Green
Chemistry

A Critical Review on the Analysis of Lignin Carbohydrate Bonds in Plants

Journal:	*Green Chemistry*
Manuscript ID	GC-CRV-11-2018-003606.R2
Article Type:	Critical Review
Date Submitted by the Author:	14-Feb-2019

Complete List of Authors:	Giummarella, Nicola; Wallenberg Wood Science Center Pu, Yunqiao; Oak Ridge National Laboratory, Joint Institute of Biological Science, Biosciences Division Ragauskas, Arthur; University of Tennessee, Lawoko, Martin; Wallenberg Wood Science Center (WWSC), Fiber and Polymer Technology

SCHOLARONE™
Manuscripts

Please do not adjust margins

Green Chemistry

Critical Review

A Critical Review on the Analysis of Lignin Carbohydrate Bonds

Nicola Giummarella,[a] Yunqiao Pu,[b,c] Arthur J. Ragauskas*[b,c,d,e] and Martin Lawoko*[a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Replacing fossil-based resources with renewable alternatives is generally acknowledged as a critical component to address several of today's environmental concerns. In this context, lignocellulosic biomass is an attractive, sustainable resource. However, the constitutional biopolymers of interest are locked in the structural complexity of the plant cell walls, which defines their properties and contributes to fractionation recalcitrance. One of the key suspects restricting fractionation of the biopolymers in high yield is the presence of lignin-carbohydrate bonds forming a matrix referred to as Lignin-Carbohydrate Complexes (LCC). Nevertheless, covalent bonds between lignin and carbohydrates, remain one of the most controversial topics in lignocellulose chemistry. This challenge can be attributed to the slow progress made in their research, which also forms the basis for this review. Herein, we will critically discuss the literature with a particular focus on the latest characterization and analytical techniques. Discussions on existing techniques and, importantly the drawbacks with them should be compelling to researchers in the area, especially at this time when crucial issues surrounding the realization of biorefineries need to be addressed.

Introduction

Environmental challenges, rural development, and energy security have steered our society towards much-needed utilization of renewable and sustainable carbon-neutral resources for the production of bio-based materials, chemicals and bioenergy.[1] Biomass, and especially lignocellulosics are an attractive candidate as they constitute the top two most abundant renewable polymers namely cellulose and lignin, with the former being number one. The separation of these polymers includes not only old traditional methods, e.g., chemical pulping methods, but also new separation concepts referred to as biorefinery concepts, which target to efficiently fractionate and yield reactive cellulose, lignin, and hemicellulose or degradation products. Promising biorefinery concepts such as co-solvent enhanced lignocellulosic extractions (CELF) processes[2] have produced valuable streams from cellulose and hemicellulose fractions, concomitantly with pure lignin fractions which could be of interest for material applications. Among all lignocellulosic resources, woody biomass plays a crucial role in a biorefinery context since it does not compete with food production and its availability could be tuned and increased by appropriate forest economy and climate change targeted policies.

The main wood biopolymers, namely lignin, hemicellulose and cellulose, are locked in the hierarchical complexity of the wood plant cell walls providing toughness, strength and resilience properties. However, the complex wood cell wall ultrastructure contributes strongly to fractionation recalcitrance. More specifically, the presence of lignin-carbohydrate bonds, forming so-called Lignin-Carbohydrate complexes (LCC) are suggested to play a crucial role in recalcitrance during biomass processing and fractionation. Such a hypothesis is not far conceived as attempts to extract carbohydrates selectively could be resisted by bonding to insoluble lignin moieties and vice-versa.

Already back in 1866, Erdmann[3] hypothesized the presence of covalent bonds between lignin and carbohydrates adopting the concept of “glycolignose” wood to explain the inability in separating the two components. Later, the low yield of pure lignins obtained by subjecting wood to ethanol extraction, referred to as "Brauns native lignin," were explained by the fact that the rest was bound to carbohydrates.[4] A comprehensive account on various LCC fractionation efforts dating before the 21st century can be found in the elegant book edited by Koshijima and Watanabe.[5] Although the complete structure of native lignin is still unknown, it is widely accepted that the nearest one that can be used to get to its structure is the lignin obtained by extraction from milled wood at ambient conditions with dioxane/water co-mixtures, commonly termed milled wood lignin (MWL); a protocol developed by Björkman.[6] LCC obtained herein from the

*a. Wallenberg Wood Science Centre, Division of Wood Chemistry and Pulp Technology, Department of Fibre and Polymer Technology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Teknikringen 56-58, SE–100 44, Stockholm, Sweden. *E-mail: lawoko@kth.se*
b. Center for BioEnergy Innovation (CBI), Oak Ridge National Laboratory (ORNL), Oak Ridge, TN 37831, USA
c. Joint Institute for Biological Sciences, Biosciences Division, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN 37831, USA
d. Department of Chemical and Biomolecular Engineering, The University of Tennessee, Knoxville, TN 37996-2200, USA. E-mail: aragausk@utk.edu*
e. Department of Forestry, Wildlife and Fisheries, Center for Renewable Carbon, The University of Tennessee Institute of Agriculture, Knoxville, TN 37996-2200, USA
** Corresponding authors*

Please do not adjust margins

residuals after MWL isolation were referred to as Bjorkman LCC. Following Björkman´s work, several studies emerged and included detailed studies on the residues left after extractions and different characterization techniques. These are discussed in this review.

Studies have suggested that virtually all residual lignin in sulphite pulps[7] as well as lignin in both unbleached and oxygen delignified softwood kraft pulps[8] is covalently bound to polysaccharides. This is consistent with the low selectivity of delignification during the final phase of Kraft pulping.[9] Similarly, the resistance to delignification during pulping was attributed to the presence of alkaline stable lignin-carbohydrate bonds when residual lignin, prepared by enzymatic treatment, was studied by Size Exclusion Chromatography (SEC).[10,11,12]

Nevertheless, despite plentiful indications of LCC presence, even after 150 years, unequivocal proofs of their existence in wood still lacks due to debatable isolation and analytical protocols. At the cornerstone of these difficulties, there are confusing claims in regards to LCC characterization. Even the most prominent analytical methodology for LCC, namely NMR, has not unequivocally assigned all lignin carbohydrate (LC) linkages. Over the last decade, the Lignin-Carbohydrate Complexes (LCC) remains an issue of conflicting claims, especially as it applies to their contribution to biomass recalcitrance. Five types of lignin-carbohydrate bonds are suggested in the literature, namely, phenyl glycosides (PG),[13,14,15,16,17] benzyl ethers (BE),[18,19,20,21,22] γ-esters (GE),[17,23,24,25,26] ferulate/coumarate esters (FE/CE)[27,28,29,30,31,32,33,34] and hemiacetal/acetal linkages[35,36] (Figure 1).

