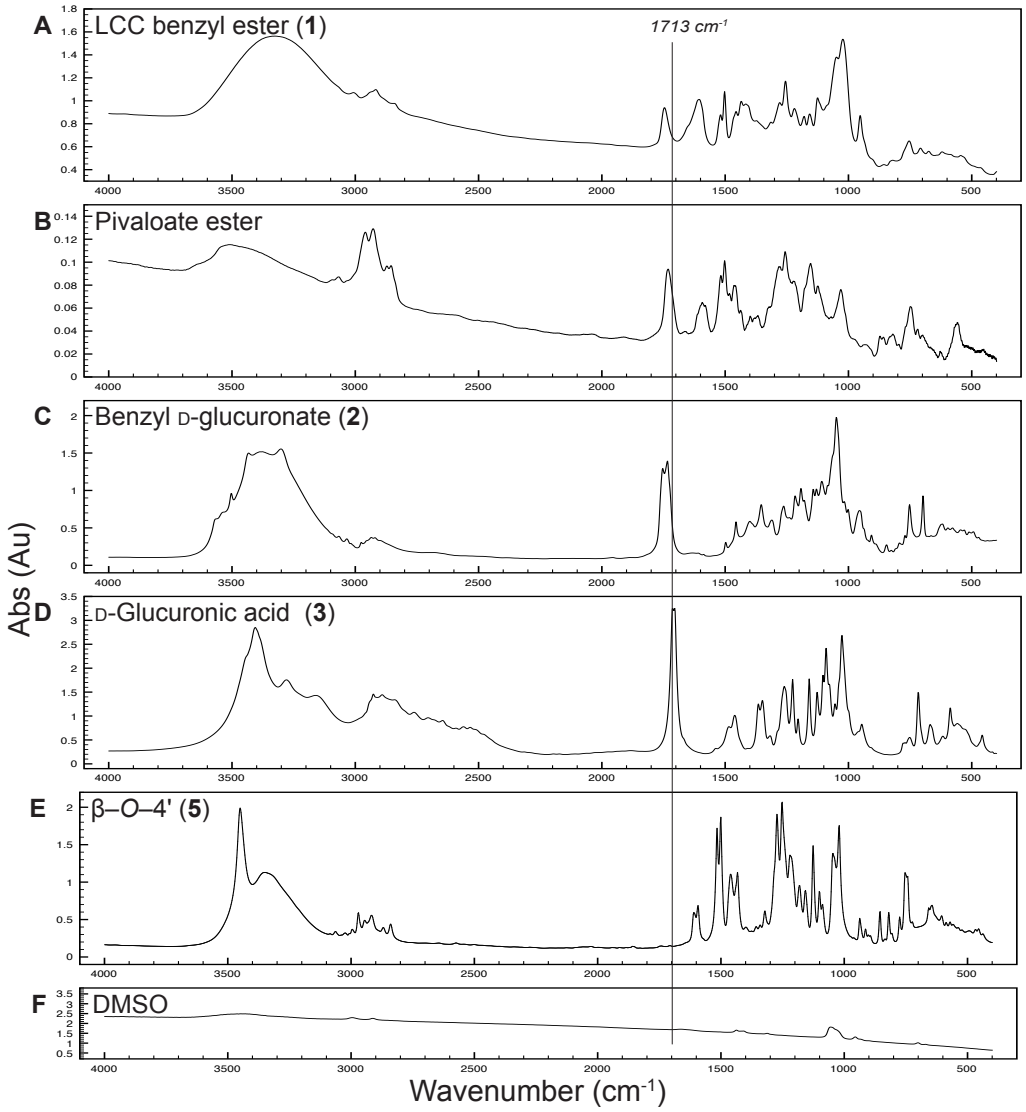


Figure 5.2: FTIR spectra. Identification of the synthesized (**A**) **1**, (**B**) pivaloate ester II, and(**C**) commercially obtained **2** serve as references for the wavenumber shift of (**D**) the GlcA carbonyl upon ester bonding. The line at 1713 cm^{-1} marks the carbonyl double peak in **D**. For reference, the spectra of (**E**) a β-O-4′ structure as well as (**F**) DMSO, with which the synthesized samples may be contaminated, are also presented. The spectra for **2** and **5** are also shown in Figure 6.1.

The chromatography itself only provides a means to separate different analytes (on the basis of their total structure and how it affects the K_D). To turn it into an analytical method, the column effluent is monitored by one or more detectors suitable for specific analytes, typically including a spectrophotometric UV detection for analytes with aryls or larger conjugated systems (Lambeth and Muhonen, 1994). Papers II and III used reverse-phase HPLC as part of the detection method in LC-ester assays, and in the preparative batch-mode adsorption chromatography setups of Paper I.

5.2.2 Ion Chromatography

Ion chromatography (IC) is often used as the trade name for high performance anionic-exchange chromatography with pulsed amperometric detection (HPAE-PAD), a particular variant of ion-exchange HPLC suited for carbohydrate analysis with high sensitivity. In IC (*i.e.* HPAE-PAD), the analysis is run at an elevated pH, where the carbohydrate primary hydroxyl is deprotonated and thus interacts with the anionic column material. The column effluent then reacts in an electrochemical cell with very precise amperometry, to give some level of selectivity while providing a very high sensitivity (Cataldi et al., 2000). Understanding *eg.* the principle of separation in IC is important for planning experiments and for evaluating data in the cases where standards are not available (*eg.* analysis after borohydride reduction, referenced below). IC analysis was used for the study of enzymatic hydrolytic products of AGX Paper IV and in the design a borohydride-reduction-based GE assay (sections 5.5 and 6.1.1).

5.2.3 Size-exclusion chromatography

In SEC the selectivity is provided directly by the porous matrix of the columns. The retention factor is determined by the size (*i.e.* hydrodynamic radius) of the analyte in relation to the distribution of the pore sizes of the matrix. In an ideal scenario, the retention factor of an analyte is determined only by the hydrodynamic radius, but matrix-analyte interactions are possible (*eg.*, van der Waals forces), slowing the analyte's passage and resulting in a decrease of its apparent size. This type of unwanted interaction was noted by I, where SEC was used to verify the presence of LC–complexes in a sample preparation by noting the concomitant decrease of lignin and carbohydrate molecular weight upon treatment with a hemicellulase mixture. This methodology is further discussed in Section 7.4.5 and its application to lignin analysis discussed in Baumberger et al. (2007).

5.2.4 Thin-layer chromatography

In TLC the mobile phase is wicked through a thin layer of dry stationary phase (*plate*) on which the sample as been applied. Analytes move differentially depending on the individual K_{D}s, resulting in different retardation factors (R_f), defined as the relative distance travelled by the analyte compared to that travelled by the solvent. Before the analytes reach the end of the plate, this is removed from the mobile phase and dried to fix the analytes.

To visualize colourless analytes, the dried plate can be *developed* by soaking or spraying with a chromogenic reagent, typically combined with heating or charring. By selectively extracting parts of the plate with solvent, the analytes can optionally be recovered after chromatographic separation. TLC was used as detection method for **2** in a GE assay (Paper III). A summary of TLC can be found in Spangenberg et al. (2010).

5.3 Mass spectrometry-based methods

MS is based on the interaction of ionic species with an electric field. Analytes are vaporized starting from solid or liquid phases. The mass-to-charge ratio (m/z) of each analyte can be calculated with high accuracy from the acceleration of ionic species in an electrical field onto a detector. In the process, analytes with different m/z are separated to yield a *mass spectrum* containing the signal intensity for each detected m/z. By various techniques, ions can be fragmented and the fragments subjected to an additional round of MS analysis (MS^n), giving information about the analyte's structure. MS^n have been used for the analysis of biological heterooligomers such as proteins (Lewis et al., 2006) and carbohydrates (Bauer, 2012) and, to some extent, for analysis of lignin and LCCs (Glasser et al., 1973; Iversen, 1985; JH Banoub and Delmas, 2003; J Banoub et al., 2015). MS^n can also be used as a detector after chromatographic separation, adding one or more analytical dimensions to the analysis, which has found widespread use in studies of the proteomics (Davis et al., 2001), metabolomics (Xiao et al., 2012) and glycomics (Zaia, 2008), but has attracted relatively little interest for use in lignin structural analysis (Morreel et al., 2010a; Morreel et al., 2010b; J Banoub et al., 2015)

MS was used for the analysis of residual structures in the end-point multi-enzyme assays of Paper IV. In addition, the mannanase-degraded GGM LCC-fraction of Paper I (see Section 3.7.1) was analysed in a preliminary assessment of LC-MS as the detection method for LC-bond assays. However, identification of fragments would have required a structural model of the analyte to which m/z:s could be matched. Such a model would Thus, a qualitative analysis was not possible.

5.4 Nuclear magnetic resonance (NMR) methods

NMR is based on magnetic excitation by radio-frequency (RF) of nuclei with integer quantum spin numbers in a strong magnetic field. For two of the nuclei that are relevant for lignin and LCC studies the NMR-active isotopes ^{1}H and ^{31}P, are the naturally dominant species. However, a natural abundance of 1.11% for ^{13}C limits NMR sensitivity for analytes that have not been been ^{13}C-enriched. By magnetic transfer between proximal nuclei, chemical bonding is reflected in the NMR spectra, and by adapting the excitation pattern, specific molecular features can be selected for analysis. An comprehensive overview of NMR for lignin analysis is found in Ralph and Landucci (2010) and (Argyropoulos, 2010).

In the context of LC-bond assays NMR complements chromatography and MS in that it provides structural information of a molecule regardless of its extent. In addition, NMR is an important technique for natural substrates as it can allow definite identification of a substructure such as an LC-bond in a complex matrix. ^{1}H and ^{13}C NMR are also fundamental as a method for establishing the molecular structure of synthetic substrates and assay products. To this effect, NMR was one of the analytical techniques utilised in Paper II.

5.5 Chemical derivatization methods

As outlined in sections 3.3 and 3.4, borohydride reduction and oxidation by 2,3-dichloro-5,6-dicyano-l,4-benzoquinone (*DDQ*; Watanabe (1989) and Imamura et al. (1994)) have been two important chemical derivatization methods used in the study of LCCs. These methods are generally combined with alkaline and acid treatments, as well as other derivatizations, to confer stability, selectivity, and to make them compatible with the analysis employed. Other methods, such as the selective cleavage of the β-*O*-4 lignin interunit linkage, in a process reported to retain LC-ether bonds (but not -esters; Ando et al. (γ-TTSA; 2015)), may also prove useful for LCC research.

In my work the borohydride reduction methodology, previously used for LC ester analysis (Comtat et al., 1974; Ö Eriksson et al., 1980; Das et al., 1981; Obst, 1982; Das et al., 1984b), was applied in an attempt to create a natural substrate GE assay based on GGM (Section 6.1.1).

Chapter 6

Glucuronoyl esterase assays

This chapter outlines and discusses the series of GE assays that were designed as part of this thesis work. GEs provided a natural starting point for LC-bond assay development, as it is a known enzyme, with little explored substrate specificity, while being expected to have LC-bond-breaking activity.

At the time when the work on this thesis started, the key questions regarding the activity of GEs were (see Section 4.1): i) Can GEs hydrolyse naturally occurring LC-bond esters? ii) How does GE activity depend on the structure of the aryl part of the ester substrate? and iii) What are suitable substrates and methods for assaying GE activity?

6.1 Assays developed as part of this work

6.1.1 Natural-substrate assays

In addressing the possibility of GEs hydrolysing naturally-occurring LC-ester bonds, I made several attempts at designing natural substrate GE assays. Tested substrates were mainly different HWEs and TMP fluids (Section 3.7.1), which were known to contain LC-ether bonds, but in which the presence of LC-ester bonds had not been demonstrated. Two detection methods were evaluated on *St*GE2-treated substrate, with un-treated and alkaline-treated substrate as references: FTIR analysis of the enzyme-treated material and borohydride reduction followed by hydrolysis and IC detection.

Natural-substrate GE assay using FTIR

While the FTIR difference spectrum between GE-treated and untreated material contained a signal that could be attributed to loss of the GlcA ester (a carbonyl stretch bond at ~1740 cm^{-1}), it was too small compared to the background (see Figure 6.1). There are three possible interpretations: a) GE did not hydrolyse GlcA LC-esters, b) there were no GlcA LC-esters present, c) the sensitivity of the method was too low compared to the number of hydrolysed bonds. As FTIR analysis of alkali-treated material has been proven as a method for the detection of LC-esters

(Obst, 1982; Takahashi and Koshijima, 1988; Singh et al., 2005), this method should nonetheless be a usable detection method for GE activity on natural substrate under the right circumstances (see Section 7.4.6).

Natural-substrate GE assay using borohydride-reduction and IC

As borohydride does not reduce the C-6 of non-esterified GlcA, GE activity should be detectable by the loss of 4-*O*-Me GlcA or the gain of its reduction product 4-*O*-Me Glc (**4** and **29**). However, while there were clear differences in the IC chromatograms between GE-treated, alkali-treated, and untreated material, a proper qualitative assessment of these differences could not be made. The IC chromatograms were complex and there were no IC standards neither for mGlcA nor 4-*O*-Me Glc available (**4** and **29**). In addition, the presence of LC-esters in the substrate was not known and the activity of GEs on these structures had not been established. Thus, a GE could not be used as positive control and the assay setup could only be verified by its successful outcome (see Section 7.2). Without any of these factors, the optimisation of the assay procedure, specifically adjusting the conditions of the borohydride reduction and of the subsequent hydrolysis step, would be an additional challenge.

While none of these efforts were conclusive, the practical experience was valuable and led to better understanding of how borohydride reduction (Section 7.4.4) and FTIR (Section 7.4.6) could be successfully used as detection methods for natural-substrate GE assays.

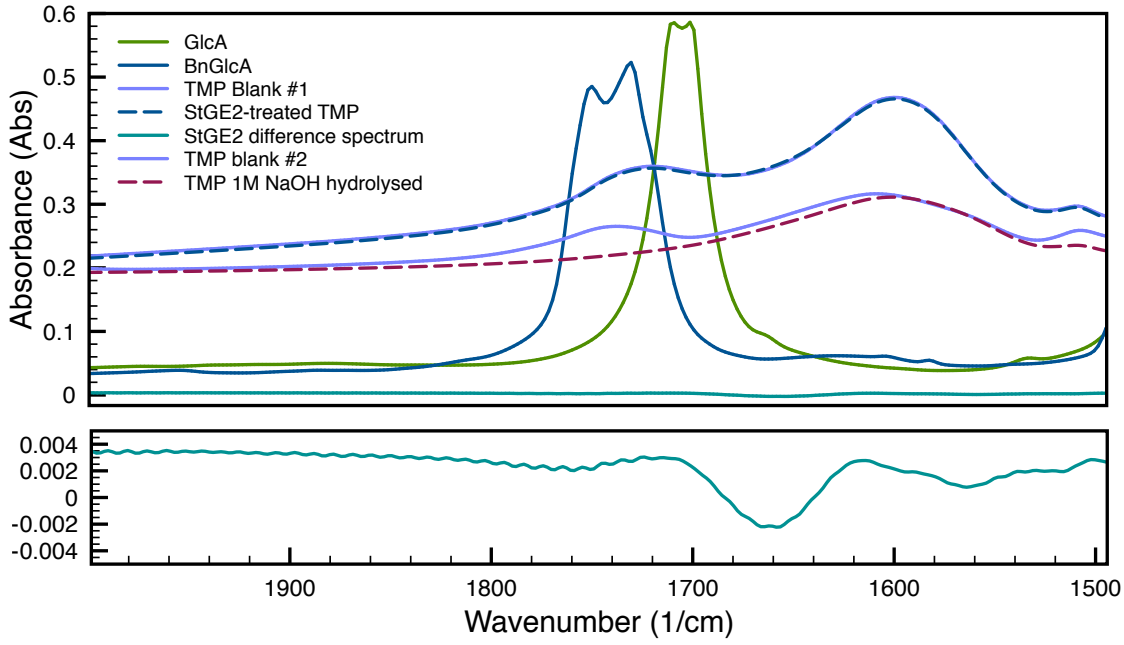


Figure 6.1: The carbonyl stretch region of FTIR spectra of TMP (Section 3.7.1) treated by alkali or *St*GE2 in two different experiments. In the lower frame, difference spectrum for the *St*GE2 sample has been magnified. The GlcA and BnGlcA spectra are included as reference. Spectra are linearly scaled to superimpose.

6.1.2 Synthetic-substrate assays

To investigate the chemo- and regio-selectivity of GEs towards the aryl moiety of the substrate, we decided to synthesize the β-O-4′ glucuronate (**1**) for use as a GE assay substrate. At the same time, the simpler BnGlcA **2** was used as a reference compound for method development to minimize use of the more costly **1** (see Section 3.4.1). At this point, GE activity on **2** had not been demonstrated, but it was inherent to the hypothesis of this work and was suggested by evidence of activity on **19** and **21** (XL Li et al., 2007; Špániková et al., 2007).

Synthesis of LC-ester substrates benefitted from the similarities between structure, synthesis, and assay detection methods for LC-ester and LC-ether substrates. Thus, development and demonstration of a synthetic LC ester assay would generate experience and information that could translate into developing synthetic LC ether assays. As will be shown, this was largely the case, even if applicability was somewhat limited by the specifics of ester bond properties (see sections 6.2.2 and 6.3).

GE assays using β-O-4′ glucuronate (1) as substrate

Reverse-phase HPLC proved to be effective for detecting both **1** and **2** and their hydrophobic breakdown products, β-O-4′ and BnOH (**5** and **6**, respectively) and activity of *St*GE2 and *Pa*GE1 could quickly be demonstrated on both **1** and **2**. However, both substrates were much less stable than anticipated and to sufficiently improve stability, the assay on **1** had to be carried out at pH 4, at which point the enzymatic activity could be expected to be significantly lower than at slightly higher pH (20% vs. 75% of max activity at pH 5; Topakas et al. (2010)). At this pH, kinetic assays could be carried out, with relatively high background hydrolysis despite efforts to minimize it, and resulted in a reasonable well-made estimation of kinetic parameters for *St*GE2 on **1** (Paper II).

In hindsight, the effects of self-hydrolysis and limited solubility on this type of assays (see sections 6.2.2 and 6.3) have become increasingly clear, and as this substrate can be expected to have a limited solubility in comparison to its apparent K_{M}, these two factors should be taken into account when interpreting kinetic parameters, as discussed in Section 6.3.1.

While use of **1** as a GE substrate was nominally successful it also clearly demonstrated that, in order to assay GE activity and kinetic parameters on similar structures, stability and solubility had to be much higher (5–10 times the actual K_{M} value; (Bisswanger, 2012b)). Thus, further experimentation on similar substrates, synthesized with the same methods as **1**, had to be abandoned. However, if sufficiently stable and soluble, β-O-4′ glucuronates are highly interesting substrates for GE assays.