Dr. Nicola Giummarella comes from Italy. After getting a bachelor's degree in Material Science from the University of Bari, he joined the KTH Royal Institute of Technology in Sweden and earned an M.Sc. in Macromolecular Materials and a Ph.D. in Fiber and Polymer Technology. His expertise includes fundamental aspects of wood chemistry as well as advanced chromatographic and spectroscopic characterization methods. Presently, he is working as a Postdoctoral Researcher whose research is primarily focused on the structural elucidation and upgrading of technical lignin for valorisation.

Dr. Martin Lawoko is an Associate Professor of Wood Chemistry at KTH, Royal Institute of Technology, Sweden. After receiving a PhD in Fiber and Polymer Technology at KTH, he completed a postdoc investigating lignocellulose biorefining at the University of Maine, USA before returning to KTH. Currently, Dr. Lawoko leads lignin research encompassing fundamental aspects of polymerization, structure, biorefining and modification at the Wallenberg Wood Science Center, KTH. His research aims to utilize green chemistry to fractionate and modify biomass-derived polymers.

Dr. Yunqiao Pu is a Senior R&D Staff in the Biosciences Division at Oak Ridge National Laboratory (ORNL). Before joining ORNL in 2014, he was a Senior Research Scientist in the Renewable Bioproducts Institute at Georgia Institute of Technology. His research interests include plant cell wall structure/biosynthesis, wood/lignin chemistry, cellulose/ hemicellulose chemistry and structures, forest biorefinery, biofuels, renewable biopolymers for biomaterials and biocomposites, nanocellulosics, and fiber engineering. He has expertise in solution- and solid-state NMR, HPLC, GPC, GC, FTIR/NIR analysis of biopolymers and natural products.

Professor Arthur Ragauskas held the first Fulbright Chair in Alternative Energy and is a Fellow of American Association for the Advancement of Science, the International Academy of Wood Science and TAPPI. In 2014, he assumed a Governor's Chair for Biorefining based in University of Tennessee's Department of Chemical and Biomolecular Engineering, with a complementary appointment in the UT Institute of Agriculture's Department of Forestry, Wildlife, and Fisheries and serves in the US Energy and Environmental Sciences Directorate, Biosciences Division, at ORNL. His research program is directed at understanding and exploiting innovative sustainable bioresources.

This journal is © The Royal Society of Chemistry 20xx

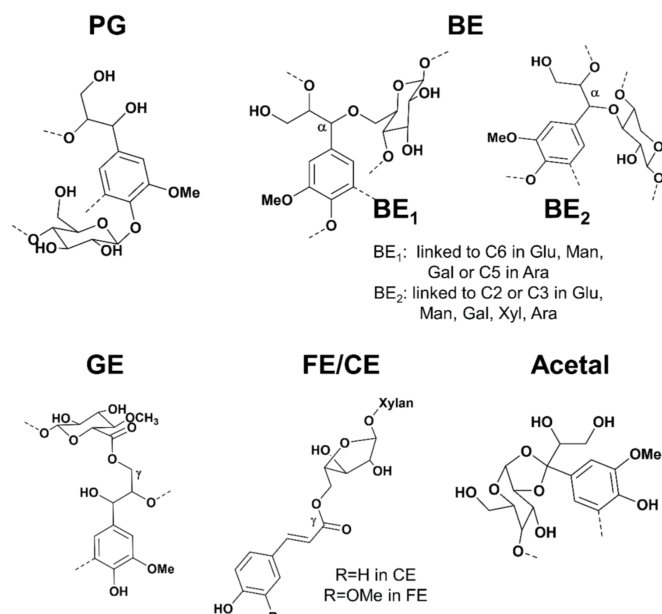


Figure 1: Suggested structures of Lignin-Carbohydrate bonds (LC) in wood and grass. In the figure, PG=phenyl glycosides, BE=benzyl ethers; GE=γ-esters; FE=ferulate esters; CE=coumarate esters.

With the goal of inspiring progressive LCC research, we critically discuss the literature with special focus on isolation, characterization and analytical techniques. The content of the review is outlined below:

- A recent development of LCC isolation techniques from the 21st century, evaluating the advantages and disadvantages. Although some work from the previous century is included in order to contextualize our discussions, details of these former studies can be obtained in the excellent collection by Koshijima and Watanabe;[5]
- A summary of the indirect methods used for lignin-carbohydrate bond analysis and the pros and cons of the methods are reviewed;
- Direct analysis of the lignin-carbohydrate bond by NMR techniques and a critical discussion on them;
- A future outlook on progressive LCC studies.

Overview of LCCs isolation protocol

The hierarchical and complex structure of plant cell walls has led to the unavoidable adoption of harsh conditions to isolate LCC or lignin for analysis. Accordingly, a valid scientific question is whether Lignin-Carbohydrates (LC) bonds are native or artefacts created during extraction. Unfortunately, a definitive answer to this question may have to wait until solid state characterizations with high sensitivity and accuracy at the molecular scale is developed. Till then, for solution-state analysis, it is ideal that reliable isolation protocol should adopt conditions as mild as possible in order to minimize structural changes and, at the same time, should provide a quantitative yield to represent the bulk of starting substrate. Lastly, the possibility to utilize a universal protocol (applicable to all lignocellulosic biomasses) is another significant advantage for analytical and technical purposes.

Isolation of LCC from milled wood

The mechanochemistry of ball milling of wood has been studied to some extent. Specific to lignin, homolytic cleavage with consequent formation of mechano-radicals leads to an increase of carbonyl and phenolic OH groups in lignin. Moreover, a decrease of molecular mass of lignin takes place together with some cleavage of aryl ether linkages.[6,37] On the other hand, LCC fragments can be released and extracted from the cell wall matrix. The question then is if they could form through radical quenching reactions between polysaccharides and lignin. Indeed, the ball milling remains an event that requires more in-depth studies on structural changes. Nevertheless, vibratory ball milling was adopted to obtain the so-called Björkman LCC, defined as the "classical" LCC preparation from a finely milled wood (Figure 2). In this protocol, crude milled wood lignin is obtained as the dioxane: water (96:4 in v:v) extract from finely grounded wood. The LCC enriched fraction is obtained by extraction of the residual solid material with dimethylsulfoxide (DMSO) or dimethylformamide (DMF).[38] Higher yield was obtained with DMSO,[39] however, in both cases, the LCC product consisted mostly of carbohydrates with lignin accounting for 20-30% of the fraction. The moderate yields usually achieved in this procedure can be increased at longer milling time.[40] Notably, the inability to refine and purify lignin from this fraction was seen as indirect evidence of covalent bonds between lignin and carbohydrates. However, in more recent studies, LCC enriched fractions were obtained by one of the classical purification process for crude MWL (Figure 2).[41,42] It has been shown that precipitation in ethanol of the crude MWL would be beneficial in recovering the bulk of hemicelluloses and therefore increase the yield of recovered lignin bound to them. Later on, with the purpose of avoiding the use of high boiling point solvents (DMSO, DMF), Watanabe and co-workers[43] adopted cold- (20 °C) and hot- water (80 °C) extractions, sequentially, to obtain an LCC fraction (LCC-WE) from the extracted finely divided wood residue with 80% aqueous dioxane (Figure 2). Approximately, the same chemical characteristics and lignin content of Björkman LCC were achieved. Recently, different analogues of modified Bjorkman methods to enrich/obtain LCC fractions more suitable for characterization with modern techniques have also been developed. For instance, in one approach, crude milled wood lignin is first dissolved in acetic acid and then precipitated by addition of water to obtain enriched LCC fraction referred to as LCC-AcOH[44,45,46,47] (Figure 2). This LCC-enriched fraction (sum of LC linkages~0.15–0.20/aromatic unit)44,45,[47] contained up to 15% of original wood lignin and proposed spectroscopic evidence of LC bonds were presented. However, it is worthy to note that yields, structure, and lignin-carbohydrate bond quantities change with milling condition (e.g., time and supplied energy).[47] In another case, acetylated cell walls of ball milled wood substrates (Ac-CW, Figure 2) were used to obtain insights on chemical linkages in wood polymers due to

complete solubilization in N-methylimidazole/dimethyl sulfoxide (NMI/DMSO) which permitted direct solution-state NMR analysis.[48] Although representative structural information of the whole ball milled cell wall can potentially be obtained by this method, the predominance of signals from carbohydrate- and lignin- linkages compromises the spectroscopic analysis of LC linkages, which are presumed to be at lower concentrations.

Enzyme as a tool to isolate and enrich LCC

The selectivity- and mild conditions of enzymatic hydrolysis are attractive for LCC isolation. Enzymatic approaches (endoglucanase, hemicellulase) have been extensively used during isolation of LCC since the first trials during the 60s.[49,50] In pioneering work, Pew49 proposed the presence of LC linkages when he realized that 5% of the carbohydrates present in finely grounded spruce wood were resistant to the action of glycosidases. More recently, the utilization of enzymes to enrich LC linkages has helped the NMR characterization. This occurs by selective cleavage of the carbohydrate chains and subsequent reduction in their respective signal intensity thereby improving the resolution of the intact LC signal which otherwise would be too low due to their low concentration.26 With this aim, a procedure producing milled wood enzymatic lignin (MWEL, Figure 2)26 by enzymatic treatment of milled wood was followed by extraction with 96% aqueous dioxane to produce the soluble cellulolytic enzyme lignin (CEL, Figure 2).[51] Although CEL was found to be structurally similar to MWL,49,51 the advantage is the higher yield, making it more representative of wood lignin than MWL. In another recent study,[52] MWLs, LCC-AcOH and CEL was prepared from *Eucalyptus grandis × E. urophylla* at different time of maturation. It was found that LCC-AcOH was enriched in phenyl glycosides. Conversely, benzyl ethers were the most predominant LCCs in CEL and MWL fractions. In other work, a similar concept to CEL is also applied only that the enzymatic treatment is performed on a regenerated ball milled wood solution which is then purified to get the product called regenerated cellulolytic enzyme lignin (RCEL, Figure 2). In this case, the associated LCC moieties are enriched when contrasted with the CEL preparations, probably due to more effective enzymatic hydrolysis.[53] Key factors were the combination of dissolution in DMSO + 6% LiCl (wt/wt), regeneration with water as anti-solvent, enzymatic hydrolysis, and extraction of the enzymatically treated residue with 80% dioxane (v/v) in water. Although mild selective hydrolysis is exerted, the approach compromises detailed structural information of the carbohydrate component of LCC. Also, when cellulase culture filtrates are applied, selectivity control is compromised. For instance, the presence of glycosidases in cellulase cultures could lead to the formation of phenyl glycosides through the transglycosylation reaction.26[54] Furthermore, it has also been suggested that β-glycosidases, present in cellulases preparations, can cleave the phenyl glycoside linkages.26