GE assays using BnGlcA (2) as substrate

Compared to the use of **1** as substrate, the HPLC-based assay on **2** was significantly simpler, as **2** was both orders of magnitude more stable (Section 6.2.2) and commercially available. This created the opportunity to better optimize the developed assays and critically assess its properties.

As part of an ongoing collaboration with Prof. Christakopoulos and Dr Topakas of the National Technical University in Athens, Greece, similar HPLC-based detection methods for **2** and its hydrophobic hydrolysis product **6** were developed in both labs. These procedures relied on stopping the assay with acetic acid (see Section 7.4.2). Kinetic parameter assays for *St*GE2 and *Pa*GE1 were made by this method by my collaborators, and the data was analysed by me (Paper III).

Given the availability of commercial uronate dehydrogenase-based kits for detection of GlcA and α-glucuronidase activity, my collaborators and I developed coupled–enzyme GE assays using **2** as substrate (presented and critically evaluated in Paper III).

Based on TLC (Section 5.2.4) methods previously used for GE assays (Špániková and Biely, 2006; Nylander et al., 2016), I developed a qualitative TLC-based GE assay on **2** (Paper III). The advantages of this method are that it is simple, fast, requires tiny volumes, and can easily be scaled up from a few to a moderate number of samples. Moreover, this method could be further developed into a simple quantitative method (Section 7.4.3) with potentially better precision than the spectrophotometric method.

6.2 Critical evaluation of the developed assays

During the development of the GE assays based on the substrates **1** and **2**, several unexpected but partly foreseeable properties of these assay systems became apparent. These properties – stability and solubility of the substrates, and their dependence on pH, temperature and buffer composition – complicated analysis and assay interpretation. Consequently, systematic study of, *i.e.*, the stability of **2** was done at a late stage, and other well-warranted experiments investigating the substrate and assay properties are missing. Nonetheless, I have collected the available data to provide the following, preliminary but systematic, overview of the properties of GlcA ester substrates and the GE assays based on them.

6.2.1 Substrate forms and buffer interaction

In solution, **2** appears in at least two interconvertible species. This can be seen in Figure 6.2. At pH 5 and in unbuffered solution, **2** appears as a peak doublet separated by a horizontal stretch of augmented baseline. The two peaks display indiscernible absorbance spectra (not shown) and are separated by ~15% retention difference. Despite their distance, there is no baseline separation between the peaks, but instead a flat trough with the same absorbance spectrum as the main peaks, indicating that interconversion occurs during the analysis. Interestingly, at pH 6, this distinction collapses and there is a continuous transition from the first to the second peak. In addition, the hydrolysis product (**6**), eluting at 6.7 min, is visible as a distinct peak at pH 6.

As **2** is free to epimerize at the C-1 carbon, the species visible in Figure 6.2 could be either the open-ring form or α and β closed-ring forms. While identity of the peaks can not be ascertained from these data, at least one of the peaks must be a

closed form and the absorbance in the between-peak area (1/3 of the total area in the unbuffered sample) must originate from substrate that has, at least, transitioned once through the open-chain during passage through the HPLC column.

Figure 6.4 shows the time-resolved spectra of the pH 5 and pH 6 samples, revealing a much higher background in the pH 6 sample. Two possible reasons for the increased background around 200 nm could be: a) absorbance from the additional disassociated acid; or b) absorbance from **2** which, upon interacting with the buffer, changes absorbance spectrum. In addition, two absorbance ridges appear in the pH 6 sample around 270 nm and 370 nm in the 5–9 min interval. A spectral section at 5.5 min elution time (Figure 6.3) shows that, in comparison, there is only a weak absorbance band at 260 nm in the unbuffered and pH 5 samples. As larger electronic systems are required for higher absorbance wavelengths, we can speculate that the 270 nm and 370 nm absorbance bands in the pH 6 are caused by intermolecular interactions, either substrate–substrate interactions, *i.e.* aggregation (Section 6.3) or buffer–substrate interactions.

In conclusion, experimental data indicate that **2** is present in multiple forms, that the substrate interacts with itself or with the buffer, and that these effects are pH- and buffer-dependent.

6.2.2 Stability of GlcA esters

I initially assumed that the alkali stability of **1** and **2** was sufficient for practical experimentation at relevant pH ranges. This was supported by previous literature. Špániková and Biely (2006) performed *Sc*GE1 kinetic assays on **11** at pH 6 "to avoid spontaneous de-esterification of the substrates at higher pH". Wong et al. (2012) choose the same pH for all assays "because the ester linkages become unstable under more alkaline conditions". However, significant spontaneous hydrolysis of the substrate proved to be a main drawback when using **1** and **2** as substrates.

It was observed that **1** hydrolysed unexpectedly fast in aqueous solutions. The rate would decrease at lower pH and lower water content, but many factors appeared to have an influence. The instability of **1** was difficult to investigate further, due to the limited availability of substrate and the consequent issues with dissolving it in defined concentrations and storing stock solutions.

For the publication of Paper III, I investigated the pH dependency of the hydrolysis of **2**. Figure 6.5 shows a graph of hydrolysis rate at varying pH for two temperatures. The data in the figure show that there is a near-linear relationship between the hydrolysis rate and the hydroxide ion concentration at both temperatures in the tested pH range. The linearity shows that hydroxide ion hydrolysis is the main mode of hydrolysis with rate constants (k_B) of -4 s^{-1} mM^{-1} at 30°C and -11 s^{-1} mM^{-1} at 40°C (see Mabey and Mill, 1978, for a summary of ester hydrolysis kinetics). As individual hydrolysis rate measurements in the range of pH 3–4 display glacial hydrolysis rates ($k < -1 \times 10^{-7}$ at 4°C), it can be concluded that there is no significant acid- or water-catalysed hydrolysis of **2** at the pH range relevant for GE assays.

Having positively determined the mode of hydrolysis for **2**, I have re-examined data from experiments on **1** and estimated the apparent k_B for hydroxide hydrolysis

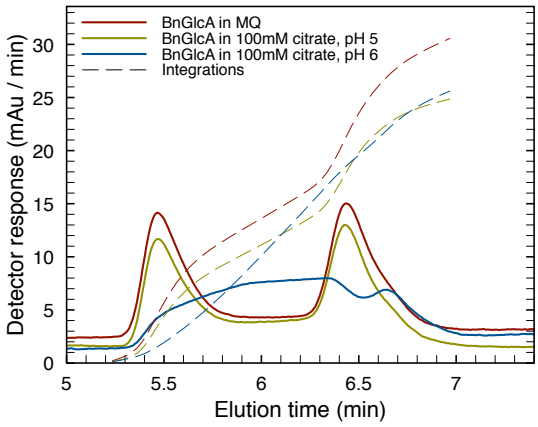


Figure 6.2: Absorbance chromatogram at 210 nm showing samples of **2** immediately after dissolving in water or in citrate buffer at pH 5 or pH 6. The two peaks at 5.5 and 6.5 min elution time are different forms of **2** whereas the peak forming in the pH 6 sample at around 6.5 min is the hydrolysis product **6**. The peak integrals are shown as dashed lines to validate that all three samples have the same total absorbance.

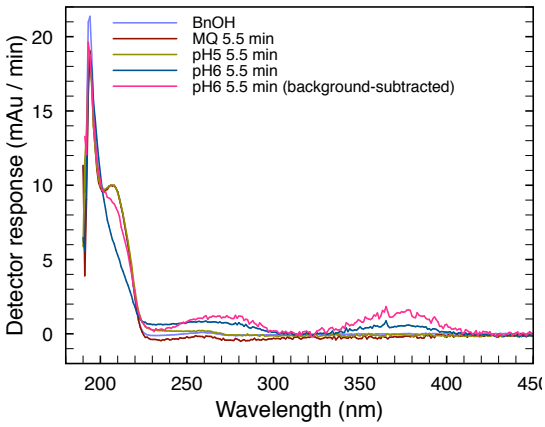


Figure 6.3: Absorbance spectra of chromatogram peaks at 5.5 min elution time in Figure 6.2. The absorbance spectrum of the product peak (at 6.5 min in a fully hydrolysed sample) is also shown. Absorbances have been normalized to superimpose at the 206 nm shoulder or at the 194 nm peak (pH 6 sample) for easier comparison.

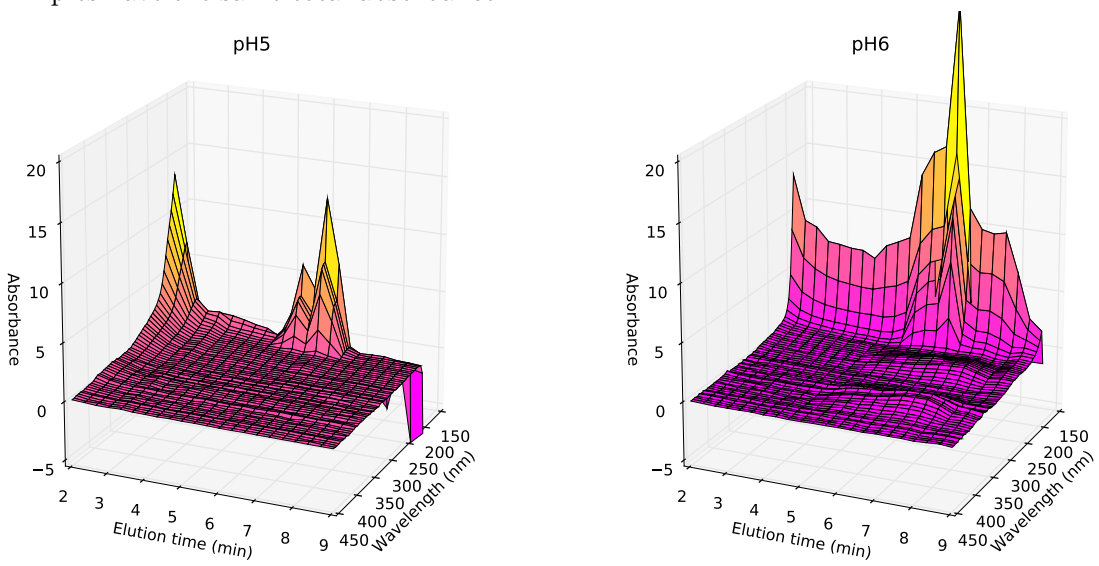


Figure 6.4: Full-spectra chromatograms of two of the **2** samples in Figure 6.2 at around 40 min after preparation. Both samples are prepared in 100 mM citrate buffer. No adjustments have been done except down-sampling.

of **1** to be ~100X larger than for **2**. While this is consistent with my observations, it is possible that the rate of acid- and water-catalysed hydrolysis is significantly higher for this substrate, as the linearity with respect to [OH^-] has not been checked.

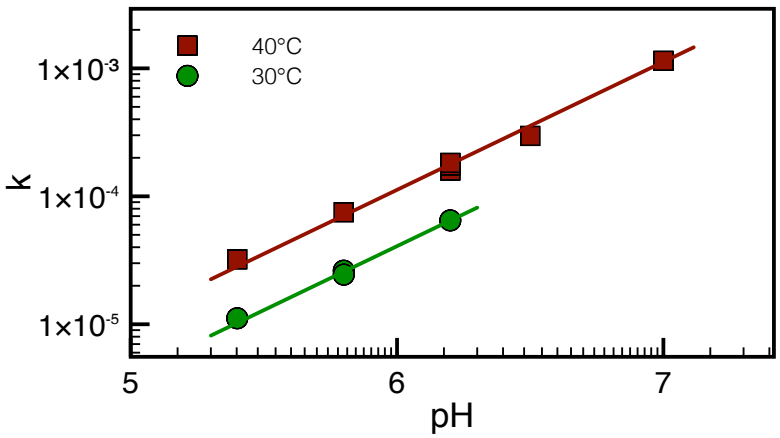


Figure 6.5: Full logarithmic plot of the hydrolysis rate of **2** (k) vs. pH at two temperatures. The measurements were made by repetitive HPLC injections of a solution of 1 mM of **2** in 100 mM acetate/phosphate buffers of varying pH, incubated in the HPLC at elevated temperature. At least six injections were made of each sample at 40°C (as well as at least three injections at 30°C). The inset lines show the regression for the hydroxide ion ester hydrolysis $k = k_B \times [OH^-]$.

6.2.3 The treatment of GlcA ester stability in the literature

The high instability shown by the esters **1** and **2** begs the question of why this feature has not received more attention in the literature on GE assays. In addition to the two articles mentioned above, the only comment on stability substrate ability GE that has come to my attention is a footnote in (d'Errico et al., 2015) about the reason for the instability of **17**. Three reasons may explain the limited evidence: a) the substrates used in literature are not as susceptible to hydrolysis as are **1** and **2**; b) in the experimental setups used in literature, the instability is lower; or c) the instability is there, but goes un-noted. I have put considerable effort into this question as it is particularly relevant for the design of GE assays.

The differences between **1** and **2** and other substrates used in GE assays depend on the nature of the esterified alcohol, the presence of a methylation at the C-4 hydroxyl, and whether the substrate has a C-1 substitution, (see Table 4.1 and the corresponding structures in the Appendix, *pp.* 149). Assuming that these factors cause the difference in hydrolysis rates, there are four ways in which ester hydrolysis can start, depicted in Figure 6.8. Deprotonations b) and c) are directly prevented by substitution at the corresponding hydroxyls and d) is indirectly prevented by any C-1 substitution, as this prevents the required ring-opening. In addition, susceptibility to hydrolysis can be modified by substitution effects, steric hindrance and alternative hydrolysis mechanisms (see discussion in Paper II on the spontaneous hydrolysis of **1** through the reverse synthesis mechanism).

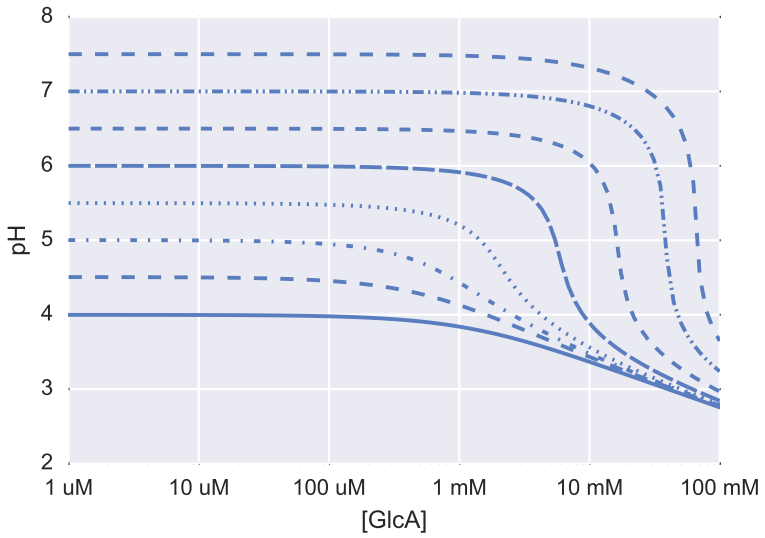


Figure 6.6: Calculated titration curve for increasing concentrations of GlcA in 50 mM phosphate buffers of varying starting pHs from pH 4.0 to pH 7.5. (Note that, to emulate a hydrolysis reaction, the buffer concentration is constant during the reaction.)

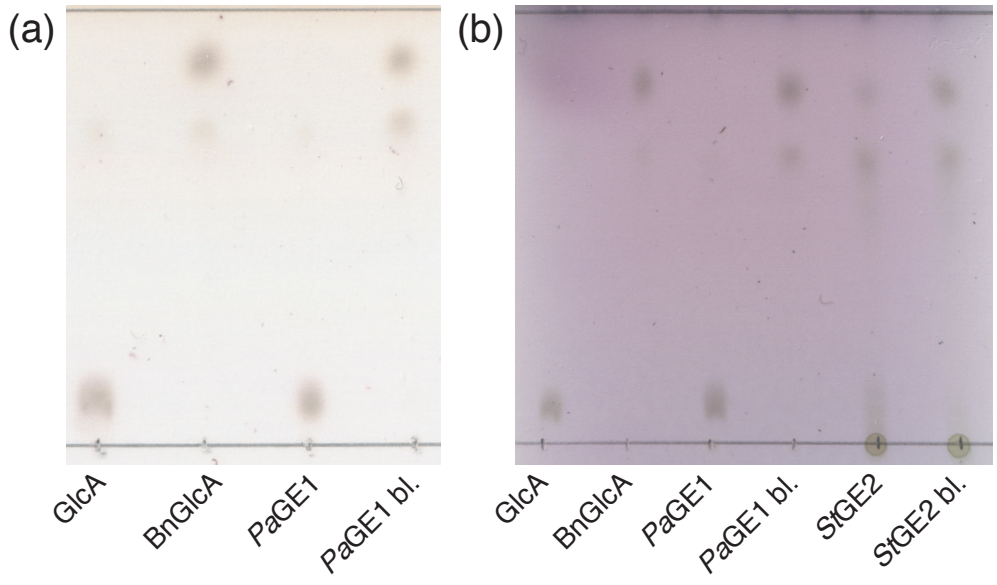


Figure 6.7: TLC analysis of a GE assay for *St*GE2 and *Pa*GE1 using boiled enzyme preparation (blanks) and **2** as substrate. Plate (a) was developed using the *N*-(1-naphthyl)ethylenediamine dihydrochloride reagent and Plate (b) using the H_2SO_4-anisaldehyde reagent. (From Paper III.)

The only data for hydrolysis rates on uronic acid esters appears to be the studies by JA Brown and Fry (1993) on *O*-galacturonosyl derivatives, in which half-times for hydroxide ion hydrolysis were comparable those we reported in Paper III for **2**. However, without data on hydrolysis rates for the substrates used in literature, no firm conclusion can be drawn on the effect of substitutions. However, if differences in substitution is the reason why substrate instability is not discussed in the literature, each of the points of substitution (C-1, C-4 and C-6) should confer sufficient stability for assaying as several substrates used in literature differ from **2** at only one of these positions.

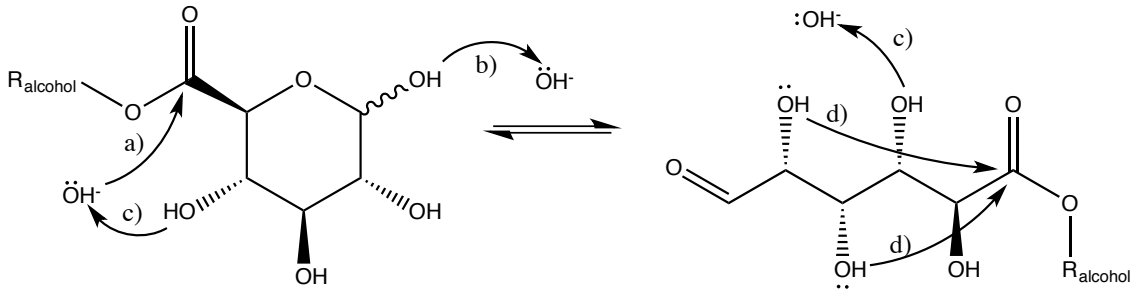


Figure 6.8: Possible initiation steps for alkaline ester hydrolysis: a) Nucleophilic attack on the ester carbonyl. b) Deprotonation of the C-1 hydroxyl. c) Deprotonation of the C-4 hydroxyl. d) Intra-molecular nucleophilic attack on the ester carbonyl by C-2 or C-3 hydroxyls.

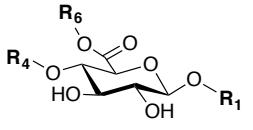
Assuming that the experimental design can confer substrate stability, and that this is the reason for the absence of discussion on stability in the literature, it would make sense to try to understand how to design GE assays to obtain increased stability. However, most assays in the literature use similar conditions to the ones I have tested, with 50 mM pH 6 sodium phosphate as the most common one (Špániková and Biely, 2006; Špániková et al., 2007; Duranová et al., 2009a; Duranová et al., 2009b; Vafiadi et al., 2009; Wong et al., 2012; d'Errico et al., 2015; d'Errico et al., 2016), though pH 5.5 (XL Li et al., 2007) and pH 7 (Katsimpouras et al., 2014) are also used.

It should be noted that phosphate buffer is not an ideal buffer for an acidifying reaction at pH 6, as its pKa values are ~2, 7.2 and >12 and the pKa of GlcA is ~3.2 (see titration curve of GlcA in phosphate buffer in Figure 6.6). Thus, for alkaline hydrolysis, the actual pH will decrease throughout the reaction and depend on the starting concentration of the ester, its k_B-value, the length of the reaction and the amount of enzyme present (and, for high substrate concentrations, the pH dependency of the enzyme's activity). The pH will in turn affect the alkaline hydrolysis rate.

Table 6.1: The computed aqueous solubilities of actual and hypothetical GlcA esters for use as GE assay substrates, as computed by the ESOL method (Delaney, 2004, see text).

	R_6	R_6 p	R_1	R_2	Log$P$	Sol	References
3					-3.13	21 M	
	Me				-3.04	15 M	
	Me		Me		-2.39	4.1 M	
11	Me			Me	-2.39	4.1 M	(Špániková and Biely, 2006; XL Li et al., 2007; Špániková et al., 2007; Duranová et al., 2009a; Duranová et al., 2009b; Vafiadi et al., 2009; Topakas et al., 2010; Wong et al., 2012; Charavgi et al., 2013; Katsimpouras et al., 2014)
	Me		Me	Me	-1.73	1.1 M	
2	Bn				-1.47	285 mM	Paper III
22	Aryl Propane	OH			-1.33	129 mM	(Katsimpouras et al., 2014)
17	Ph		Me		-0.95	118 mM	(d'Errico et al., 2015)
23	Aryl Propane				-1.04	96 mM	(Katsimpouras et al., 2014)
24	Aryl Propene				-0.96	83 mM	(Katsimpouras et al., 2014)
30	Bn		Me		-0.82	78 mM	(d'Errico et al., 2015)
	Bn			Me	-0.82	78 mM	
16	Me		pNP		-1.05	71 mM	(Špániková and Biely, 2006; XL Li et al., 2007; Špániková et al., 2007; Duranová et al., 2009a; Duranová et al., 2009b; Vafiadi et al., 2009; Topakas et al., 2010; Wong et al., 2012; Charavgi et al., 2013; Katsimpouras et al., 2014)
19	Aryl Propane	OMe			-1.03	65 mM	(XL Li et al., 2007; Špániková et al., 2007)
20	Aryl Propane		Me		-0.38	26 mM	(d'Errico et al., 2015)
31	Bn		Me	Me	-0.16	22 mM	(d'Errico et al., 2015; d'Errico et al., 2016)
	Aryl Propane	OMe	Me	Me	0.28	4.9 mM	
1	β-O-4′				-0.77	4.4 mM	Paper II
	Bn		pNP		0.52	1.4 mM	
32	β-O-4′		Me		-0.11	1.2 mM	
33	β-O-4′		Me	Me	0.54	333 μM	
34	β-O-4′	Me	Me	Me	0.84	166 μM	
35	β-O-4′	OBn			1.11	45.8 μM	
36	β-O-4′	OBn	Me		1.76	12.6 μM	
15	β-O-4′	OBn	Me	Me	2.42	3.4 μM	(d'Errico et al., 2016)

For compounds mentioned in the text, the first column gives the compound number. The following columns indicate GlcA substituents, with the "R_6 p" indicating the substituent for the *para*-position of arylic esters. The Log$P$ column gives the computed octanol:water partition coefficient (Wildman and Crippen, 1999, as implemented by rdkit) and the "Sol" column gives the computed solubility.



Common structure

6.3 Substrate solubility

During my work on GE esters, I observed several indications of substrate aggregation. One was the increased wavelength of absorbance seen in samples of **2** figs. 6.2 and 6.3 and discussed in Section 6.2.1. Similarly, TLC analysis from Paper III revealed a second, less pronounced spot of **2** at a lower R_f and with a lower intensity, depending on sample background (Figure 6.7). For the kinetic assays on *St*GE2 (Paper II and Section 6.3.1), substrate aggregation could explain our observations of a slightly lower background hydrolysis observed at increasing substrate concentrations. When I later attempted to determine the solubility of **2** by making saturated solutions at various temperatures and diluting them for analysis by HPLC, I could not achieve saturation, and the samples showed a concentration of 600 mM in the stock that was to be saturated. However, there was considerable variation between replicates, hinting that the stock solution was not a completely homogeneous solution.

As LCCs are amphoteric molecules, aggregation and micelle formation are not improbable, and several studies have discussed LCC aggregation (Yaku et al., 1979; Gradwell et al., 2004; Uraki et al., 2006; Westbye et al., 2007; Esker et al., 2009). Of these, the experiment by Yaku et al. (1979) is the most extensive. Accordingly, aggregation and micelle formation was studied in a natural LCC preparation by several methods, including gel filtration, conductivity measurements and by studying the interaction of LCCs with hydrophobic chromogens. The authors also studied micelle formation by looking at the wavelengths of the absorbance maxima, which decreased markedly when the LCC concentration was decreased below a critical concentration. These latter observations are in line with the interpretation that aggregation may be the cause of the spectral changes noted for **2** at pH 6 (Section 6.2.1).

Substrate aggregation would not necessarily be a problem in an enzyme assay if the enzyme was saturated. Indeed, Michelis-Menten-like kinetics can be observed by lipolytic enzymes acting on suspended droplets, where surface area is critical for saturation; (Marangoni, 2003). However, in the case of GEs, the Michaelis constant is often in the mM range (Table 4.1), and K_{M}-values up to more than 50 mM have been reported. Saturating substrate concentrations would therefore not always be feasible for synthetic LC-bond substrates like **1** and **2** (as the suggested **7** and **9**), both as a consequence of substrate availability and, possibly, as a consequence of limited solubility. If an assay is run below saturation on an aggregating substrate, the assay metric will also depend on how much aggregation there is in a certain sample (as it lowers the enzyme-available substrate concentration). Conversely, if micelle formation could measured and controlled, the reaction may work as well, or even better as in interface catalysis. After all, the presence of CBMs on several GEs (XL Li et al., 2007), the nature of their substrate and the fact that many of the measured substrate affinities require LC-bond concentrations that are difficult to create in the lab (Table 4.1).

While better techniques for studying micellar formation exist today (*eg.* Uraki et al., 2006; Olesen et al., 2014), I have not had any practical means to study LCC aggregation. To have a better understanding of the relative solubilities, I have made simple theoretical calculations of the solubilities of a number of actual and possible GE substrates (Table 6.1). The calculations are based on the logP value, as

calculated by the chemoinformatics toolkit rdchem (Wildman and Crippen, 1999, by the method of) and used the ESOL method (Delaney, 2004), but with updated factors, fit to a different training set[1]. While the tabulated solubilities should be interpreted with caution, the analysis is informative as per the effect of various substitutions. Of particular interest are methylations (and a benzylation), whose addition to **2** and **1** may improve substrate stability (Section 6.2.2). However, the calculated solubilities indicate that these substitutions have significant effects on solubility. While it is possible to measure activity on a substrate that has a solubility comparable to the K_{M}, it would be difficult to make correct quantifications under those circumstances.

6.3.1 Kinetic assays on unstable substrates

Estimations of kinetic parameters were made while studying enzyme activity on substrates **1** and **2** (papers II and III, respectively). Some reflections and insights that originated in those studies have general relevance in the context of LC-bond assays and will be discussed hereafter. They concern the effect of substrate stability and initial presence of hydrolysed substrate, as well as substrate solubility vs. K_{M} and the effect on certainty.

As discussed previously in this chapter, the GE assays are potentially affected both by the stabilities and solubilities of the assay substrates. In addition, they depend also on substrate concentration and buffer type, among other things. The enzyme adds another level of complexity, as it affects both its own hydrolysis rate and the background rate. In addition, while the hydrolysis of **2** or **1** is assumed here to be effectively irreversible, making the concentrations of hydrolysis products irrelevant, it cannot be assumed that the enzyme is not inhibited by the reaction products (specifically GlcA, which appears to be what the enzyme recognizes; (Charavgi et al., 2013)). Hence, instead of the simple first-order reaction, there are several additional factors that should be taken into account (Figure 6.9): i) substrate aggregation, which lowers the effective substrate concentration; ii) spontaneous hydrolysis of the substrate; and iii) product inhibition, which should be minimized in order for the results to be valid.

The system depicted in Figure 6.9 is difficult to analyse as there is a multitude of factors affecting it in various ways. The analysis is additionally complicated if the substrate concentration is below saturation, as both enzyme and background activities dynamically affect i) enzyme rate and ii) background hydrolysis rate (assuming the effect of a pH decrease on enzyme activity is negligible).

Given the uncertainty added to the system by the unquantified factors of stability, solubility and product inhibition, GE-assay construction should abide by the following steps. i) Critically assess substrate solubility. ii) Estimate background hydrolysis rate for the substrate. Select pH accordingly and choose a buffer that prevents pH drift. iii) Adapt the assay protocol to minimize hydrolysis before the assay. Always

[1]Dr Christos Kannas, "Estimate Water Solubility of molecules". Presentation at 2^{nd} RDKit UGM 2013. Available from github.com/rdkit/UGM_2013.git. Retrieved commit 20b875276d6db8516b44303fd7d0cf35ce1d646d

measure the amount of product present in the substrate stock or at the start of the reaction. If significant amounts of product (*eg.* GlcA) cannot be avoided, as may be the case for unstable synthetic substrates (*e.g.* **1**) that are difficult to purify, consider assessing the product inhibition. While synthetic LC-ether substrates could be expected to be significantly more stable than LC-esters, ether bonds of LC-ether models can still be susceptible to alkaline hydrolysis (Paper II and Taneda et al. (1987)). The advice given here is thus also applicable to enzyme assays for LC-ether scission.

In the work on the kinetic assay on **2** (Paper III), I made a considerable effort to make realistic estimates of the confidence intervals for K_{M} and k_{cat}: As the interdependency of these variables is high when the maximum substrate concentration compared to K_{M} is low, the usual procedure of estimating the confidence-interval by Student's t-distribution is no longer effective. While I show in Paper III how error estimation needs to be done when $\frac{[S]}{K_{\mathrm{M}}}$ is too low to clearly identify where enzyme saturation occurs, I would argue that, in light of the topics discussed in this chapter, this procedure is not necessarily relevant to GE assays. Rather, precise error estimation is misleading in this case, where a more immediate concern would be the application of a simple model to a system of this complex. In this case, assessing model validity by experiments and data analysis would be a more relevant effort. Also, a better response to the general uncertainty of kinetic parameter estimation in low solubility–high K_{M}–low stability situations would be to disclose more of the underlying data when publishing results, and avoid over-analysis. Especially as the details actually provide the basis to understand the system, regardless of the validity of the model or the interpretation of its parameters.

6.3.2 Spectrophotometric GE assay

A spectrophotometric coupled-enzyme GE assay on **2** was developed as part of the work on Paper III. The reaction system of the assay is complex as it involves an initial enzyme reaction followed by a detection reaction at elevated pH, at which the substrate clearly continues to hydrolyse. Figure 6.10 displays a graph of the absorbance development of the assay during the detection reaction.

The assay, thoroughly discussed in Paper III has a low detection threshold, making it useful for activity comparisons and screening (at least 4 mU/mL in the measured conditions). However, the pH-dependence of the spontaneous hydrolysis rate during the detection reaction means that the pH and buffering capacity of the sample can affect the absorbance readout during the detection phase.

These assays are relatively fast and simple but, in my experience, they are not particularly reliable. Hence, they should be used mainly for comparing the activity of similar samples, rather than as a method to infer general properties.

6.4 Substrate specificity

When I started my project on GEs, GE activity had been demonstrated on two aryl propane esters of GlcA , showing that it could act on GlcA esters of hydrophobic

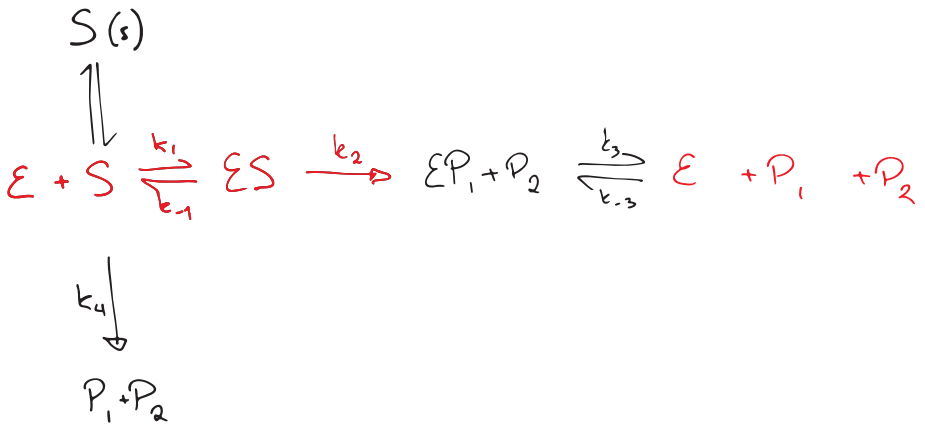


Figure 6.9: A first-order enzyme reaction (red), extended by the additional factors of substrate aggregation (top), spontaneous hydrolysis (bottom) and product inhibition (right) by one of its products.

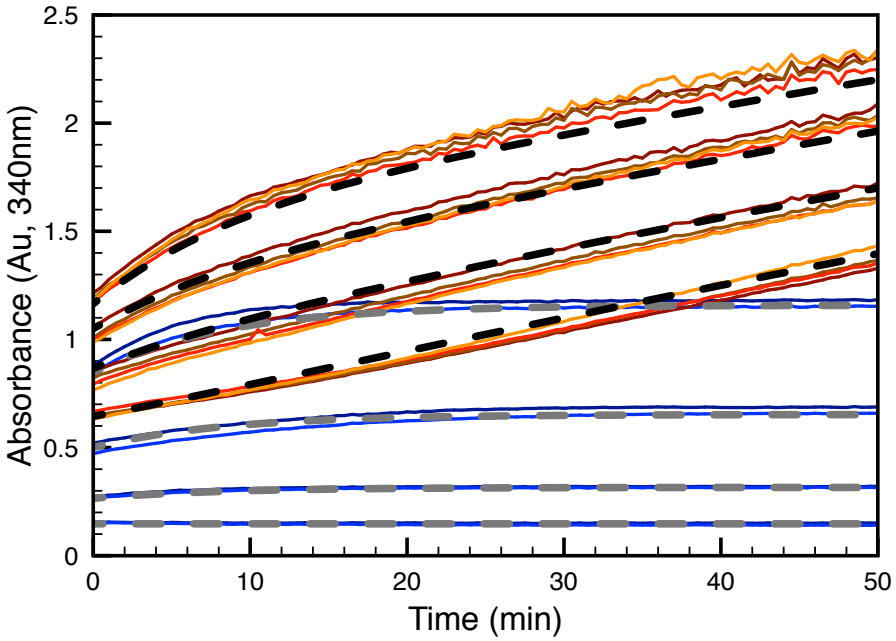


Figure 6.10: Results of an assay run of the spectrophotometric assay at four different concentrations of *St*GE2 in quadruplicates (0%, 5%, 10% and 15% culture filtrate diluted in boiled culture filtrate; brown–orange colours), with four GlcA standards in duplicates (blue colours). Enzyme samples (0–15% diluted culture filtrate). A computer model of the reactions was made and the results are shown as thick dashed lines, (black for the enzyme reactions and grey for the standards).

alcohols (Špániková et al., 2007), and on various methyl esters of GlcA (Špániková and Biely, 2006; XL Li et al., 2007; Špániková et al., 2007; Wood et al., 2008; Duranová et al., 2009a; Duranová et al., 2009b; Vafiadi et al., 2009; Topakas et al., 2010; Pokkuluri et al., 2011). It had been suggested that substrate recognition was primarily on the sugar moiety (Špániková et al., 2007) while selectivity for the aryl moiety of GE substrates was largely unknown.

My work on **1**, demonstrating GE hydrolysis on this diaryl GlcA ester model

(Paper II), the development of GE assays on **2** (Paper III), and the work of others on similar synthetic GE esters (d'Errico et al., 2015; d'Errico et al., 2016) unquestionably indicate that GEs are able of hydrolysing the kind of structures that are believed to give rise to LC esters bonds in nature. In addition, a recent study has indicated that GE can act in synergy with glycoside hydrolases for complete hydrolysis of a lignocellulosic feedstock (d'Errico et al., 2016). In light of this, what remains to be done is: a) develop a natural-substrate LCC assay, in which the actual enzyme activity can be positively identified and b) develop new assays. The latter should be improved in terms of: i) more stable substrates; ii) testing regio-selectivity further, by synthetic or naturally-derived materials; and iii) synthesizing stereochemically pure substrates for testing stereoselectivities.

Chapter 7

LCC assays

The previous chapters have been devoted to, first, outlining the enzyme assay and each of its components and, second, how I have combined these to create LC-ester bond assays for GE enzymes. This chapter will further the analysis of possible LC assay components and how they can be combined into different types of LC-bond assays. The chapter's analysis is done in light of the work that I have done and the experience that I have gained from my work on the LC-ester bond assays.