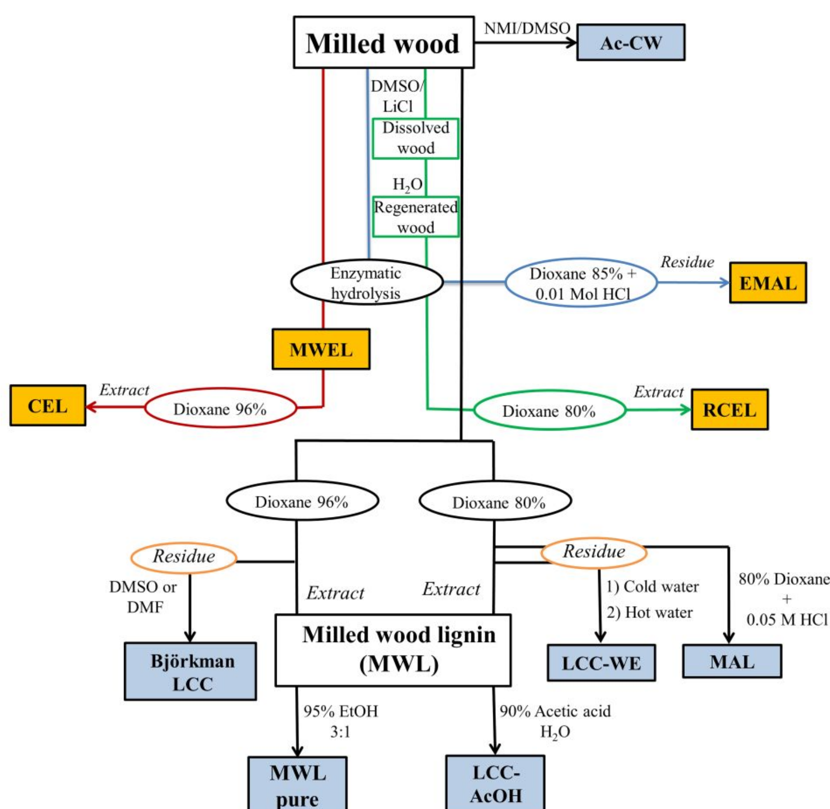


Figure 2: Isolation protocol of Lignin-Carbohydrate bonds (LCC) from wood. In yellow: lignin protocol including both ball milling and enzymatic hydrolysis; in light blue: lignin isolated after ball milling. MWL =Milled Wood Lignin, Ac-CW=Acetylated cell walls, CEL=Cellulolytic Enzyme Lignin, MWEL= Milled Wood Enzymatic Lignin, EMAL=Enzymatic Mild Acidolysis Lignin, RCEL=Regenerated Cellulolytic Enzyme Lignin, AcOH=Acetic acid, WE=Water Extract, MAL=Mild Acidolysis Lignin.

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

Green Chemistry

Critical Review

Acid and Base hydrolysis of LC bonds

As the proposed LC linkages are carbon-oxygen bridges, it is expected that at certain acidic or basic conditions they could be cleaved. Yet, for lignocellulose where reaction conditions can change quite rapidly depending on the nature of released degradation products, it is difficult to study. In the case of lignin, it is well known than condensation reactions could occur subsequent to formation of benzylic cation at certain acidic conditions leading to repolymerization.[55,56] Acid hydrolysis (0.1 N HCl, 70°C) of LC ether bonds in model compounds and in LC enriched fractions from Spruce (0.2 N HCl, 60 °C, 24 h) has been studied.[20] The difference in conditions for models and true wood substrates are clearly discernible and in support of that harsher conditions are required for the latter. Other conditions for selective acid hydrolysis of LC bonds in lignocellulose will be re-visited in connection with analytical techniques for LC bonds

Similarly, for alkaline reactions, the condition for LC bond cleavage have been investigated primarily for ester bond identification between lignin and carbohydrates after saponification reaction in sodium hydroxide (0.1 M NaOH solution, 1.5–2 h, 20°C).[26,35,57] After neutralization and centrifugation, the precipitate has been analyzed by both spectroscopic techniques such as FT-IR[26] and NMR[25] as well as chromatographic methods such as HPLC.[58]

It is however important to note that for alkaline processing of lignocellulose, reaction conditions can change due to release of various degradation products including organic acids. Such changes could favour the reformation of LC ether bonds through addition reactions to quinone methide derived through resonances with phenolate structures. Indeed, LC ethers have been proposed to exists in kraft pulps.[59,60,61] In summary, the creation of non-native LC bonds or degradation of existing LC bonds will be determined by the exact reaction conditions.

Acid hydrolysis for LCC enrichment

As said, the selectivity of acid hydrolysis of LCCs can be controlled through optimized hydrolytic conditions. Yuan et al.[62] isolated, by this method, a lignin fraction containing relatively high levels of associated carbohydrates named mild acidolysis lignin (MAL, Figure 2). The fractionation was achieved through successive treatment of the residual milled wood following extraction of MWL with 80% aqueous dioxane containing 0.05M HCl. In another approach, mild acid hydrolysis is combined with enzymatic hydrolysis to yield enzymatic mild acidolysis lignin (EMAL, Figure 2).[63] EMAL, however, was designed for procurement of purer lignin than those obtained through previous protocols, e.g., cellulolytic enzyme lignins (CEL, Figure 2) and milled wood lignin (MWL, Figure 2),[64] based on the concept of selective cleave LC bonds and hydrolysis of carbohydrate moieties.[63,64] In general, higher recovery yield could be obtained, but harsher milling conditions and/or more severe enzymatic treatments are required which could modify and/or break LCC structures.

Base hydrolysis for LCC enrichment

Optimized base hydrolysis to cleave the backbone of the constituting polymers and at same time enrich the LC bond have been studied. For instance, in a recent protocol applied on *Maize* stem,[58] sequential mild alkaline extraction, combined with endoglucanase treatment, has been adopted to isolate two alkali-soluble hemicellulose-based LCC fractions with quite distinct structural properties. After isolation of the first LC alkali soluble fraction in 0.5 M aqueous NaOH (20 °C for 24 h), the obtained residue was treated with Novozyme 476 enzyme (pH 5 for 90 h at 45 °C). Finally, the hydrolyzed solid was then extracted with 2 M NaOH to obtain a second LC enriched fraction. Subsequent alkaline hydrolyses followed by analysis by HPLC-PDA of the hydrolysates revealed that both alkali-soluble fractions contained substantial amount of ether-bound hydroxycinnamic acids to lignin. However, mass balance analysis reveals that the bulk of starting material is lost during the several hydrolysis steps since only 47% of total lignin and 25% of initial mass ended up in those two fractions. This, together with the possibility that LC ester and ether bonds, which are also relatively labile in alkali conditions,[20] could be cleaved during fractionation, makes the overall quantification of ester/ether bonds not representative of the whole starting material.

Sequential fractionation of LCC enriched fractions

Although the techniques mentioned above have been useful for LC linkage analyses by spectroscopic methods, the detailed structure of the carbohydrate component in LCC moiety is lost due to the hydrolysis by either enzymes or acids. These details include the degree of native acetyl decorations, acidic group distributions, side chain substitutions, etc. when compared to the pristine carbohydrate fractions involved in LCC. Such information is an essential contribution to fundamental studies on cell wall formation. Most of the fractionations discussed above do not target individual purified LCC components for more detailed analysis.