7.1 Development of LCC-bond assays

This chapter describes how LC-bond assays should be designed, in theory as well as in practice. The underlying assumption of this chapter is that assay development starts with the intention of finding enzymes that can break a certain type of LC-bonds (the *target activity*) for which no known enzyme activities exist (*eg.* LC-ether bonds), or for which enzymes have been discovered, but knowledge is scarce (*eg.* LC-ester bonds). Hence, some of the discussion in this chapter may not be relevant or valid further down the process from discovery to application, where more application-specific assays may be required.

The main focus points in this chapter deal with assay design: a) what components can be used to make an LC-bond assay; b) what can we use it for; and c) how do we do that, practically? This chapter also starts from the realization that, in order to take LC-bond-breaking enzymes from discovery to application, a range of different assays would be required for each chemospecificity, including both synthetic- and natural-substrate assays

7.2 Interaction analysis

In order to better understand the challenges of developing enzyme assays for LC-bond-breaking enzymes and to better direct our efforts, this section will attempt to detail the factors that influence this process and how they interact. A superficial analysis would note that the lack of LC-bond assays is due to the lack of its components, mainly LC-bond-containing substrates and LC-bond-breaking enzymes but also, in

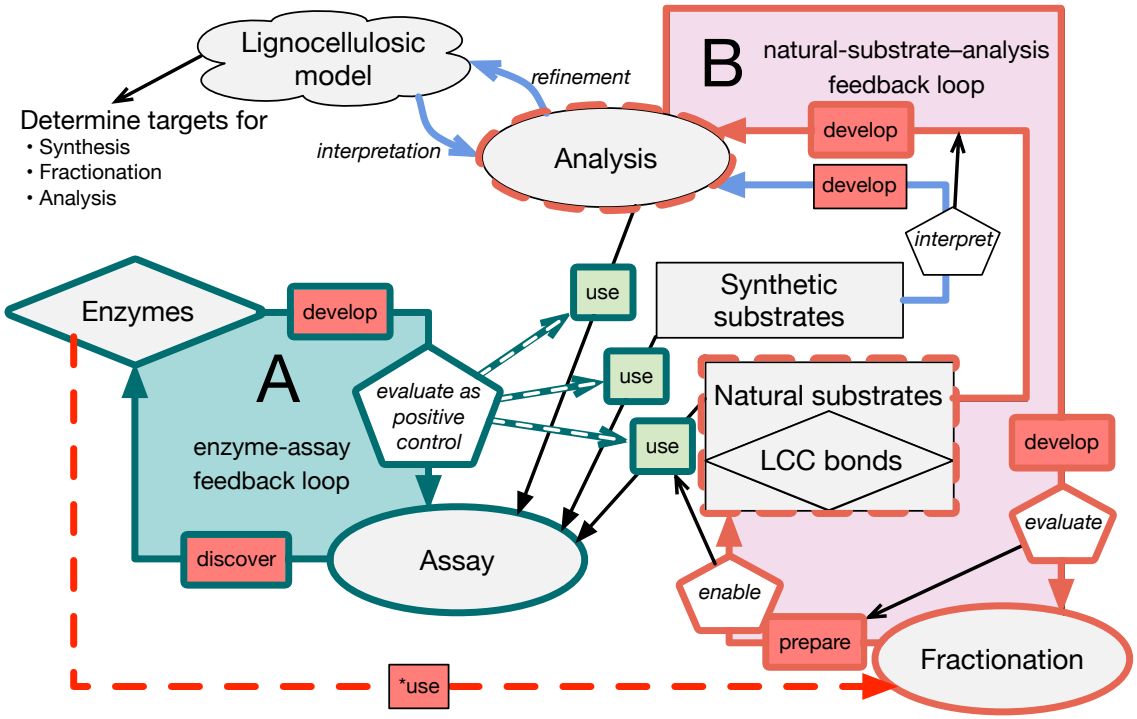


Figure 7.1: Interaction model showing factors relevant for the development of LC-bond assays and how they affect each other. Sub-figure A depicts the enzyme–assay feedback loop (Section 7.2) and how that loop affects the validation of analysis methods and assay substrates. Sub-figure B depicts the natural-substrate–analysis feedback loop (Section 7.2 and Figure 7.2) and how it affects assaying natural-substrate as well as the analysis methods, refining the understanding of the lignocellulosic model in the process. (*) As some enzymes can be used for selective degradation of lignocellulosics (see Sections 7.3.1), the discovery of relevant enzyme activities can create new possibilities for the preparation of substrates (and increase structural understanding), thus loosely connecting the two feedback loops (dashed line).

the case of natural LC-bond substrates, suitable assay detection methods. In turn, some of these shortcomings are rooted in our limited understanding of lignocellulosic material. However, this is a simplified account that does not include all factors. In addition, the complicating factors interact in various ways to create catch-22-like circumstances, further challenging the creation of LC-bond assays. In order to better understand the complications of LC-bond assay design, the main factors and their interactions are presented in figures 7.1 and 7.2 and analysed here. The objectives of the analysis are to give a realistic description of the difficulties involved in LC-assay design and to briefly discuss how these difficulties can be managed.

The enzyme-assay feedback loop

The most obvious interaction effect omitted from the analysis above is the catch-22-style negative feedback loop (A in Figure 7.1) stemming from the fact that enzymes are required as positive controls in the development of enzyme assays, while enzymes require enzyme assays to be discovered. The depiction in Figure 7.1 shows that while the enzyme's role in assay development is to validate the assay setup in its entirety, it

also validates both the substrate and the analysis method. Starting from a working (*i.e.* validated) assay, the components and conditions can then be changed, one at a time, with the assay outcome validating the alterations. Conversely, without a known enzyme to act as positive control, the choice of analysis, substrate and assay setup cannot be positively verified. The choice of components and setup of such an assay can not be verified until a positive sample is found (although a positive result is not sufficient for validation except in simple cases; see the first section of Chapter 2), which may not happen if the assay design is wrong. While the importance and consequences of having an enzyme as positive control may appear obvious, it is important to emphasize as it severely affects the uncertainty of the development effort.

The natural-substrate–analysis feedback loop

The most important factor, common to all assays, is the frequency of the target structure. The frequency must be sufficient at least to distinguish activity from non-activity using the chosen detection method. The target frequency can be quantitatively increased through enrichment procedures, however, these are also hampered by a set of issues.

The interplay of analysis methods and naturally-derived substrates, generated through various preparation methods, creates a second catch-22 feedback loop (B in Figure 7.1 and further elaborated in Figure 7.2). This is based on the fact that: a) LC-bonds cannot be observed or quantified in intact material, but require enrichment; b) enrichment of LC-bond containing structures (*target structures*) requires a fractionation method that is selective for the target structure; c) to evaluate a fractionation method, an analytical method capable of observing the target structures is required.

It should be realised that there exists sufficiently powerful methods for analysis and fractionation to allow enrichment for LC-bonds until they can be easily observed. Otherwise, LC-bonds would never have been observed in the first place. Nonetheless, the natural-substrate–analysis feedback loop will invariably constitute a practical problem when implementing these methods. Unless the methods are trivial, or there is substantial experience already, the only way out of the feedback loop would be by collaborating or trading with partners where the methods have already been implemented.

In the absence of a direct observation method, development of natural-substrate assays has to proceed without analytical confirmation. However, at some point in the assay development, the presence of the target structure in the substrate must be affirmed, and the earlier that happens during the following three development phases, the better:

1. Determining the applicability of a substrate to an assay (ie. verifying presence of target)

2. Enrichment of target structures in a substrate

3. Detection of activity in an assay

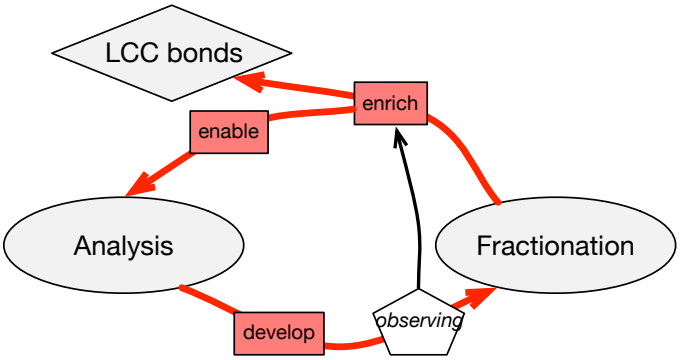


Figure 7.2: A expanded scheme of the natural-substrate–analysis feedback loop (Section 7.2 and Figure 7.1) showing how analysis methods can support the development of fractionation methods by quantifying the enrichment procedure. In turn, enrichment of LC-bonds enables them to be observed by the analysis method. As a benefit, observations of the enriched bonds can then provide the basis for interpretation of future analyses (part of the natural-substrate–analysis feedback loop of Figure 7.1).

To the very least, the target activity must be detectable in the last of these steps (or the assay would not work). If an enzyme known to hydrolyse the target bond is available, then, in this case, the assay setup relies entirely on validation by positive control. If not, the assay will only be possible to validate by its own positive result.

Another way in which the natural-substrate–analysis feedback loop can be bypassed is by using synthetic substrates. These generally allow for much-simplified detection methods. In addition, structural identification of natural substrates can be aided by analysis of homologous synthetic substrates (a similar methodology was employed in Paper II, where various GlcA esters homologous to **1** were analysed to aid the identification of the latter). This double-effect of synthetic substrates confers them an important role in the development of LC-bond assays.

The importance of the lignocellulosic model

One of the key factors of LC-bond assays derives from our understanding of lignocellulosic structures in general (the *lignocellulosic model*, of which Chapter 3 provides an overview), and how it relates to the target LC-bond and any sources of natural or synthetic substrates we have available. This understanding is directing both our selection of assay targets, including regio- and stereo-selectivities, as well as our choice of analysis methods. While the lignocellulosic model provides the framework for interpreting the results, the same results feed into the model to improve it (creating a positive feedback loop). Also, while analysis of synthetic substrates can help improve the lignocellulosic model, the validity of synthetic-substrate assays is determined by the validity of the model – *eg.* whether the bond that they mimic exists in nature or not (see Section 2.1). Consequently, the resolution and precision of the lignocellulosic model is an integral part of the development of LC-bond assays. This makes refining the lignocellulosic model a valuable side-goal when developing LC-bond assays.

7.3 Substrates for LCC assays

For a biochemist, obtaining assay substrates may be the biggest challenge in the design of LC-bond assays. Preparation of natural substrates requires equipment

and resources not commonly found in biochemistry laboratories, such as mills or specialized autoclaves for extraction. Also, the generally low yields of LCC preparations lead to large amounts of starting materials that have to be processed or small amounts of substrates in the end. While the synthetic route for substrate generation is simpler in some respects, it requires considerable experience. In addition, the large number of hydroxyl groups and the complex stereochemistry of synthetic substrates makes synthesis a challenge. However, once a substrate is available in sufficient quantities, designing an assay around the substrate may be the easier parts.

7.3.1 Natural-substrate preparations

There are numerous methods for generating natural LCC substrates. However, apart from my work on the Azuma procedure (sections 3.7.1 and 3.7.1) production of natural substrate has not been a focus of my work. Based upon my experience, production of natural substrates is practically challenging for the following reasons.

First, implementing a method for substrate preparation generally requires experience, particularly since there is a scarcity of methods to objectively evaluate the progress of the procedure. Methods requiring various types of milling appear to depend on the particulars of the setup – equipment size, sample loading, frequency, the presence of solvent or inert atmosphere, what ball is used in the ball mill and run time – and these factors are generally difficult to replicate (see Kleinert, 1970; Koshijima et al., 1972; Ikeda et al., 2002). Second, sooner or later, a method must be scaled up to produce target LC-bonds at a sufficient frequency to detect them. Third, many LCC-preparation methods have rather small yields, so large amounts of starting material and large-scale equipment may be required to get sufficient amounts of substrate for the intended assays.

Enzymatic treatment has been used extensively for the preparation of LCCs, usually with commercial enzyme mixtures or culture filtrates (for example Yaku et al., 1976; Ö Eriksson et al., 1980; Iversen, 1985; Du et al., 2014). Use of specific enzymes with defined enzymatic activities enables the production of more specific LCC fragments (Lawoko et al., 2003). Such an enrichment could start either from a carbohydrate-rich or from a lignin-rich LCC material. In the first, and more straightforward case, a material such as the LCC fraction (called HWE-XB) of the HWE of Paper I or a filtered TMP process fluid (Oinonen et al., 2013) could be selectively degraded by a selection of GGM-degrading enzymes (*i.e.* *exo*- and *endo*-mannanases, acetyl mannan esterase, α-galactosidase, β-mannanase and β-glucosidase). After a round of fractionation (by hydrophobic or size-exclusion chromatography or by dialysis) to remove the released monomers, relatively small LCCs should remain. For a sample such as the LCC fraction made in Paper I, the average lignin (guaiacyl unit) to carbohydrate ratio was approximately 1:7. The size of the resulting LCCs depended on how close to the lignin substitution the glycolytic enzymes could act. While I have not studied this for GGM enzymes, Paper IV shows (Table 1) that 8.3% (% weight) more xylose and 13.8% mGlcA was released with the addition of Agu115, corresponding to 0.83±0.37 released xylose per released mGlcA unit. Thus, in this case, it appears as if the concerted action of the hydrolytic enzymes used in Paper IV were capable of hydrolysing the glycosidic bonds adjacent

to the substituted xylose. While there is no evidence supporting the possibility of main chain hydrolysis so close to an LC-bonded side chain, neither in AGX nor GGM, hydrolysis of the main chain appears possible up to the substituted residue. Based on this argument and data in papers I and IV, Figure 7.3 presents a hypothetical structure for the LCC substrates that could be attained by selective enzymatic degradation of the LCC fraction. As with many LCC preparation methods, the yields for the LCC fraction of Paper I were meager, 1.2 g LCC preparation from 900 g of wood. With a selective polysaccharide hydrolysis, the resulting yield could be expected to be less than half of that.

The same principle of enzymatic degradation of an LCC substrate could be applied to the lignin portion of the substrate. Radical depolymerization does not appear to be suitable for the type of HWE and TMP samples discussed here (Oinonen et al., 2013; Oinonen et al., 2015). However, the β-aryl ether degradation (see Section 3.3) that has recently been applied by many research groups (Ferreira et al., 1993; Reiter et al., 2013; Ohta et al., 2015; Rosini et al., 2016) is an exciting possibility in this direction. It would allow for milled-wood lignin or for the high lignin-fraction of Paper I (HWE-LCC) to be used as starting material for a more defined LCC preparation.

7.3.2 Synthetic substrates

Synthetic substrates have an important advantage over natural substrates in that the substrates and products of the assay reaction are generally separable by chromatographic methods (*i.e.* by HPLC and TLC). In addition, if the LC-bond substrate includes aromaticity, there is a good chance that HPLC with UV-detection would be sufficient in terms of detection. Furthermore, evaluating the progress of the synthesis – analysis and quantification – is usually just a matter of routine for synthetic substrates. In addition, compared with the milling step of natural-substrate preparations, synthetic chemistry procedures may be less dependent on which equipment is available.

When selecting a structure to synthesize for the use as an LC-bond substrate in an enzyme assay, the most important factors are: i) representability; *i.e.* how well the substrate represents a natural structure of interest, as well as – for handling and assaying – the ii) stability and iii) solubility. Then, practical issues should be considered: iv) versatility for use in assays, *eg.* what are the possible assay setups and how can activity be detected, v) how to synthesise and purify the substrate and vi) the consequences of chirality for synthesis, purification and assaying. In addition, the novelty of the synthesis and the utility of this substrate as a standard for structural determination of lignocellulosics can be important factors increasing the value of the synthesis effort. As synthetic models for LC-bond structures are scarce, synthesis scale-up and distribution of the product to other parties may also be worth considering. The use of biosynthetic tools, *i.e.* enzyme-catalysed steps, can bring multiple advantages to the synthesis itself, as well as provide a way to share the efforts between biochemists and synthetic chemists while providing a novelty value.

Regarding the target structures to synthesize, analyses of stability and solubility are key steps of the planning phase. The discussion and analyses in Paper II, in sections 6.2.2 and 6.3 and in Table 6.1 can be of aid. The former provides a suggestion for what hydroxyls need to be protected for LCC ester substrates, as discussed earlier some of this may apply also to ethers (see also *eg.* Taneda et al., 1987; Sipilä and Brunow, 1991a), while the latter provide some basis for understanding how these modifications can affect solubility.

Given the predicted solubility span of GE substrates of over three orders of magnitude of Table 6.1[1], it is clear that substitution is a balance between i)~stability effects, ii) solubility effects and iii) easy of synthesis. For example, the simplest β-O-4′ substrate (**1**) has a predicted solubility of 4.4 mM, and each added methyl group reduced solubility 2–5 times (**32**, **33** and **34**; Table 6.1). In contrast, according to the solubility model, the addition of a benzyl group reduces the solubility 10 times (**15**, **35** and **36**). Hence, these additions are critical for enzyme assayability, especially at the K_M:s displayed by the GEs for substrates with β-O-4′ aryl components (Table 4.1).

Examining stability factors and prioritising substitutions by stability–solubility–assayability would be an important part of the design of synthetic substrates. In this process should be considered alternative substitutions (*eg.* carbohydrate residues where appropriate) and, as synthetic carbohydrate chemistry is intricate, how such substrates can be realised with a combination of commercially available or naturally-occurring starting materials, through the combined efforts of synthetic and biosynthetic chemistry. These efforts would be worthwhile for many reasons apart from the direct use in enzyme assaying, as synthetic substrates can be used to improve analytical methods and, in turn, the lignocellulosic model, on which model selection ultimately depends.

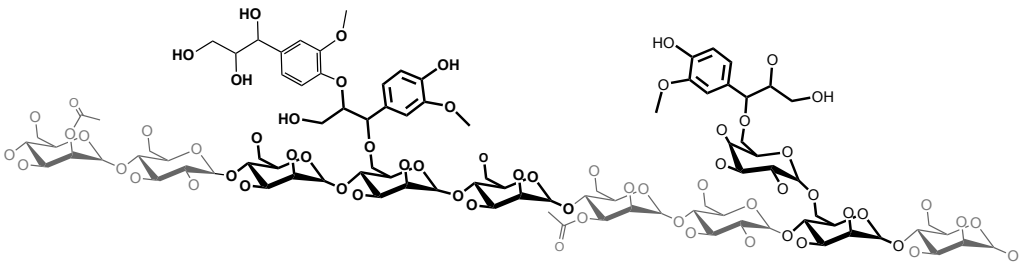


Figure 7.3: A general GGM LCC fragment. By selective degradation using enzymes, part of the fragment, represented by the shaded residues, can be degraded and removed, *eg.* by chromatography, leaving the potential LC-bond assay substrates drawn in black. The extent of the remaining structures would depend on a great many factors, such as the source of the material, the extraction procedure and treatment before and after extraction as well as the enzymes employed and the conditions of the reaction. Hence, the process can be challenging to reproduce.