More recent approaches of LCC fractionation involve complete and/or sequential dissolution of finely divided wood, followed by sequential and selective precipitations of dissolved solutes achieved through anti-solvent- or salt- addition.[65]

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

Though the approaches are more tedious than those described earlier, by multiple and sequential solvents fractionation (Figure 3), all of the lignin is recovered stepwise in a quantitative way, grounding the representativeness for further characterization.

The first quantitative fractionation method of LCCs isolated from softwood wood and pulps was achieved with a high recovery yield of all LCC-lignin after selective endoglucanase hydrolysis and swelling in aqueous urea of the milled wood followed by re-dissolution in strong alkali with the addition of borate (Figure 3, block A).[66,67] Nevertheless, this method involves many steps since, in some cases, the enzyme treatment had to be repeated and it is only applicable for softwoods limiting its universality.

Subsequently, another important fractionation method was established to isolate LCC from ball milled native- and processed- hardwoods[68] and softwood.[69] Complete dissolution was achieved in 40% tetra-n-butyl ammonium hydroxide (TBAH) with dimethyl sulfoxide (DMSO) as solvent (Figure 3, block B). Later, it was shown[70] that the use of a base, namely TBAH, cleaved labile ester linkages in the biopolymers, in a study where grass species were subjected to fractionation adopting the protocol. Accordingly, to circumvent the cleavage of acid- or base- sensitive lignin-carbohydrate linkages, a near-neutral pH quantitative LCC fractionation protocol was developed[71,72] (Figure 3, block C). First, the bulk of hemicelluloses (and bounded lignin) is simply extracted with warm water from extractive-free ball milled wood (BMW). After that, a neutral solvent system, utilizing the superior dissolution power of ionic liquid [AmimCl] with DMSO as co-solvent, was adopted to obtain complete dissolution of the water-insoluble residue. Selective precipitation of various LCC fractions were obtained by sequential anti-solvent additions. Due to the preserved hemicellulose structure in both LCC and lignin-free hemicellulose, two structurally different xylan populations were defined: one with a lower degree of acetylation that was linked to lignin and the other with higher acetylation degree that was not.

More recently, in another quantitative protocol (Figure 3, block D), *Eucalyptus* LCCs were firstly extracted with 80% neutral aqueous dioxane and then with 80% acidic dioxane. Then, sequential precipitates were obtained with different ethanol concentrations (70% and 100%), and acidic water (pH = 2) with the advantage to grade precipitated hemicelluloses based LCC fractions.[73]

In general, regarding all LCC fractionation to date, the single most common criticism is the ball milling step. Detailed studies

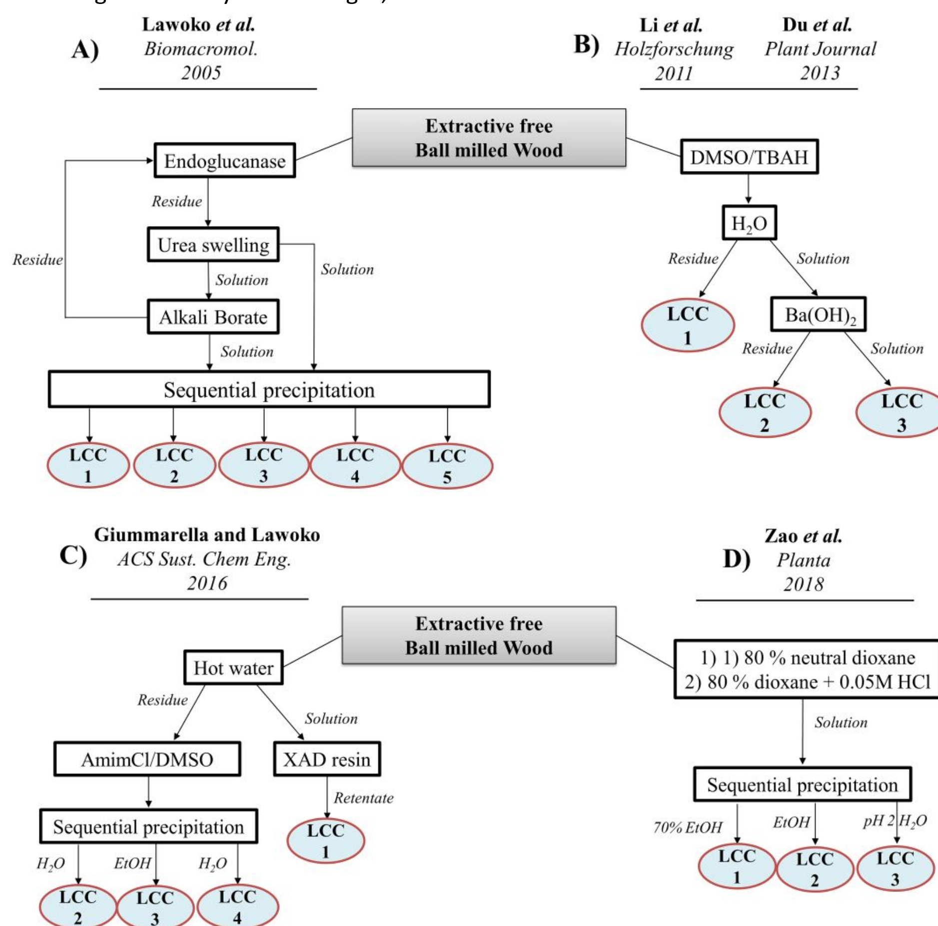


Figure 3: Review of quantitative protocols for LCCs sequential fractionation after mechanical ball milling.

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

Green Chemistry

Critical Review

are required to understand this process. Fortunately for lignin, NMR studies on biomimetically synthesized analogues[37,47,74,75] have provided evidence to the contention that lignin structure in milled wood lignins are actually native albeit a few modifications. Similar studies bio-mimicking LCC formation have been done,[76,77] but still lack the advanced structural analyses required to identify the LC linkage unequivocally. Such analysis would constitute a crucial advancement in LCC fundamental studies.