[1]Note that the ESOL model does not take chirality into consideration, counts all substitutions by the same group as equal, regardless of substitution point, and have not been trained on LCC-like substrates specifically.

Detection	*appl.*	*Synthetic LCC ester*	**specific.**	*Natural LCC ester*	**Pre-treatment**
Coupled-enzyme	(S,Q)	(UV) UDH		(UV) Agu + UDH	
	(S,Q,K,C)	(UV) Agu + UDH (for glucuronopyranoside substrates)		(IC) Agu	
			α,E	(UV) LigL,N or LigD,O	
			α,S	(UV+HPLC) LigL,N or LigD,O + LigE,F	
HPLC-UV	S,Q,K,C	UV-absorbing substrates (*eg.* **2** and **16**)	(L)	Possibly, for substrates with defined lignin moities	Lignin depolymerization
TLC	S,Q,(K),C	Yes		Possibly (cmp HPLC-UV)	
IC		Only with pre-analysis separation or derivatization		Probably with BH_4 reduction (Section 7.4.4)	BH_4 reduction + complete hydrolysis
FTIR				Probably (Section 7.4.6)	
NMR			α,γ,(S)	Probably	Derivatization by P-substituent
MS-based					
		Synthetic LCC ether		Natural LCC ether	
Coupled-enzyme	S,Q, (K,C)	(UV) dehydrogenase		(UV) glycosidase + dehydrogenase	
	(S,Q,K,C)	(UV) glycosidase + dehydrogenase (for glycosidically linked substrates)		(IC) glycosidase	
			(α),E	(UV) LigL,N or LigD,O	
			(α),S	(UV+HPLC) LigL,N or LigD,O + LigE,F	
HPLC-UV	S,Q,K,C	For UV-absorbing substrates	(L)	Possibly, for substrates with defined Lignin moities	Lignin depolymerization
TLC	S,Q,(K),C	Yes		Possibly (cmp HPLC-UV)	
IC		Only with pre-analysis separation or derivatization		Probably with BH4 reduction (Section 7.4.4)	BH_4 reduction + complete hydrolysis
SEC				Possibly (Section 7.4.5)	
FTIR				Probably (Section 7.4.6)	
NMR			α/γ,(S)	Probably	Derivatization by P-substituent
MS-based			(α/γ)	Probably	Depolymerisation (lignin/carbohydrate)

Abbreviations: UDH, uronate dehydrogenase; Agu, α-glucuronidase; BH_4, borohydride; P-substituent, phosphor-substituent (*eg.* phosphitylation). The listed "Lig" enzymes are stereospecific enzymes that are part of the β-etherase lignin degradation pathway (see Helmich et al., 2016).

Table 7.1: Suggestions for possible LCC-bond assays on natural and synthetic LCC-esters and ethers. The application column suggests (for synthetic substrates) which assay types a certain combination would be applicable to, the specificity column indicates if the combination would any structural specificity (keys for these columns are tabulated in the right; parentheses indicate validity under certain conditions) and the pre-treatment-column indicates what pre-treatments would be required for a certain assay (the latter two columns applicable to natural-substrate assays only).

keys for the above table

key	Synthetic subst	key	Natural subst
S	screening	α	benzylic LCCs
Q	quantification	γ	γ-LC bonds
K	kinetic parameters	E	Some stereospecificity
C	characterization	S	Stereospecific
		L	Lignin-specific

7.4 Suggested LCC assays

Based on published evidence and personal experience with LC-ester assays, a number of setups for LC-bond assays appear feasible. These are listed in Table 7.1 and discussed in the following sections. Some of the strategies in the table have already been sufficiently discussed in previous chapters and will not appear further in this chapter. The go-through builds on the assumption that synthetic substrates are generally preferable over natural ones for assay types for which both substrate classes can apply. The discussion on MS-based analyses, as well as ^{13}C NMR are more speculative in nature and are therefore found in Chapter 9. A few LCC assays not specifically included in Table 7.1 are also treated in this chapter: coupled-enzyme assays (Section 7.4.1) and glycosidic LC-bond assays (Section 7.4.8).

7.4.1 Coupled-enzyme LCC assays

Coupled-enzyme assays make good candidates for LCC assays. As shown in Table 7.1, this assay type is applicable to both ether and ester LC-bonds, synthetic as well as natural substrates of many types. The setup used for the spectrophotometric GE assay of Paper III can serve as a starting point. Moreover, by replacing the uronate dehydrogenase (UDH) with another suitable dehydrogenase (for galactose, arabinose, glucose or xylose), assays on synthetic substrates such benzylic or β-O-4′ ethers to sugar residues would be possible (Section 3.3.1, Figure 3.4, and generalized structures **7**, **8**, **9** and **10** in the Appendix). Assuming that other dehydrogenases, like UDH, require an unsubstituted C-1 hydroxyl, for substrates with glycosidic bonds at this position, an additional glycosidase would be required. Interestingly, if the glycosidase is able to remove side-chains from polymeric substrates (such as the Agu115 used for removal of GlcA side chain to xylan in Paper IV), the coupled-enzyme assay would have the prospect of working on natural as well as on synthetic substrates. The coupled-enzyme assay is not as easily applicable when LC-bonds bind directly do the main chain, rather than to a side chain, of oligo- or polymeric substrates, as detection would require more elaborate enzyme combinations.

The specifics of the enzymes that are included in the assay setup would determine the possible applications and the range of usable conditions for the coupled-enzyme assay, but in principle it would be usable for everything from screening to kinetic characterization. However, a coupled enzyme-assay can be a complex reaction system, that may or may not allow continuous assays. In these cases, end-point assays would be advisable and boilable substrates would be preferable as they provide a simple method for stopping.

β-etherase pathway coupled-enzyme assays

It is quite possible that the stereospecific enzymes of the β-etherase pathway (see Section 4.3 and Helmich et al. (2016)) could be harnessed for coupled-enzyme assays. The first step in this pathway is an NAD^+-dependant oxidation by a dehydrogenase that acts on and is specific for the stereoconfiguration of the Cα of the lignin diaryl unit **5**. This activity would not be available on α-LC bonded (and perhaps not even for γ-LC bonded) diaryl units. With a suitable substrate, these enzymes would provide chemo- and stereospecificity as well as a straightforward detection method. While this may work on synthetic as well as on natural substrates, the

latter would have to be pre-treated to remove non-LCC-bonded diaryls, *i.e.* lignin (see Section 7.3.1).

It may be possible to add one more level of stereospecificity to this assay system by including the next step in the β-etherase pathway, a glutathione-*S*-transferase specific to the stereoconfiguration of the Cγ of the lignin diaryl, though another detection method would be required for this reaction step. One advantage of this system is that the substrate is the main difference between the LC-ester and LC-ether assays, thus allowing development of an LC-ether assay by way of an LC-ester one so that existing GEs can be used as positive controls for several critical development steps (*cf.* "The enzyme-assay feedback loop" of Section 7.2).

Complex coupled-enzyme assays

An assay setup that could be useful for LCC assays on natural substrates are the multi-step coupled-enzyme reactions that are often referred to as “synergy assays”. A straightforward application of this type of assay was presented in Paper IV, which studied the effects of side-chain removal for the enzymatic degradation of the main-chain of an AGX substrate. The setup used in this study could be extended to an LC-bond assay, if used with a substrate where LC-bonded side chains (or LC-bonded main chain) significantly impedes substrate hydrolysis (this was not the case for the sodium-chlorite extracted xylan in this study). For a LCC xylan substrate, the same combination of enzymes – α-glucuronidase, α-arabinofuranosidase, *endo*- and *exo*-xylanases (see Figure 4.4) – in combination with an LC-bond-degrading enzyme would be a possible setup for an LC-bond assay.

The complex coupled-enzyme assay can also be used to investigate features of the substrate. For example, the MS analysis in Paper IV (Figure 3, *pp.* 137) shows that three oligosaccharides are not degraded when all enzymes are present: a tetra-pentose, an aldopenturonic acid and a di-aldohepturonic acid. Several hypotheses can be formed from these simples observations: i) The Ara*f* sidechains can either be on adjacent or on the same Xyl (as a 2,3-diAra*f*). The first case would indicate that the xylosidases can cleave up to the substitution and, in the second case, that they can cleave up to the substitution on of the sides only. ii) The lack of aldotetrauronic acid and the presence of aldopenturonic acid indicates that mGlcA and Ara*f* are either not adjacent, or that they are, and that the xylosidases cannot cleave up to the substitution. iii) Two mGlcA can prevent degradation of five xylose residues, if we assume the assertion that mGlcA and Ara*f* are not adjacent. Several alternative interpretations are also possible, but the conclusion is, that this type of study is informative both on substitution patterns in the substrate, as well as on enzyme regioselectivity.

While the complex coupled-enzyme setup makes a straightforward assay, the LC-bond frequency must still be sufficient for detection of LC-bond scission, which may require enrichment. On the other hand, there is an amplification effect involved when an LC-bond blocks an oligomer from depolymerization as each additional monomer that is released as a consequence of LC-scission adds to the signal (if included in detection). For GEs, a study using a complex coupled-enzyme assay on the industrially relevant substrate corn fibre was recently published (d’Errico et al.,

2016). The study showed that inclusion of GEs in the combination of hydrolysing enzymes had the potential to increase monosaccharide yield. Increases in the range of 5–20% were reported, but no analysis on statistical significance was made.

7.4.2 HPLC detection in LC-ether and -ester assays

As an assay detection method, reverse-phase HPLC separation with UV-absorbance detection is generally applicable to any assay on a UV-absorbing synthetic substrate. HPLC-UV can be applied for stopped assays or for following a reaction by periodic sampling, as was done to measure the hydrolysis rate for **2** (Section 6.2.2). Thus, HPLC-UV is applicable to all types of LC-bond assay, though some care is required for assays on impure enzymes or culture-filtrates not to foul the column or to introduce UV-absorbing species influencing the analysis.

HPLC-UV detection could also be applicable to assays on naturally-derived substrates with small and well-defined lignin structures, as long as the substrate structures always have a distinctly smaller K_{D} than the lignin moiety of the product, when it has been cleaved off the LCC substrate. Such a setup could be useful as it creates a de-facto multiplexed assay, *eg.* for assessing the relative cleavage of different side-chain enantiomers in an HPLC setup with some enantiomeric separation.

There are probably other HPLC systems besides reverse-phase HPLC-UV that could be used, depending on substrate. However, detector sensitivity is likely an important aspect as various factors can reduce analyte concentrations: low availability for custom-synthesized substrates (as in Paper II), low product concentrations in kinetic studies, requirements for pre-analysis dilutions (as for unstable substrates in Paper II and Paper III). Low detection sensitivity may be countered by larger injection volumes and higher-capacity columns, but only to a limited extent; thus low-sensitivity detectors such as RI would generally not be applicable.

7.4.3 TLC for semi-quantitative LC-ether and -ester assays

TLC has been used for analysis of GE assays on various substrates (Špániková and Biely, 2006; Wong, 2006; Katsimpouras et al., 2014; d'Errico et al., 2015), including the assay on **2** that I presented in Paper III. TLC is simple, cheap, fast and generally suitable for model–compound assays on many samples in parallel, providing results in a very short time, typically an hour. TLC assays are suitable for a rough screening of activity in many samples and are thus suitable for following protein expression or purification; when screening for many different activities in parallel, or for quick estimations of relative activities when testing enzyme stability, or the effects of pH and temperature on the enzyme activity.

While TLC is often used as a qualitative analysis it is also routinely applied as a quantitative method. For this purpose, and for preparatory TLC, automated systems are available. In addition, it is reasonable to assume that a manually performed TLC can be made sufficiently quantitative to replace other assay detection methods: precise sample spotting, optimized protocols and densitometric analysis of scanned plate images, where, specifically, colour fading and scanner exposure must be actively managed.

Extending the TLC-based GE assay on **2** in Paper III to a quantitative assay would provide a number of advantages of the coupled-enzyme spectrophotometric assay on the same substrate. Subjecting a sample to analysis would effectively stop substrate hydrolysis, thus reducing spontaneous hydrolysis as well as allowing the reaction to be followed by TLC for periodic sampling of a reaction (only ~1 μL is required per sample). In addition, the TLC analysis would require much smaller reaction volumes than the spectrophotometric assay in cuvettes, while being easier to set up for small sample series compared to the plate-reader version of the spectrophotometric assay. Also, the TLC assay does not suffer the complex reaction dynamics that the coupled-enzyme reaction does.

TLC analysis would be applicable to synthetic LC-ether bond substrates in the same way as for the ester substrates, with the same advantages, bar the ones related to substrate stability. It could also be applicable to assays on natural substrates if one of the moieties, liberated upon LC-bond scission, is sufficiently well-defined to be identifiable in the TLC analysis.

7.4.4 IC analysis in LCC assays

As the analytical power of IC decreases with increasing oligomer length, IC is mainly applicable to analysis of hydrolysed or relatively well-defined samples (in terms of oligosaccharide length and branching). As such, it is mainly useful for analysis in assays using chemical derivatization or for assays based on partial or complete enzymatic hydrolysis (Section 7.4.1). As IC columns are sensitive to hydrophobic contaminants, substrates with considerable amounts of lignin should not be analysed without prior fractionation (which could, if included, both increase and decrease selectivity, depending on circumstances).

In conjunction with IC, borohydride reduction as attempted in this work (Section 5.5) could be a valuable analytical technique for GE assays as it would allow for identifying LC-ester-containing substrates (by comparing the amount of reduced product in alkali-treated and untreated substrate). Information about the presence and esterification of GlcA are key metrics in evaluation and in preparation of natural substrates for GE assays. This development would, however, require IC standards for at least one of mGlcA or 4-*O*-Me Glc (**4** and **29**, respectively) and ideally both (see Section 5.5).

Once established, the application of borohydride reduction and IC analysis to a substrate that had already been verified to have LC-ester bonds (*i.e.* by the same technique), would constitute a reasonable GE assay, similar to what I attempted to implement (Section 5.5). In practice, the procedure is better suited for smaller samples series as it requires autoclaving in suitable glassware and as the reagents are reactive and have to be handled accordingly.

7.4.5 Size-Exclusion chromatography

When size-exclusion chromatography (SEC) was used in Paper I to demonstrate the presence of LCCs in one of the material fractions, interactions between the lignin analytes and the column matrix were observed in the form of lignin model

monomers eluting after carbohydrate monomers of similar masses. While this effect may be ameliorated by applying protocols for lignin SEC, optimized for reduced interactions (*eg.* Ziebell (2008)), the resolution attainable by SEC may not be improvable to the point where analyses become either qualitative or quantitative. It can, however, be used to study the polymerization (Oinonen et al., 2013) or depolymerization (Paper I, Lawoko et al. (2006)) of LCCs without giving detailed structural information. As such, it could be used either to assess LCC-degrading enzymes (on partly characterized substrates) or to investigate the presence of LCCs in substrates (by partly characterized enzymes). However, given its caveats, it may not be the method of choice except under specific circumstances.

7.4.6 Natural-substrate FTIR-based GE assay

In the case of FTIR, the inconclusiveness of my efforts to construct a natural-substate GE assay (Section 5.1.2) was due to two factors. i) GlcA esters in the substrate could not be identified with certainty due to limited resolution and absence of an alternative method of quantification. ii) The difference in signal between the untreated and GE-treated substrate was too small in comparison to the inter-sample variation to infer activity while alkaline hydrolysis resulted in massive changes to the substrate (Figure 6.1). However, these efforts could probably be addressed to make FTIR a viable method for detection in natural-substrate GE assays. Solid-phase FTIR resolution appears to be dependent on sample preparation, and would improve with experience. Better reproducibility may thus improve both resolution and inter-sample comparisons, increasing the significance of smaller differences. Using standards (such as the samples in Figure 5.2) and a substrate in which carbonyls or esters and glucuronic acids could be quantified by other means (*eg.* ^{31}P NMR, Section 7.4.7) would increase confidence in peak identification, as would using a positive control.

In addition, real-time FTIR with multi-variate data analysis (applied in, *eg.* Richards et al., 2004; Kumar and Barth, 2010; Baum et al., 2013; Merayo et al., 2013) could allow for a continuous GE assay on natural substrates that would yield a substrate-product differential spectrum. However, as Figure 6.1 shows, a high resolution around 1740 cm^{-1} is required to resolve the interconversion of the carbonyl from ester to carboxyl. In addition, a time-resolved spectrum shows what gradual changes are brought about in the material, as opposed to being caused by the addition of signal sources or to differences in the preparation of samples.

7.4.7 ^{31}P NMR for general LCC analysis

^{31}P NMR analysis of hydroxyls in lignin and carbohydrates, after derivatisation by phosphitylation reagents, is a structural analysis method applicable to LCC substrates.

Phosphor has the advantage that its main isotope (100%) is magnetically active. Phosphitylation, the derivitisation of hydroxyls with phosphite esters, has been available as a method for the structural analysis of lignin for the last decades (Schiff et al., 1986; Argyropoulos et al., 1993; Jiang et al., 1995; Argyropoulos, 2010). Using a suitable phosphitylation reagent, resonance shifts in ^{31}P NMR for various types of

lignin hydroxyls can be resolved and distinguished from the shifts of carboxylic acid hydroxyls in a quantitative way (Wroblewski et al., 1988; Jiang et al., 1995). As reagent, 2-chloro-4,4,5,5-tetramethyldioxaphospholane (Granata and Argyropoulos, 1995) has been extensively used for lignin (Argyropoulos et al., 1993; Crestini and Argyropoulos, 1997; Pu et al., 2011; Oinonen et al., 2015) and recently also for tannin (Melone et al., 2013).

The selectivity and quantitative nature of ^{31}P NMR of phosphitylated materials could be used for various purposes in LCC assays. Primarily, it can serve to characterize potential substrates. This would be especially useful for GlcA ester assays, where the hydroxyl of free carboxylic acid can be quantified. In theory, the method could be used to obtain structural information and give proof to enzyme activities in other types of assays but in practice the carboxylic acids are the most obvious targets - the resolution of distinct lignin hydroxyls is limited, judging from the spectra of naturally-derived lignins in the literature. The ability to quantify free carboxyls, renders ^{31}P NMR a feasible detection method that has analytical selectivity for ester substrates, though it may not reveal to what those esters are bonded. The sensitivity for this method is dictated by the frequency of LC-bonded vs. non-LC-bonded carboxylic acids in the substrate.