Characterization of LCC from Hot water extraction (auto-hydrolysis) processes

Hot water or subcritical water extraction has in recent years been commonly applied to extract hemicelluloses that could subsequently be converted to chemicals, fuels and polymers.
In one study, the protocol in Figure 3, block C, was adopted to study the role of LCC in recalcitrance of hemicellulose extraction during subcritical water extraction. It was suggested[78] that LC bonds were indeed preserved at extraction conditions of 160 °C, 2 h. The phenyl glycosides LCC were simply dissolved, ending up in the liquid phase while the benzyl ethers and esters remained in the residual wood, apparently responsible for hemicellulose recalcitrance. In other related work,[79] when auto-hydrolysis was instead performed at 170 °C, benzyl ethers and esters were dissolved in addition to the phenyl glycoside, indicating the non-stability of the LC linkages at such high temperatures. The dissolution of LCC fragments was an indication that fragmentation of lignin structure and other linkages in the hemicelluloses had occurred. Recently, Tarasov et al.[80] applied a flow through reactor to study auto-hydrolysis of spruce wood chips. From the gel permeation chromatography studies, the authors proposed that the percentage of lignin associated to carbohydrates as LCC in the dissolved fraction increased when the temperature was increased from 180 °C to 190 °C. Tunc et al.[81] also applied GPC to characterized LCC in accelerated solvent extracted hemicellulose from hardwoods. A criticism of using gel permeation chromatography with dual UV/RI detection to deduce LCC however is that other non-lignin chromophores could be present in solution. This is especially true for subcritical extraction at conditions were pseudo-lignin maybe present. Pseudo-lignin has been proposed to form from recondensation of carbohydrate degradation products forming a lignin-like polymer.[82] Zhu et al.[83] designed a protocol for obtaining xylo-oligosaccharides and lignin through a sequential extraction, first with hot water and then ethanol containing 1% hydrochloric acid. Both extraction stages were performed at 180 °C. It was proposed by ^{13}C-^{1}H HSQC NMR studies that the residue after subcritical water extraction contained LCC of benzyl ether type. However, the lignin obtained in the subsequent ethanol extraction was devoid of LC ethers, presumably due to acid hydrolysis.
In other works,[71,72] the use of hydrophobic interaction resins (Amberlite XAD resins) have been applied directly on warm or hot water extracts to obtained LCC enriched fractions in the retentate and purer hemicelluloses in the permeate. It has been reported[71] that phenyl glycosides were enriched in such retentates from *Norway Spruce* extracts, consistent with reports by Narron et al. who performed their study on a hardwood.[84] In both cases, linkage analyses were carried out by using 2D NMR techniques. On the other hand, Zhang et al.[85] detected LC esters from a retentate of Spruce extract using similar NMR techniques. Mass spectroscopy could be a useful tool in studying low molar mass LCC fragments, especially in hot water extracts where they are common, depending on the processing conditions. Indeed, in a recent study, Boes et al.,[86] utilized such resins to enrich LCC fractions from auto-hydrolyzates of sweetgum and switch grass processed at 180 °C. The LCC fractions were studied by a newly developed method applying chloride as an ionization dopant in combination with tandem mass spectroscopy, to reveal a unique type of saturated LC ester. The concept of this unique analysis is based on selective tagging of carbohydrate fraction with chloride. It is reported that the basicity of lignin in the gas phase prohibits chloride tagging[87] while that of carbohydrates permits the formation of stable anionic adducts.[88]

Indirect analysis of LC bonds: historical overview and challenges

This section will outline the indirect methods of Lignin-Carbohydrate (LC) bond analyses. Generally, the approaches adopt protective and selective degradation chemistries as well as selective functionalization at the LC binding site for analysis. Even though such approaches are consistent, the selectivity is not fully accepted, hence still debatable.

Indirect evidence of LCCs existence

The most common indirect approach consists of selective acid hydrolysis of the benzyl ether (BE) bond[19,20,25] or saponification (alkaline hydrolysis) of γ-ester linkages between glucuronic acid and lignin.[25,26,57,89] However, the questionable selectivity of these indirect methods is manifested in the lack of detailed structural evidence. For instance, the presence of bonds such as lignin-based alkyl benzyl ethers (Figure 4)[54,90] can contribute to the increase of aliphatic benzyl hydroxyls not

Please do not adjust margins

Please do not adjust margins

created by cleavage of benzyl ethers LC bonds as proposed for the

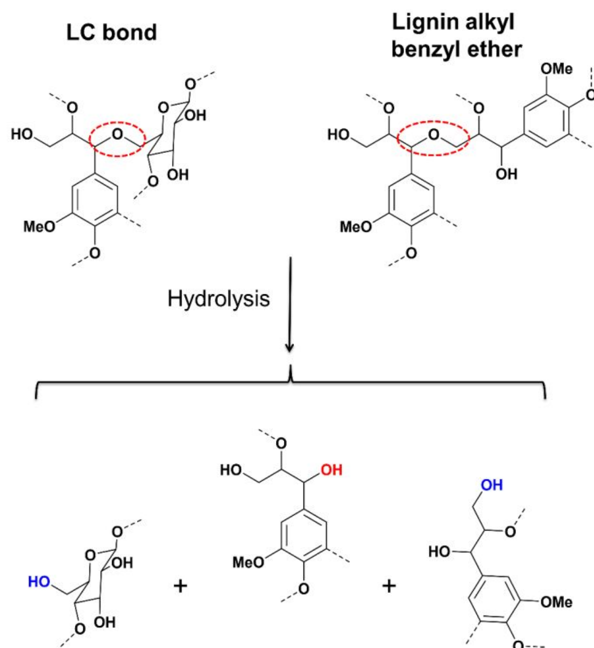


Figure 4: Hydrolysis of the benzyl ether (BE) bond in LC (up) and lignin-alkyl benzyl ether structure such as α–O–γ. Selectivity of the approach is arguable since in both cases, the cleavage in Cα leads to the formation of new benzyl hydroxyl (red).

abovementioned techniques. Similar arguments could be made for phenyl coumaran and pinoresinol structures.

Smith degradation, consisting of periodate oxidation (IO_4^-), reduction (BH_4-), and mild hydrolysis[91,92] has also been applied on LCCs substrate to detect LC bonds (Figure 5).21,57 This technique is conventionally utilized in structural investigations of polysaccharides. In lignin carbohydrate (LC) linkage study context, the glycosidic bonds in carbohydrate moiety of the LCC are first oxidized to acyclic acetal bonds which are more susceptible to mild acid hydrolysis and can thus be selectively targeted. The hydrolysis products include glycerol from xylose units, erythritol from six-carbon sugars and importantly enriched LC oligomers. However, the efficiency of oxidation has been questioned and several loops of oxidation may be required. More importantly, no LC linkage analysis was done, and deduction of LC bond is made through routine sugar analysis of the product.

Permethylation is a well-known analytical protocol adopted to study the linkage pattern in carbohydrates[93,94] and to detect methyl ester of uronic acid which is otherwise degraded in harsh acidic condition when not protected. This method has been adopted to detect LC ether bonds both in pulps[95] and wood.[96,97,98,99] However, the procedure is tedious, involving several steps. It starts with permethylation of enzyme-hydrolyzed ball milled wood, performed using CH_3I in DMSO, followed by mild acid hydrolysis of glycosidic bonds which yields new hydroxyls. These products are reduced with labelled borodeuteride which avoids mutarotation of the monosugars and marks the reducing end. Finally, acetylation of the new hydroxyl group distinguished them from the original ones and was also used to access LC ether binding site

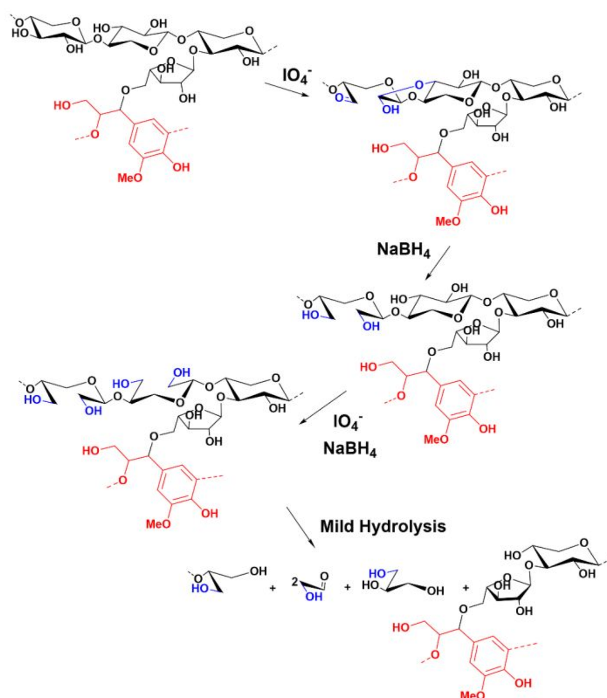


Figure 5: Graphical representation of Smith degradation pattern for Benzyl ether LC to C5 in arabinose attached to xylan chain adapted from reference 57.

on sugar moiety.

The conclusion was that these LC ethers were formed through C6-hydroxyls in hexoses and C5-hydroxyls in arabinose. The analysis was performed by gas chromatography with flame ionization detection (GC-FID) and/or mass spectrometry (GC-MS). A major problem in attempting permethylation analysis on wood or pulp substrates is the poor solubility in any suitable solvent which requires that the methylation step be repeated at least twice. As a consequence, the yields and the quantification are directly proportional to the solubility of the substrates.[95] The LC linkage analysis by this method is questionable since the method does not take into account the native acetylation of hemicellulose at carbon-2 and carbon-3 positions. In addition, it assumes that the C6 or C5 hydroxyls in native structures are either free or attached to lignin only.

DDQ Oxidation

Generally, the disadvantage of the indirect methods previously described all target only the carbohydrates side of LCCs. Indirect elucidations on the binding sites between lignin and carbohydrates evaluating both sides were obtained when acetylated cellulose-rich LCC fragments were subjected to oxidation which was suggested to be selective for benzylic (α-) ether22 and ester-24 bonds. The oxidation was performed with 2,3 dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) to yield α carbonyls on the lignin backbone[5] which can be quantified by ^{13}C NMR. The DDQ-oxidized polysaccharide fraction, carrying a new hydroxyl group, was methylated by the Prehm procedure,[94] hydrolyzed with 2M TFA, reduced with $NaBH_4$

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

and then analyzed by GC MS after acetylation. By this method, qualitative as well as quantitative information could be obtained since the position and the relative proportion of groups in the mono-sugars reveals the position where lignin is attached through benzyl ether and therefore, the frequency of benzyl ethers (BE). Unlike other methods, DDQ oxidation was proposed to target specifically benzyl ethers and esters showing structural differences from both lignin and carbohydrate side. However, when the DDQ method was applied to acetylated MWL,[100] the selectivity and the completeness of oxidation were lower when compared with those obtained in the presence of LCC model compounds leading to uncertain outcomes by some researchers.[101]

Size Exclusion Chromatography

Size Exclusion Chromatography (SEC) studies have been used as an indirect proof of LC linkages. Using a dual detector system, consisting of a differential refractive index (RI) and ultraviolet (UV), the co-elution patterns of carbohydrates (detected by RI) and lignin (detected by both detectors) imply a possible linkage between them. In other words, with the known concentrations of the carbohydrate and lignin in each fraction as a gauge, superimposed RI and UV signals imply either that the hydrodynamic volumes of the carbohydrates and lignin are equal or that the lignin is chemically linked to the carbohydrates.[66,102] In addition, it must be considered that solubility, consequent to extraction/fractionation of the LC substrates, together with the SEC system adopted leads to different outcomes. For instance, a comparative study focused on the chemical interaction between wood biopolymers,[103] showed that Intrinsic Viscosity-Differential Pressure (IV-DP) detector, being proportional to the intrinsic viscosity of dissolved solutes, is the most selective for linear free carbohydrates and LCC. Conversely, it has been shown that Right Angle Laser Light Scattering (RALLS) detector is more sensitive to high molar mass polymers with significant contribution to the signal intensity from branching points.[104] To sum up, the selectivity of different detectors towards different polymers considering their mass, shape and composition can be triggered by LC isolation adopted. However, no insight about the nature and amount of lignin–carbohydrate linkages present in the analysed materials could be obtained by SEC analysis.

Fourier transform infrared spectroscopy (FT-IR)

Fourier transform infrared spectroscopy (FT-IR) is a useful technique in determining the functional groups of the substrate investigated. The advantage is that the technique is rapid, nondestructive and highly sensitive.