7.4.8 Glycosidic LC-bond assays

While specific phenol glycosidic LC-bond-degrading enzymes (phenyl glycosidases) may exist, it is also possible that these bonds are simply hydrolysed in nature by glycosidases, that generally accept substrates with phenolic aglycons. However, if specific phenyl glycosidases exist they would be preferable over other glycosidases for enzymatic de-lignification, as the latter would not be specific to LC-bonds, but also be active on bonds between monosaccharides. In this scenario, the carbohydrate glycosidases would be selective for the structure of the sugar moiety, whereas the phenyl glycosidases would be selective for the structure of the aglycon, possibly in conjunction with specific carbohydrate structures.

Keeping in mind differences in specificity, the standard synthetic-substrate assays employed for glycosidases, based on 4-nitrophenol glycosides (Bisswanger, 2012a), could constitute a simple spectrophotometric assay for glycosidic LC-bond-degrading enzymes. They nevertheless present two caveats: i) the phenyl glycosidases may not recognize the 4-nitrophenol as a substrate and ii) the assay will not be specific for the phenyl glycosidases. Both circumstances make the 4-nitrophenol substrates unsuitable, at least for screening assays, in the enzyme-discovery phase. To address these two concerns, we would need to synthesize a substrate in which a more specific phenolic structure (*i.e.* guiacol, syringol or derivatives) is conjugated to a sugar for which we would not expect excreted glycosidases (*eg.* epimers like D-talopyranoside or D-ribopyranoside). However, constructing these assays and screening for LC-bond specific phenyl glycosidases is a risky venture as i) it is not known if this type of enzymes exist and ii) it may be difficult to validate the assay properly unless enzymes known to hydrolyse the substrate exist, in which case such enzymes would provide false positives.

Chapter 8

Conclusions

Despite intensive study, understanding of lignocellulosic structure and structural variation is far from complete. Production and analysis of natural substrates requires experience and expertise and the same is true for synthetic assay substrates, which must be custom-made for the assay application. LC-bond-breaking enzymes that can act as positive controls when developing assays are lacking, and though GEs are a strong candidate, LC-ester hydrolysis in natural substrates is yet to be demonstrated.

By this work, the knowledge of the substrate specificity of GEs and their affinity for arylic glucuronate esters was expanded by showing that GEs have activity on the LCC model compound β-O-4′ glucuronate (**1**) and on the commercially available substrate benzyl glucuronate (**2**). Both models display high rates of hydroxide ion hydrolysis, which has to be accommodated for when designing and interpreting GE assays on these substrates.

The synthesis of LC-bond model **1** and the investigations on the stability and solubility of **1** and **2** have increased our understanding of how LC-ester assays and -assay substrates should be designed and showed the importance of *background hydrolysis*, *solution acidification*, *solubility*, and *low substrate affinity* to the design of GE assays. This was demonstrated by the design of a series of GE assays that use **2** as substrate and that are applicable to screening, quantification and to kinetic characterization. For the developed assays, HPLC, TLC and spectrophotometry were used as the detection methods.

To separate opportunities from the dead ends I have made an analysis of the factors and interactions that determine the feasibility of LC-bond assays, and on the basis of my experiences from the work on GE assay design, this thesis presents a number of techniques and ideas for LC-ester and -ether bond assays.

For production of natural substrates for use in LC-bond assays, a substantial amount of crude LCC preparation is required. For a carbohydrate rich-preparation, as was produced by hot-water extraction, ultrafiltration and chromatography, LCCs should be enriched by selective degradation of the polymer by hemicellulases, followed by separation of the LCCs. For a lignin-rich preparation, as was prepared by jet-milling of wood, there are no established methods, but recent developments within lignin-degrading β-etherases show great potential for this application. In addition, high-resolution analysis methods, currently HSQC NMR, would be required to qualitatively determine the presence of LCCs.

For the production of synthetic substrates for LC-bond assays, the simple quinone methide-based syntheses used for **1** may result in unstable substrates also for the corresponding ethers. However, given the low K_M that GEs display on **1** and **2** and the limited solubility of the methyl and benzyl-protected, synthetic efforts need to find ways to address both solubility and stability. One option may be, given that LCCs are amphiphilic, to assay a substrate emulsion. Another option may be adding carbohydrates instead of methyls, where possible. However, this technically challenging and carbohydrates may aggregate.

Closer at hand are natural-substrate assays using ^{31}P, FTIR or borohydride reduction for GE assays on natural substrates. Also synergy-type assays are readily applicable, if suitable substrates are at hand.

Chapter 9

Outlook

Wood is a striking demonstration of making a useful material using very simple and renewable factors: solar energy, water and carbon dioxide. *Thibaut – Mechanics of Wood and Trees*

Photosynthesizing plants make very good solar panels, but they store their energy not only in readily available carbohydrates, but also in the recalcitrant lignin. Lignin- and LC-bond degreading enzymes could help us get at that vast and underutilized resource. As this thesis shows, for the study of the LC-bond-degrading enzymes and the related fields that it depends on, three roads lead forward: i) synthesizing model substrates ii) increasing analytical resolution and iii) expanding the enzymatic toolbox.

The best stimulant to future research would probably be commercial offerings of synthetic and natural substrates. While the required equipment is, in both cases, quite standard in the right facility, there would clearly be an economy of scale, at least to the natural-substrate preparation. Also, inter-study comparisons would be feasible. Production of synthetic substrates require substantial experience, and does not necessarily scale easily. Also, as my work shows, it can be difficult to forsee what substrates will work. Although the challenges and opportunities should be clearer now.

Despite intensive study for more than a century, understanding of lignocellulosic structure and variation is far from complete, but there is significant progress in the recent observations of LC-bonds in a variety of LCC preparations by NMR and applications of MS^2 on lignin also show promise.

In the long term, better structural resolution and importantly, a much more refined structural model of lignocellulosics, is the preferred outcome of NMR spectroscopy for the research on lignocellulosic degradation. The better structural resolution we have for our assay substrates, the better we can either model them, in synthetic-substrate assays, or discriminate similar enzyme activities, in natural-substrate assays. This is true also for lignin-degradation assays.

One possibility for increased structural resolution is the use of ^{13}C-enriched substrates. These can be prepared by feeding a plant with ^{13}C-labelled compounds, as was demonstrated for Ginko shoots watered with ^{13}C-labelled lignin monomers (Xie et al., 2000), or by cultivation in a chamber where the $^{13}CO_2$ content of the atmosphere can be increased. It is also possible that tailored pulse sequences and suitable sample preparation, including ^{13}C enrichment, could allow for new ways to analyse lignin structure similar to how protein NMR developed in previous decades (Bax and Ikura, 1991). That the developments of protein NMR could be paralleled in lignin is not obvious, as lignin is unique among the biopolymers in its heterogeneity, its mixed chirality and its many branching structures. However, at some level, it is also a set of recurring features that would allow generalizations.

While it has not been extensively applied to lignin analysis, the combination of liquid chromatography with multi-dimensional mass spectrometry (MS^n) is a very potent analytical method that could very well have applications on lignocellulosic substrates not yet developed. Identification of sample components can be done even in complex mixtures if a sufficiently large collection of analytical data – retention times, m/z spectra, fragmentation spectra – are collected. In metabolomics and proteomics, putative identification of components can be automated (Davis et al., 2001; Theodoridis et al., 2008; M Brown et al., 2011; Xiao et al., 2012). There are also large libraries for identification of compounds by MS.

For the analysis of lignin by MS^n, several interesting works have been published in the last two decades (Jacobs and Dahlman, 2001; Maslen et al., 2007; D′Auria et al., 2012; Yoshioka et al., 2012). These range from structural determination of isobaric carbohydrate oligomers to studying the repeating patters of lignin. Fragmentation patterns of lignin structures have also been investigated (Morreel et al., 2010a; Morreel et al., 2010b). With data on ionization species, mass spectra and fragmentation patterns, it should be possible to construct a sufficiently detailed model of the structure of residual lignin in various samples to develop automated analytical methods based on MS.

Taken together, it appears that multi-dimensional MS is an under utilized method for the analysis of complex lignocellulosic structure. There are also opportunities for derivatization (Zhao et al., 1997) and enzymatic degradation as aides in analysis. While lignin and carbohydrate analysis may require somewhat different procedures, the LC-bond-degrading enzymes could find an analytical application in this type of analysis.

The regiospecificity of enzymes and their ability to degrade specific structures can be used to create better-characterized substrates by degrading and removing defined parts of the material. This makes enzymes important tools for LC-bond assays. However, until enzymes for selective modification of lignin are available, we must rely on specialists in analytical, wood, and organic chemistry to provide support with both substrates and analytics. Therefore, recent developments in lignin-degrading β-etherases may give biochemists an opportunity to complement the technology of glycosidic enzymes to further contribute to the common effort of elucidation of lignocellulosic structure.

For natural assay substrates, there is a two-fold challenge in applying an analytical tool and a preparatory method that will enrich the target bond at a sufficient frequency

for detection. For this, glycoside hydrolases are a valuable tool. These enzymes can also be applied directly to a complex substrate, enabling the use of a simple detection method, such as quantification of monosaccharide release, provided the target frequency is sufficient. The fact that it is largely possible to foresee the activity of a combination of enzymes with defined activities on a defined substrate can be exploited for analytical purposes. As the mass spectrometry and IC data of Paper IV shows, analysis of the recalcitrant structures can reveal information about substrate structure and enzyme specificity.

As LC-ether bonds may be an important step towards tailored lignocellulosic materials, we can speculate what kind of enzymes would be required for this activity. While no LPMOs have been suggested to act on lignin or LC-bonds, the kind of oxidative regio-specific extracellular activity that they display can serve as one possible model for how LC-ether degrading enzymes could function. The functional relevance of the abundant genes for this class of enzymes (Busk and Lange, 2015) is just starting to be explored, meaning there is ample room for finding activities of LC-bond lysis.

It should be noted that LPMOs' requirements for copper, oxygen, and an electron donor should be shared by any hypothetical LC-lytic mono-oxygenase. What a hypothetical LC-bond degrading LPMO can use as an electron donor can only be guessed. However, as inclusion of copper and one or many of the known electron-donating compounds in a screening assay is not complicated, it would make sense to do so, to enable LPMO-like activity.

Interestingly, given the enzyme's ability to use ferulic and caffeic acid and other lignin-like compounds as electron donors, the LC-bond structure could act as donor rather than acceptor, lysing the LC-bond as a consequence. However, it should be realised that while a hydrolase can combine high chemoselectivity with low regioselective (*cf.* GE activity on BnGlcA vs methyl 4-*O*-GlcA), for an oxidoreductase, this may not be the case: the wrong methylation or aryl substituent could effectively prevent activity. Thus, as for the stabilities of GlcA ester substrates, a methylation may be the difference between a travellable road and a *cul-de-sac*.

Acknowledgments

I wish to express my deepest gratitude to my main supervisor, Professor Lisbeth Olsson, for her patience, generosity and guidance during these years. Thank you for believing in me from start to finish and for supporting me when I needed it the most.

I also wish to thank my dear friends and colleagues Drs Maurizio Bettiga and Valeria Mapelli for their invaluable help and guidance in my frequent times of need. I wouldn't have made it without you.

I would like to acknowledge Dr Evangelos Topakas, for putting down time to teach me the tricks of the trade, and to Drs Sylvia Klaubauf and Silvia Hüttner for your reliance, hard work and for your stoicism in the face of delays.

I've had the pleasure of working with many inspiring and knowledgeable research partners, both in Sweden and internationally. I especially want to acknowledge Professors Paul Christakopoulos, Gunnar Henriksson and Gunnar Westman for all the help you have given me over the years, Dr Jean-Guy Berrin and Gabriela Ghizzi D.Silva for your kind milling efforts, Drs Niklas Westerberg, Mariana Charagvi, Lauren McKee, and Petri Oinonen, Associate Professor Francisco Vilaplana and my brother-in-arms in the entrenched battle of lignin chemistry, Filip Nylander.

A big "Thank you" goes out to all current and former members of the IndBio group for enthusiasm, teamwork and camaraderie in all times, and to many fine members of the Sysbio group with whom I've shared labs, laughs and boats. Thanks Christian, Calle and Johan for great outdoors adventures, and thanks George for being my supportive friend at all times.

Finally, my family, and above all Jenny, deserve the deepest thanks for supporting me and standing by me all the way to the finishing line. Thanks for your endless patience and love. Now it is time to start living for real

Bibliography

Adler, E (1977). Lignin chemistry-past, present and future. *Wood Sci Technol* 11.3, pp. 169–218.

Agger, JW, T Isaksen, A Várnai, S Vidal-Melgosa, WGT Willats, R Ludwig, SJ Horn, VGH Eijsink, and B Westereng (2014). Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci USA* 111.17, pp. 6287–6292.

Ando, D, F Nakatsubo, T Takano, and H Yano (2015). Elucidation of LCC bonding sites via γ-TTSA lignin degradation: crude milled wood lignin (MWL) from Eucalyptus globulus for enrichment of lignin xylan linkages and their HSQC-NMR characterization. *Holzforschung*.

Argyropoulos, DS (2010). Heteronuclear NMR Spectroscopy of Lignins. In: *Lignin and Lignans*. CRC Press, pp. 245–265.

Argyropoulos, DS, HI Bolker, and C Heitner (1993). ^{31}P NMR spectroscopy in wood chemistry part V. Qualitative analysis of lignin functional groups. *J Wood Chem Technol* 13.2, pp. 187–212.

Azuma, JI and T Koshijima (1988). Lignin-carbohydrate complexes from various sources A2. In: *Biomass Part B: Lignin, Pectin, and Chitin*. Academic Press, pp. 12–18.

Azuma, JI, N Takahashi, and T Koshijima (1981). Isolation and characterization of lignin-carbohydrate complexes from the milled-wood lignin fraction of *Pinus densiflora* Sieb. et Zucc. *Carbohydr Res* 93.1, pp. 91–104.

Balakshin, MY, EA Capanema, and A Berlin (2014). Isolation and Analysis of Lignin–Carbohydrate Complexes Preparations with Traditional and Advanced Methods: A Review. In: *Studies in Natural Products Chemistry*. Ed. by Atta-ur-Rahman. Elsevier, pp. 83–115.

Balakshin, MY, EA Capanema, and Hm Chang (2007). MWL fraction with a high concentration of lignin-carbohydrate linkages: Isolation and 2D NMR spectroscopic analysis. *Holzforschung* 61.1, pp. 1–7.

Balakshin, MY, EA Capanema, HS Gracz, Hm Chang, and H Jameel (2011). Quantification of lignin-carbohydrate linkages with high-resolution NMR spectroscopy. *Planta* 233.6, pp. 1097–1110.

Banoub, JH and M Delmas (2003). Structural elucidation of the wheat straw lignin polymer by atmospheric pressure chemical ionization tandem mass spectrometry and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J. Mass Spectrom.* 38.8, pp. 900–903.

Banoub, J, GH Delmas Jr., N Joly, G Mackenzie, N Cachet, B Benjelloun-Mlayah, and M Delmas (2015). A critique on the structural analysis of lignins and application of novel tandem mass spectrometric strategies to determine lignin sequencing. *J. Mass Spectrom.* 50.1, pp. 5–48.

Bauer, S (2012). Mass spectrometry for characterizing plant cell wall polysaccharides. *Front. Plant Sci.* 3.45.

Baum, A, PW Hansen, AS Meyer, and JD Mikkelsen (2013). Simultaneous measurement of two enzyme activities using infrared spectroscopy: A comparative evaluation of PARAFAC, TUCKER and N-PLS modeling. *Anal. Chim. Acta* 790, pp. 14–23.

Baumberger, S, A Abaecherli, M Fasching, G Gellerstedt, R Gosselink, B Hortling, J Li, B Saake, and E de Jong (2007). Molar mass determination of lignins by size-exclusion chromatography: towards standardisation of the method. *Holzforschung* 61.4, pp. 1–10.

Bax, A and M Ikura (1991). An efficient 3D NMR technique for correlating the proton and ^{15}N backbone amide resonances with the α-carbon of the preceding residue in uniformly ^{15}N/^{13}C enriched proteins. *J Biomol NMR* 1.1, pp. 99–104.

Beeson, WT, VV Vu, EA Span, CM Phillips, and MA Marletta (2015). Cellulose Degradation by Polysaccharide Monooxygenases. *Annu. Rev. Biochem.* 84.1, pp. 923–946.

Bennati-Granier, C, S Garajova, C Champion, S Grisel, M Haon, S Zhou, M Fanuel, D Ropartz, H Rogniaux, I Gimbert, E Record, and JG Berrin (2015). Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by Podospora anserina. *Biotechnology for biofuels* 8.1, pp. 1–14.

Bidlack, J, M Malone, and R Benson (1992). Molecular structure and component integration of secondary cell walls in plants. *Proc Okla Acad Sci* 72, pp. 51–56.

Biely, P, A Malovíková, I Uhliariková, XL Li, and DWS Wong (2015). Glucuronoyl esterases are active on the polymeric substrate methyl esterified glucuronoxylan. *FEBS Lett* 589.18, pp. 2334–2339.

Biely, P and V Puchart (2006). Recent progress in the assays of xylanolytic enzymes. *Journal of the Science of Food and Agriculture* 86.11, pp. 1636–1647.

Biely, P, RP de Vries, M Vrsanská, and J Visser (2000). Inverting character of alpha-glucuronidase A from *Aspergillus tubingensis*. *Biochim Biophys Acta* 1474.3, pp. 360–364.

Bisswanger, H (2012a). Enzyme Assays. In: *Practical Enzymology*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, pp. 93–264.

Bisswanger, H (2012b). General Aspects of Enzyme Analysis. In: *Practical Enzymology*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, pp. 5–91.

Bisswanger, H (2014). Enzyme assays. *Perspectives in Science* 1.1-6, pp. 41–55.

Björkman, A (1954). Isolation of Lignin from Finely Divided Wood with Neutral Solvents. *Nature* 174.4440, pp. 1057–1058.

Bolker, HI (1963). A Lignin Carbohydrate Bond as revealed by Infra-red Spectroscopy. *Nature* 197.4866, pp. 489–490.

Bolker, HI and N Terashima (2009). Infrared Spectroscopy of Lignins. IV. In: *Lignin Structure and Reactions*. Washington, D.C.: AMERICAN CHEMICAL SOCIETY, pp. 110–124.

Brown, JA and SC Fry (1993). The preparation and susceptibility to hydrolysis of novel O-galacturonoyl derivatives of carbohydrates. *Carbohydr Res* 240, pp. 95–106.

Brown, M, DC Wedge, R Goodacre, DB Kell, PN Baker, LC Kenny, MA Mamas, L Neyses, and WB Dunn (2011). Automated workflows for accurate mass-based putative metabolite identification in LC/MS-derived metabolomic datasets. *Bioinformatics* 27.8, pp. 1108–1112.

Brunow, G and K Lundquist (2010). Functional Groups and Bonding Patterns in Lignin (Including the Lignin-Carbohydrate Complexes). In: *Lignin and Lignans*. CRC Press, pp. 267–299.

Busk, PK and L Lange (2015). Classification of fungal and bacterial lytic polysaccharide monooxygenases. *BMC Genomics* 16.1, pp. 6287–13.

Cataldi, TRI, C Campa, and GE De Benedetto (2000). Carbohydrate analysis by high-performance anion-exchange chromatography with pulsed amperometric detection: The potential is still growing. *Fresenius J Anal Chem* 368.8, pp. 739–758.

Charavgi, MD, M Dimarogona, E Topakas, P Christakopoulos, and ED Chrysina (2013). The structure of a novel glucuronoyl esterase from *Myceliophthora thermophila* gives new insights into its role as a potential biocatalyst. *Acta Crystallogr D Biol Crystallogr* 69.Pt 1, pp. 63–73.

Comtat, J, JP Joseleau, C Bosso, and F Barnoud (1974). Characterization of structurally similar neutral and acidic tetrasaccharides obtained from the enzymic hydrolyzate of a 4-O-methyl-D-glucurono-D-xylan. *Carbohydr Res* 38, pp. 217–224.

Cragg, SM, GT Beckham, NC Bruce, TDH Bugg, DL Distel, P Dupree, AG Etxabe, BS Goodell, J Jellison, JE McGeehan, SJ McQueen-Mason, K Schnorr, PH Walton, JEM Watts, and M Zimmer (2015). Lignocellulose degradation mechanisms across the Tree of Life. *Curr Opin Chem Biol* 29, pp. 108–119.

Crepin, VF, CB Faulds, and IF Connerton (2004). Functional classification of the microbial feruloyl esterases. *Appl Microbiol Biotechnol* 63.6, pp. 647–652.

Crestini, C and DS Argyropoulos (1997). Structural Analysis of Wheat Straw Lignin by Quantitative ^{31}P and 2D NMR Spectroscopy. The Occurrence of Ester Bonds and α-O-4 Substructures. *J Agric Food Chem* 45.4, pp. 1212–1219.

Das, NN, SC Das, A Dutt, and A Roy (1981). Lignin-xylan ester linkage in jute fiber (*Corchorus capsularis*). *Carbohydr Res* 94.1, pp. 73–82.

Das, NN, SC Das, and AK Mukherjee (1984a). On the ester linkage between lignin and 4-O-methyl-d-glucurono-d-xylan in jute fiber (*Corchorus capsularis*). *Carbohydr Res* 127.2, pp. 345–348.

Das, NN, SC Das, AK Sarkar, and AK Mukherjee (1984b). Lignin-xylan ester linkage in mesta fiber (*Hibiscus cannabinus*). *Carbohydr Res* 129.C, pp. 197–207.

D′Auria, M, L Emanuele, and R Racioppi (2012). FT–ICR–MS analysis of lignin. *Natural Product Research* 26.15, pp. 1368–1374.

Davies, G and B Henrissat (1995). Structures and mechanisms of glycosyl hydrolases. *Structure* 3.9, pp. 853–859.

Davis, MT, J Beierle, ET Bures, MD McGinley, J Mort, JH Robinson, CS Spahr, W Yu, R Luethy, and SD Patterson (2001). Automated LC–LC–MS–MS platform using binary ion-exchange and gradient reversed-phase chromatography for improved proteomic analyses. *Journal of Chromatography B: Biomedical Sciences and Applications* 752.2, pp. 281–291.

Delaney, JS (2004). ESOL: Estimating Aqueous Solubility Directly from Molecular Structure. *J. Chem. Inf. Comput. Sci.* 44.3, pp. 1000–1005.

d'Errico, C, J Börjesson, H Ding, KBRM Krogh, N Spodsberg, R Madsen, and RN Monrad (2016). Improved biomass degradation using fungal glucuronoyl—esterases—hydrolysis of natural corn fiber substrate. *J. Biotechnol.* 219, pp. 117–123.

d'Errico, C, JO Jørgensen, KBRM Krogh, N Spodsberg, R Madsen, and RN Monrad (2015). Enzymatic degradation of lignin-carbohydrate complexes (LCCs): Model studies using a fungal glucuronoyl esterase from Cerrena unicolor. *Biotechnol Bioeng* 112.5, pp. 914–922.

Dimmel, D (2010). Overview. In: *Lignin and Lignans*. CRC Press, pp. 1–10.

Du, X, M Pérez-Boada, C Fernández, J Rencoret, JC del Río, J Jiménez-Barbero, J Li, A Gutiérrez, and AT Martinez (2014). Analysis of lignin–carbohydrate and lignin–lignin linkages after hydrolase treatment of xylan–lignin, glucomannan–lignin and glucan–lignin complexes from spruce wood. *Planta* 239.5, pp. 1079–1090.

Duranová, M, J Hirsch, K Kolenová, and P Biely (2009a). Fungal Glucuronoyl Esterases and Substrate Uronic Acid Recognition. *Biosci Biotech Bioch* 73.11, pp. 2483–2487.

Duranová, M, S Špániková, HAB Wösten, P Biely, and RP de Vries (2009b). Two glucuronoyl esterases of *Phanerochaete chrysosporium*. *Arch Microbiol* 191.2, pp. 133–140.

Enoki, A, F Yaku, and T Koshijima (1983). Synthesis of LCC model compounds and their chemical and enzymatic stabilities. *Holzforschung* 37.3, pp. 135–142.

Eriksson, KE (2010). Lignin Biodegradation. In: *Lignin and Lignans*. CRC Press, pp. 495–520.

Eriksson, Ö, DAI Goring, and BO Lindgren (1980). Structural studies on the chemical bonds between lignins and carbohydrates in spruce wood. *Wood Sci Technol* 14.4, pp. 267–279.

Escalante, A, A Gonçalves, A Bodin, A Stepan, C Sandström, G Toriz, and P Gatenholm (2012). Flexible oxygen barrier films from spruce xylan. *Carbohydr Polym* 87.4, pp. 2381–2387.

Esker, A, U Becker, S Jamin, S Beppu, S Renneckar, and W Glasser (2009). Self-Assembly Behavior of Some Co-and Heteropolysaccharides Related to Hemicelluloses. In: *Hemicelluloses: Science and Technology*. Washington, DC: American Chemical Society, pp. 198–219.

Ferreira, LM, TM Wood, G Williamson, C Faulds, GP Hazlewood, GW Black, and HJ Gilbert (1993). A modular esterase from Pseudomonas fluorescens subsp. cellulosa contains a non-catalytic cellulose-binding domain. *Biochem J* 294.Pt 2, pp. 349–355.

Fry, SC (1982). Phenolic components of the primary cell wall. Feruloylated disaccharides of D-galactose and L-arabinose from spinach polysaccharide. *Biochem J* 203.2, pp. 493–504.

Gall, DL, H Kim, F Lu, TJ Donohue, DR Noguera, and J Ralph (2014). Stereochemical features of glutathione-dependent enzymes in the Sphingobium sp. strain SYK-6 β-aryl etherase pathway. *J Biol Chem* 289.12, pp. 8656–8667.

Gibson, LJ (2012). The hierarchical structure and mechanics of plant materials. *J R Soc Interface* 9.76, pp. 2749–2766.

Gierer, J and HF Huber (1964). The Reactions of Lignin during Bleaching. Part I. Chlorination of Model Compounds of the beta-Arylether Type. *Acta Chem Scand* 18, pp. 1237–1243.

Gierer, J and S Wännström (1986). Formation of Ether Bonds between Lignins and Carbohydrates during Alkaline Pulping Processes. *Holzforschung* 40.6, pp. 347–352.

Glasser, WG, JS Gratzl, JJ Collins, K Forss, and JL McCarthy (1973). Lignin. XVI. Synthesis, Nuclear Magnetic Resonance, and Mass Spectroscopy of Several Monomeric and Dimeric Lignin-like Sulfonates. *Macromolecules* 6, p. 114.

Gradwell, SE, S Renneckar, AR Esker, T Heinze, P Gatenholm, C Vaca-Garcia, and W Glasser (2004). Surface modification of cellulose fibers: towards wood composites by biomimetics. *Comptes Rendus Biologies* 327.9–10, pp. 945–953.

Granata, A and DS Argyropoulos (1995). 2-Chloro-4,4,5,5-tetramethyl-1,3,2-dioxaphospholane, a Reagent for the Accurate Determination of the Uncondensed and Condensed Phenolic Moieties in Lignins. *J Agric Food Chem* 43.6, pp. 1538–1544.

Helmich, KE, JH Pereira, DL Gall, RA Heins, RP McAndrew, C Bingman, K Deng, KC Holland, DR Noguera, BA Simmons, KL Sale, J Ralph, TJ Donohue, PD Adams, and GN Phillips (2016). Structural Basis of Stereospecificity in the Bacterial Enzymatic Cleavage of β-Aryl Ether Bonds in Lignin. *J Biol Chem* 291.10, pp. 5234–5246.

Hemsworth, GR, EM Johnston, GJ Davies, and PH Walton (2015). Lytic Polysaccharide Monooxygenases in Biomass Conversion. *Trends Biotechnol.* 33.12, pp. 747–761.

Henriksson, G (2009). Lignin. In: *Wood chemistry and biotechnology*. Ed. by M Ek, G Gellerstedt, and G Henriksson. Walter de Gruyter.

Henriksson, G, E Brännvall, and H Lennholm (2009). The Trees. In: *Wood chemistry and biotechnology*. Ed. by M Ek, G Gellerstedt, and G Henriksson. Berlin, New York: Walter de Gruyter.

Henriksson, G and H Lennholm (2009). Cellulose and Carbohydrate Chemistry. In: *Wood chemistry and biotechnology*. Ed. by M Ek, G Gellerstedt, and G Henriksson. Berlin, New York: Walter de Gruyter.

Hubbe, M and L Lucia (2007). The "love-hate" relationship present in lignocellulosic materials. *BioResources* 2.4, pp. 534–535.

Ikeda, T, K Holtman, JF Kadla, Hm Chang, and H Jameel (2002). Studies on the Effect of Ball Milling on Lignin Structure Using a Modified DFRC Method. *J Agric Food Chem* 50.1, pp. 129–135.

Imamura, T, T Watanabe, M Kuwahara, and T Koshijima (1994). Ester linkages between lignin and glucuronic acid in lignin-carbohydrate complexes from Fagus crenata. *Phytochemistry* 37.4, pp. 1165–1173.

Iversen, T (1985). Lignin-carbohydrate bonds in a lignin-carbohydrate complex isolated from spruce. *Wood Sci Technol* 19.3, pp. 243–251.

Jacobs, A and O Dahlman (2001). Characterization of the Molar Masses of Hemicelluloses from Wood and Pulps Employing Size Exclusion Chromatography and Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry. *Biomacromolecules* 2.3, pp. 894–905.

Jacques, D, E Haslam, GR Bedford, and D Greatbanks (1974). Plant Proanthocyanidins .2. Proanthocyanidin-A2 and Its Derivatives. *Perkin Trans. 1* 23, pp. 2663–2671.

Jaeger, T and C Mayer (2008). N-acetylmuramic acid 6-phosphate lyases (MurNAc etherases): role in cell wall metabolism, distribution, structure, and mechanism. *Cell. Mol. Life Sci.* 65.6, pp. 928–939.

Jeffries, TW (1990). Biodegradation of lignin-carbohydrate complexes. *Biodegradation* 1.2, pp. 163–176.

Jiang, ZH, DS Argyropoulos, and A Granata (1995). Correlation analysis of ^{31}P NMR chemical shifts with substituent effects of phenols. *Magn Reson Chem* 33.5, pp. 375–382.

Joniak, D (1995). Synthesis of p-Methoxybenzyl D-Galacturonate. *Chemical Papers*.

Joseleau, JP and R Kesraoui (1986). Glycosidic Bonds between Lignin and Carbohydrates. *Holzforschung* 40.3, pp. 163–168.

Karlsson, O, T Ikeda, T Kishimoto, K Magara, Y Matsumoto, and S Hosoya (2004). Isolation of lignin–carbohydrate bonds in wood. Model experiments and preliminary application to pine wood. *Journal of Wood Science* 50.2, pp. 1–9.

Katsimpouras, C, A Bénarouche, D Navarro, M Karpusas, M Dimarogona, JG Berrin, P Christakopoulos, and E Topakas (2014). Enzymatic synthesis of model substrates recognized by glucuronoyl esterases from *Podospora anserina* and *Myceliophthora thermophila*. *Appl Microbiol Biotechnol* 98.12, pp. 5507–5516.

Kim, H, F Lu, SD Mansfield, and J Ralph (2012). Whole plant cell wall characterization using solution-state 2D NMR. *Nature Protocols* 7.9, pp. 1579–1589.

Kleinert, TN (1970). Schlagmahlung von Holz, ihr Einfluß auf Sulfit- und Kraftkochung. *Papier* 24, pp. 623–626.

Kondo, R, T Iimori, and H Imamura (1988). Enzymatic synthesis of monomeric lignin compounds with commercial β-glucosidase. *Mokuzai Gakkaishi* 34, pp. 724–731.

Kondo, R, H Yamagami, and K Sakai (1993). Xylosylation of phenolic hydroxyl groups of the monomeric lignin model compounds 4-methylguaiacol and vanillyl alcohol by Coriolus versicolor. *Appl Environ Microb* 59.2, pp. 438–441.

Koshijima, T, T Taniguchi, and R Tanaka (1972). Lignin Garbohydrate Complex. Pt. I. The Influences of Milling of Wood upon the Björkman LCC. *Holzforschung*.

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

Košíková, B and A Ebringerova (1994). Lignin-carbohydrate bonds in a residual soda spruce pulp lignin. *Wood Sci Technol* 28.4, pp. 291–296.

Košíková, B, D Joniak, and L Kosáková (1979). On the Properties of Benzyl Ether Bonds in the Lignin-Saccharidic Complex Isolated from Spruce. *Holzforschung* 33.1, pp. 11–14.

Kumar, S and A Barth (2010). Following Enzyme Activity with Infrared Spectroscopy. *Sensors* 10.4, pp. 2626–2637.

Lam, TB, K Kadoya, and K Iiyama (2001). Bonding of hydroxycinnamic acids to lignin: ferulic and p-coumaric acids are predominantly linked at the benzyl position of lignin, not the beta-position, in grass cell walls. *Phytochemistry* 57.6, pp. 987–992.

Lambeth, DO and WW Muhonen (1994). High-performance liquid chromatography-based assays of enzyme activities. *Journal of Chromatography B: Biomedical Sciences and Applications* 656.1, pp. 143–157.

Langston, JA, T Shaghasi, E Abbate, F Xu, E Vlasenko, and MD Sweeney (2011). Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl Environ Microb* 77.19, pp. 7007–7015.

Larsson, PO (2012). Enzymes – General aspects. In: *Enzyme Technology*. Lund, pp. 1–10.

Laver, ML and HHL Fang (1989). Ferulic acid esters from bark of Pseudotsuga menziesii. *J Agric Food Chem* 37.1, pp. 114–116.

Lawoko, M (2005). *Lignin Polysaccharide Networks in Softwood and Chemical Pulps: Characterisation, Structure and Reactivity*. Stockholm: KTH.

Lawoko, M, G Henriksson, and G Gellerstedt (2003). New Method for Quantitative Preparation of Lignin- Carbohydrate Complex from Unbleached Softwood Kraft Pulp: Lignin-Polysaccharide Networks I. *Holzforschung* 57.1, pp. 69–74.

Lawoko, M, G Henriksson, and G Gellerstedt (2006). Characterisation of lignin-carbohydrate complexes (LCCs) of spruce wood (*Picea abies* L.) isolated with two methods. *Holzforschung* 60.2, pp. 151–161.

Leisola, M, O Pastinen, and DD Axe (2012). Lignin–Designed Randomness. *BIO-Complexity* 2012.3, pp. 1–11.

Levasseur, A, P Gouret, L Lesage-Meessen, M Asther, E Record, and P Pontarotti (2006). Tracking the connection between evolutionary and functional shifts using the fungal lipase/feruloyl esterase A family. *BMC Evol Biol* 6.1, p. 92.

Lewis, JK, J Wei, and G Siuzdak (2006). *Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry in Peptide and Protein Analysis*. Applications, Theory and Instrumentation. Chichester, UK: John Wiley & Sons, Ltd.

Li, J, G Henriksson, and G Gellerstedt (2007). Lignin depolymerization/repolymerization and its critical role for delignification of aspen wood by steam explosion. *Bioresour Technol* 98.16, pp. 3061–3068.

Li, K and RF Helm (1995). Synthesis and Rearrangement Reactions of Ester-Linked Lignin-Carbohydrate Model Compounds. *J Agric Food Chem* 43.8, pp. 2098–2103.

Li, XL, S Špániková, RP de Vries, and P Biely (2007). Identification of genes encoding microbial glucuronoyl esterases. 581.21, pp. 4029–4035.

Lo Leggio, L, TJ Simmons, JCN Poulsen, KEH Frandsen, GR Hemsworth, MA Stringer, P von Freiesleben, M Tovborg, KS Johansen, L De Maria, PV Harris, CL Soong, P Dupree, T Tryfona, N Lenfant, B Henrissat, GJ Davies, and

PH Walton (2015). Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nat Commun* 6, p. 5961.

Lundquist, K, R Simonson, and K Tingsvik (1983). Lignin carbohydrate linkages in milled wood lignin preparations from spruce wood. *Sven Papperstidn* 86.6, R44–R47.

Mabey, W and T Mill (1978). Critical review of hydrolysis of organic compounds in water under environmental conditions. *Journal of Physical and Chemical Reference Data* 7.2, pp. 383–415.

Marangoni, AG (2002a). *Enzyme Kinetics*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Marangoni, AG (2002b). Tools and Techniques of Kinetic Analysis. In: *Enzyme Kinetics*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 1–40.

Marangoni, AG (2003). *Interfacial Enzymes*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Masai, E, A Ichimura, Y Sato, K Miyauchi, Y Katayama, and M Fukuda (2003). Roles of the enantioselective glutathione *S*-transferases in cleavage of beta-aryl ether. *J. Bacteriol.* 185.6, pp. 1768–1775.

Masai, E, Y Katayama, S Kubota, S Kawai, M Yamasaki, and N Morohoshi (1993). A bacterial enzyme degrading the model lignin compound β-etherase is a member of the glutathione-*S*-transferase superfamily. 323.1-2, pp. 135–140.

Masai, E, Y Katayama, and M Fukuda (2007). Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds. *Biosci Biotech Bioch* 71.1, pp. 1–15.

Maslen, SL, F Goubet, A Adam, P Dupree, and E Stephens (2007). Structure elucidation of arabinoxylan isomers by normal phase HPLC–MALDI-TOF/TOF-MS/MS. *Carbohydr Res* 342.5, pp. 724–735.

Melone, F, R Saladino, H Lange, and C Crestini (2013). Tannin structural elucidation and quantitative ^{31}P NMR analysis. 1. Model compounds. *J Agric Food Chem* 61.39, pp. 9307–9315.

Merayo, N, D Hermosilla, C Negro, and Á Blanco (2013). On-line FTIR as a novel tool to monitor Fenton process behavior. *Chemical Engineering Journal* 232, pp. 519–526.

Min, Dy, Q Li, V Chiang, H Jameel, Hm Chang, and L Lucia (2014). The influence of lignin–carbohydrate complexes on the cellulase-mediated saccharification I: Transgenic black cottonwood (western balsam poplar, California poplar) *P. trichocarpa* including the xylan down-regulated and the lignin down-regulated lines. *Fuel* 119.C, pp. 207–213.

Moreira, LRS and EXF Filho (2008). An overview of mannan structure and mannan-degrading enzyme systems. *Appl Microbiol Biotechnol* 79.2, pp. 165–178.

Morreel, K, O Dima, H Kim, F Lu, C Niculaes, R Vanholme, R Dauwe, G Goeminne, D Inzé, E Messens, J Ralph, and W Boerjan (2010a). Mass spectrometry-based sequencing of lignin oligomers. *Plant Physiol* 153.4, pp. 1464–1478.

Morreel, K, H Kim, F Lu, O Dima, T Akiyama, R Vanholme, C Niculaes, G Goeminne, D Inzé, E Messens, J Ralph, and W Boerjan (2010b). Mass spectrometry-based fragmentation as an identification tool in lignomics. *Anal. Chem.* 82.19, pp. 8095–8105.

Nylander, F, H Sunner, L Olsson, P Christakopoulos, and G Westman (2016). Synthesis and enzymatic hydrolysis of a diaryl benzyl ester model of a lignin-carbohydrate complex (LCC). *Holzforschung* 70.5, pp. 385–391.

Obst, J (1982). Frequency and alkali resistance of wood lignin-carbohydrate bonds in wood. *Tappi/April* 65.4, pp. 109–112.

Ohta, Y, S Nishi, R Hasegawa, and Y Hatada (2015). Combination of six enzymes of a marine *Novosphingobium* converts the stereoisomers of β-O-4 lignin model dimers into the respective monomers. *Scientific Reports* 5, p. 15105.

Oinonen, P, D Areskogh, and G Henriksson (2013). Enzyme catalyzed cross-linking of spruce galactoglucomannan improves its applicability in barrier films. *Carbohydr Polym* 95.2, pp. 690–696.

Oinonen, P, L Zhang, M Lawoko, and G Henriksson (2015). On the formation of lignin polysaccharide networks in Norway spruce. *Phytochemistry* 111, pp. 177–184.

Olesen, NE, R Holm, and P Westh (2014). Determination of the aggregation number for micelles by isothermal titration calorimetry. *Thermochimica Acta* 588, pp. 28–37.

Otero, JM, G Panagiotou, and L Olsson (2007). Fueling industrial biotechnology growth with bioethanol. *Adv. Biochem. Eng. Biotechnol.* 108, pp. 1–40.

Parkin, KL (2002). Putting Kinetic Principles into Practice. In: *Enzyme Kinetics*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 174–192.

Pérez-Boada, M, FJ Ruiz-Duenas, R Pogni, R Basosi, T Choinowski, MJ Martínez, K Piontek, and AT Martinez (2005). Versatile Peroxidase Oxidation of High Redox Potential Aromatic Compounds: Site-directed Mutagenesis, Spectroscopic and Crystallographic Investigation of Three Long-range Electron Transfer Pathways. *J Mol Biol* 354.2, pp. 385–402.

Pew, JC and P Weyna (1962). Fine grinding, enzyme digestion, and the lignin-cellulose bond in wood. *Tappi* 45.3, pp. 247–256.

Picart, P, C Müller, J Mottweiler, L Wiermans, C Bolm, P Domínguez de María, and A Schallmey (2014). From gene towards selective biomass valorization: bacterial β-etherases with catalytic activity on lignin-like polymers. *ChemSusChem* 7.11, pp. 3164–3171.

Pokkuluri, PR, NEC Duke, SJ Wood, MA Cotta, XL Li, P Biely, and M Schiffer (2011). Structure of the catalytic domain of glucuronoyl esterase Cip2 from Hypocrea jecorina. *Proteins* 79.8, pp. 2588–2592.

Pu, Y, S Cao, and AJ Ragauskas (2011). Application of quantitative ^{31}P NMR in biomass lignin and biofuel precursors characterization. *Energy & Environmental Science* 4, pp. 3154–3166.

Quinlan, RJ, MD Sweeney, L Lo Leggio, H Otten, JCN Poulsen, KS Johansen, KBRM Krogh, CI Jørgensen, M Tovborg, A Anthonsen, T Tryfona, CP Walter, P Dupree, F Xu, GJ Davies, and PH Walton (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the National Academy of Sciences* 108.37, pp. 15079–15084.

Ralph, J, RD Hatfield, JH Grabber, HJ Jung, S Quideau, and RF Helm (1996). "Cell wall cross-linking in grasses by ferulates and diferulates". In: pp. 1–8.

Ralph, J and L Landucci (2010). NMR of Lignins. In: *Lignin and Lignans*. CRC Press, pp. 137–243.

Ralph, J and RA Young (1983). Stereochemical aspects of addition-reactions involving lignin model quinone methides. *J Wood Chem Technol* 3.2, pp. 161–181.

Reiter, J, H Strittmatter, LO Wiemann, D Schieder, and V Sieber (2013). Enzymatic cleavage of lignin β-O-4 aryl ether bonds via net internal hydrogen transfer. *Green Chemistry* 15.5, pp. 1373–1381.

Richards, S, M Ropic, D Blackmond, and A Walmsley (2004). Quantitative determination of the catalysed asymmetric transfer hydrogenation of 1-methyl-6,7-dimethoxy-3,4-dihydroisoquinoline using in situ FTIR and multivariate curve resolution. *Anal. Chim. Acta* 519.1, pp. 1–9.

Rosini, E, C Allegretti, R Melis, L Cerioli, G Conti, L Pollegioni, and P D'Arrigo (2016). Cascade enzymatic cleavage of the β-O-4 linkage in a lignin model compound. *Catal. Sci. Technol.* 6.7, pp. 2195–2205.

Rowe, JW, CL Bower, and ER Wagner (1969). Extractives of jack pine bark: Occurrence of *cis*- and *trans*-pinosylvin dimethyl ether and ferulic acid esters. *Phytochemistry* 8.1, pp. 235–241.

Schiff, DE, JG Verkade, RM Metzler, TG Squires, and CG Venier (1986). Determination of Alcohols, Phenols, and Carboxylic Acids Using Phosphorus-31 NMR Spectroscopy. *Appl Spectrosc* 40.3, pp. 348–351.

Schrodinger, LLC (2010). The PyMOL Molecular Graphics System, Version 1.3.

Singh, R, S Singh, K Trimukhe, K Pandare, K Bastawade, D Gokhale, and A Varma (2005). Lignin-carbohydrate complexes from sugarcane bagasse: Preparation, purification, and characterization. *Carbohydr Polym* 62.1, pp. 57–66.

Sipilä, J and G Brunow (1991a). On the Mechanism of Formation of Non-Cyclic Benzyl Ethers During Lignin Biosynthesis Part 3. The Reactivity of a β—O—4-Type Quinone Methide with Methyl-α-D-glucopyranoside in Competition with Vanillyl Alcohol. The Formation and the Stability of Benzyl Ethers between Lignin and Carbohydrates. *Holzforschung* 45.s1, pp. 3–8.

Sipilä, J and G Brunow (1991b). On the Mechanism of Formation of Non-Cyclic Benzyl Ethers During Lignin Biosynthesis. Part 4. The Reactions of a β—O—4-type Quinone Methide with Carboxylic Acids in the Presence of Phenols. The Formation and Stability of Benzyl Esters between Lignin and Carbohydrates. *Holzforschung* 45.s1, pp. 9–14.

Solomon, EI, UM Sundaram, and TE Machonkin (1996). Multicopper Oxidases and Oxygenases. *Chem. Rev.* 96.7, pp. 2563–2606.

Song, T, A Pranovich, I Sumerskiy, and B Holmbom (2008). Extraction of galactoglucomannan from spruce wood with pressurised hot water. *Holzforschung* 62.6, pp. 659–666.

Spangenberg, PDB, CF Poole, and C Weins (2010). Theoretical Basis of Thin Layer Chromatography (TLC). In: *Quantitative Thin-Layer Chromatography*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 13–52.

Špániková, S and P Biely (2006). Glucuronoyl esterase - Novel carbohydrate esterase produced by *Schizophyllum commune*. *FEBS Lett* 580.19, pp. 4597–4601.

Špániková, S, M Poláková, D Joniak, J Hirsch, and P Biely (2007). Synthetic esters recognized by glucuronoyl esterase from *Schizophyllum commune*. *Arch Microbiol* 188.2, pp. 185–189.

Takahashi, N, JI Azuma, and T Koshijima (1982). Fractionation of lignin-carbohydrate complexes by hydrophobic-interaction chromatography. *Carbohydr Res* 107.2, pp. 161–168.

Takahashi, N and T Koshijima (1988). Ester linkages between lignin and glucuronoxylan in a lignin-carbohydrate complex from beech (*Fagus crenata*) wood. *Wood Sci Technol* 22.3, pp. 231–241.

Taneda, H, J Nakano, S Hosoya, and Hm Chang (1987). Stability of α-Ether type Model Compounds During Chemical Pulping Processes. *J Wood Chem Technol* 7.4, pp. 485–497.

Teleman, A (2009). Hemicelluloses and pectins. In: *Wood chemistry and biotechnology*. Ed. by M Ek, G Gellerstedt, and G Henriksson. Berlin, New York: Walter de Gruyter.

Theodoridis, G, HG Gika, and ID Wilson (2008). LC-MS-based methodology for global metabolite profiling in metabonomics/metabolomics. *TrAC Trends in Analytical Chemistry* 27.3, pp. 251–260.

Thibaut, B, J Gril, and M Fournier (2001). Mechanics of wood and trees: some new highlights for an old story. *Comptes Rendus de l'Académie des Sciences - Series IIB - Mechanics* 329.9, pp. 701–716.

Thompson, JE and SC Fry (2000). Evidence for covalent linkage between xyloglucan and acidic pectins in suspension-cultured rose cells. *Planta* 211.2, pp. 275–286.

Timell, TE (1967). Recent progress in the chemistry of wood hemicelluloses. *Wood Sci Technol* 1.1, pp. 45–70.

Toikka, M, J Sipilä, A Teleman, and G Brunow (1998). Lignin-carbohydrate model compounds. Formation of lignin-methyl arabinoside and lignin-methyl galactoside benzyl ethers via quinone methide intermediates. *Perkin Trans. 1* 22, pp. 3813–3818.

Topakas, E, M Moukouli, M Dimarogona, C Vafiadi, and P Christakopoulos (2010). Functional expression of a thermophilic glucuronyl esterase from *Sporotrichum thermophile*: identification of the nucleophilic serine. *Appl Microbiol Biotechnol* 87.5, pp. 1765–1772.

Topakas, E, C Vafiadi, and P Christakopoulos (2007). Microbial production, characterization and applications of feruloyl esterases. *Process Biochemistry* 42.4, pp. 497–509.

Tsai, AYL, T Canam, A Gorzsás, EJ Mellerowicz, MM Campbell, and ER Master (2012). Constitutive expression of a fungal glucuronoyl esterase in Arabidopsis reveals altered cell wall composition and structure. *Plant Biotechnol J* 10.9, pp. 1077–1087.

Uraki, Y, Y Usukura, T Kishimoto, and M Ubukata (2006). Amphiphilicity of a lignin-carbohydrate complex. *Holzforschung* 60.6, pp. 1–6.

Vaaje-Kolstad, G, B Westereng, SJ Horn, Z Liu, H Zhai, M Sørlie, and VGH Eijsink (2010). An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330.6001, pp. 219–222.

Vafiadi, C, E Topakas, P Biely, and P Christakopoulos (2009). Purification, characterization and mass spectrometric sequencing of a thermophilic glucuronoyl esterase from *Sporotrichum thermophile*. *Fems Microbiol Lett* 296.2, pp. 178–184.

Virgili, F, G Pagana, L Bourne, G Rimbach, F Natella, C Rice-Evans, and L Packer (2000). Ferulic acid excretion as a marker of consumption of a French maritime pine (*Pinus maritima*) bark extract. *Free Radical Biology and Medicine* 28.8, pp. 1249–1256.

Vu, VV, WT Beeson, EA Span, ER Farquhar, and MA Marletta (2014). A family of starch-active polysaccharide monooxygenases. *Proc Natl Acad Sci USA* 111.38, pp. 13822–13827.

Wang, PY, HI Bolker, and CB Purves (1967). Uronic acid ester groups in some softwoods and hardwoods. *Tappi* 3, pp. 123–124.

Watanabe, T (1989). Structural Studies on the Covalent Bonds between Lignin and Carbohydrate in Lignin-Carbohydrate Complexes by Selective Oxidation of the Lignin with 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone. *Wood research : bulletin of the Wood Research Institute Kyoto University* 76, pp. 59–123.

Watanabe, T and T Koshijima (1988). Evidence for an Ester Linkage between Lignin and Glucuronic Acid in Lignin-Carbohydrate Complexes by DDQ-Oxidation. *Agr Biol Chem Tokyo* 52.11, pp. 2953–2955.

Watanabe, T, J Ohnishi, Y Yamasaki, S Kaizu, and T Koshijima (1989). Binding-site analysis of the ether linkages between lignin and hemicelluloses in lignin-carbohydrate complexes by DDQ-oxidation. *Agr Biol Chem Tokyo* 53.8, pp. 2233–2252.

Westbye, P, T Köhnke, W Glasser, and P Gatenholm (2007). The influence of lignin on the self-assembly behaviour of xylan rich fractions from birch (Betula pendula). *Cellulose* 14.6, pp. 603–613.

Westereng, B, D Cannella, J Wittrup Agger, H Jørgensen, M Larsen Andersen, VGH Eijsink, and C Felby (2015). Enzymatic cellulose oxidation is linked to lignin by long-range electron transfer. *Scientific Reports* 5, p. 18561.

Widsten, P and A Kandelbauer (2008). Laccase applications in the forest products industry: A review. *Enzyme Microb Technol* 42.4, pp. 293–307.

Wildman, SA and GM Crippen (1999). Prediction of Physicochemical Parameters by Atomic Contributions. *J. Chem. Inf. Comput. Sci.* 39.5, pp. 868–873.

Williamson, G, PA Kroon, and CB Faulds (1998). Hairy plant polysaccharides: a close shave with microbial esterases. *Microbiology* 144.8, pp. 2011–2023.

Wong, DWS (2006). Feruloyl esterase. *Appl. Biochem. Biotechnol.* 133.2, pp. 87–112.

Wong, DWS (2009). Structure and Action Mechanism of Ligninolytic Enzymes. *Appl. Biochem. Biotechnol.* 157.2, pp. 174–209.

Wong, DWS, VJ Chan, AA McCormack, J Hirsch, and P Biely (2012). Functional Cloning and Expression of the Schizophyllum commune Glucuronoyl Esterase Gene and Characterization of the Recombinant Enzyme. *Biotechnol Res Int* 2012, p. 951267.

Wood, SJ, XL Li, MA Cotta, P Biely, NEC Duke, M Schiffer, and PR Pokkuluri (2008). Crystallization and preliminary X-ray diffraction analysis of the glucuronoyl esterase catalytic domain from Hypocrea jecorina. *Acta Crystallogr F* 64, pp. 255–257.

Wroblewski, AE, C Lensink, R Markuszewski, and JG Verkade (1988). Phosphorus-31 NMR spectroscopic analysis of coal pyrolysis condensates and extracts for heteroatom functionalities possessing labile hydrogen. *Energy Fuels* 2.6, pp. 765–774.

Xiao, JF, B Zhou, and HW Ressom (2012). Metabolite identification and quantitation in LC-MS/MS-based metabolomics. *TrAC Trends in Analytical Chemistry* 32, pp. 1–14.

Xie, Y, S Yasuda, H Wu, and H Liu (2000). Analysis of the structure of lignin-carbohydrate complexes by the specific ^{13}C tracer method. *Journal of Wood Science* 46.2, pp. 130–136.

Yaku, F, S Tsuji, and T Koshijima (1979). Lignin Carbohydrate Complex. Pt. III. Formation of Micelles in the Aqueous Solution of Acidic Lignin Carbohydrate Complex. *Holzforschung* 33.2, pp. 54–59.

Yaku, F, Y Yamada, and T Koshijima (1976). Lignin Carbohydrate Complex Pt. II. Enzymic Degradation of Acidic Polysaccharide in Björkman LCC. *Holzforschung* 30, p. 148.

Yapo, BM (2011). Pectic substances: From simple pectic polysaccharides to complex pectins - A new hypothetical model. *Carbohydr Polym* 86.2, pp. 373–385.

Yoshioka, K, D Ando, and T Watanabe (2012). A comparative study of matrix- and nano-assisted laser desorption/ionisation time-of-flight mass spectrometry of isolated and synthetic lignin. *Phytochem Anal* 23.3, pp. 248–253.

Yuan, TQ, SN Sun, F Xu, and RC Sun (2011). Characterization of lignin structures and lignin-carbohydrate complex (LCC) linkages by quantitative ^{13}C and 2D HSQC NMR spectroscopy. *J Agric Food Chem* 59.19, pp. 10604–10614.

Zaia, J (2008). Mass Spectrometry and the Emerging Field of Glycomics. *Chemistry & Biology* 15.9, pp. 881–892.

Zhao, Y, SB Kent, and BT Chait (1997). Rapid, sensitive structure analysis of oligosaccharides. *Proc Natl Acad Sci USA* 94.5, pp. 1629–1633.

Ziebell, AL (2008). *Modelling lignin depolymerisation using size exclusion chromatography*. PhD thesis. Faculty of Life and Social Sciences, Swinburne University of Technology.

Zikeli, F, T Ters, K Fackler, E Srebotnik, and J Li (2014). Successive and quantitative fractionation and extensive structural characterization of lignin from wheat straw. *Industrial Crops and Products* 61, pp. 249–257.