In several work,[46,105,106,107,108,109] LCC enriched substrates have been characterized unveiling mostly the typical functional groups contained in the bulk of lignin and carbohydrates, such as hydroxyl, carbonyl, methoxyl, carboxyl, aromatic and aliphatic C-H outlining the structural motifs of the isolated fractions. It has been reported that absorption peak at 1721 cm^{-1} could be due to the C=O stretching indicating a ester linkage between lignin and carbohydrates[110] while the signal at 1170 cm^{-1} could be assigned to the C–O stretching in C–O–C in phenyl glycoside linkage.[105] In addition, Liu et al.[107] reported that the disappearance of absorption bands at 1728 and 1156 cm^{-1}, assigned to C=O stretching in unconjugated ketone and carboxyl groups and C−O stretching in ester groups, is a sign of benzyl esters cleavage. In another work,[109] ^{13}C labeled coniferin was injected into rice stalk to label the side chain of lignin. FT-IR, combined with ^{13}C-NMR spectrum, was adopted to obtain structural elucidation of the linkages at α position in lignin side with carbohydrate. It was concluded that the linkages between lignin and carbohydrate included a benzyl ether, an ester and an acetal bond. However, FT-IR hardly can reveal undoubtedly the linkage between lignin and carbohydrates. For instance, the absorbance peak at around 1730 cm^{-1} of esterified carbonyl groups, may be due to the acetyl groups attached to hemicellulose as well as esters of glucuronic acid to lignin.

All in all, FT-IR provides a quick “fingerprint” of the LC substrates, looking at the constituting polymers, but when it comes to their connectivity, it has to be combined with other analytical tool such as NMR to show unequivocally the LC linkage.

Direct analysis of LC bonds by NMR techniques

NMR of LCC: overview and challenges

Deeper insights on the molecular structure of LCC can be obtained by liquid state NMR which is, by far, the most effective tool[44,47,69,70,71,72,73,111,112,113,114] utilized for characterizing isolated wood biopolymers. Among all NMR techniques, 2D NMR is the most utilized. The main advantage of 2D NMR techniques, such as heteronuclear single quantum coherence spectroscopy (HSQC), is the higher resolution when compared with 1D NMR analysis. This result is because overlapping signals of protons directly linked to carbons having different chemical shifts are separated in the carbon dimension; similarly, in the proton dimension, carbons signals with similar chemical shift can be separated by the environment of the directly attached protons. Another advantage of HSQC[48,90,115,116,117,118] is the possibility to decipher simultaneously structural details of lignin skeleton and carbohydrates structure[119,120] as well as their connectivity, LC bonds so to speak.[44,47,52,70,71,72,73,111,112,113,114,121] In addition, diagnostic 2D NMR techniques such as total correlation spectroscopy (TOCSY) and heteronuclear multiple bond correlation (HMBC) are both indispensable tools for unravelling the spectra of complex polymeric and multifunctional molecules providing additional evidence for the HSQC assignments. HSQC-TOCSY provides cross peaks through bond correlations between virtually all spins in the same spin system, whereas HMBC correlates protons and carbon separated up to three bonds but has the added advantage of analysis of connectivity between two spin systems, which makes it attractive as a diagnostic for lignin-carbohydrate linkages. Another NMR technique well-established in the field of lignin analytical chemistry is ^{31}P NMR

spectroscopy[122] which is a powerful analytical tool for phenolic-, aliphatic- and carboxylic- hydroxyls[122] and can been used to assess these functionalities in LCCs substrates.

Although ^{13}C NMR is used to provide structural and quantitative information of the carbohydrate and/or lignin part of the LCCs, it is hardly utilized for LC bond detection because their resonance signals are heavily overlapped with those originating from lignin and/or carbohydrate signals.[44,47] However, ^{13}C NMR can be combined with HSQC analysis for quantitation[116] since 2D NMR experiment *per se* is not quantitative. In other words, the intensities of cross peaks in HSQC spectra are not directly proportional to analyte concentration. This result is because, during HSQC experiment, the efficiency of the polarization transfer through J-coupling of ^{13}C and ^{1}H nucleus depends not only on the value of the ^{13}C/^{1}H coupling constant but also on specific resonance signal attenuation during the pulse sequence such as T_1 and T_2 relaxation times, peak multiplicity and homo-nuclear coupling. Nevertheless, in the last decades, efforts have been made to render HSQC analysis quantitative by *ad hoc* pulse sequence which achieve, at same time, a reasonable signal to noise ratio (S/N) within a reasonable experiment time. This is the case of Quick, Quantitative HSQC (QQ-HSQC)[123] an improved version of four-times longer Q-HSQC[124] and Q-CAHSQC.[125,126] In QQ-HSQC the efficiency of the transfer steps is constant over a range of heteronuclear coupling constants for a number of scans as a conventional HSQC experiment. In an alternative and more recent approach[127], the time-zero ^{13}C HSQC ($HSQC_0$) spectrum, obtained by extrapolating peak intensities from a series of measured HSQC spectra acquired with incremented repetition times, is quantitative since the cross-peaks are linearly proportional to the concentrations of the analytes. $HSQC_0$ method seems to be simpler to implement but requires more time than the QQ-HSQC method.[128] However, both methods have been successfully adopted in MWL analysis[124,128,129] showing that both quantitative HSQC methods deliver very similar, and thus comparable results. Nevertheless, for more complex substrates such as LCC, such quantitation could be compromised by overlapping signals.

Setting the standard for unequivocal proof of linkages in lignocellulosics: the case of lignin

The assignments of lignin inter-units by HSQC technique is supported by robust diagnostic NMR such as multiple bonds connection (HSQC-TOCSY, HMBC) and 3D NMR analysis.[130,131,132] For instance, details on the connectivity at the atomistic level going from one spin system to the next have been thoroughly studied for milled wood lignins.[132,133] More recent examples include structures such as spirodienone[134] and dibenzodioxocin[135] (see Figure 6). In contrast to lignin analysis, analysis of LC bonds by 2D NMR methods requires long experimental time due to the low LC linkage abundance. Therefore, the need for high sensitivity NMR spectrometer, possibly equipped with a cryogenic probe, is crucial for identification and especially for quantification of LC linkages. Alternatively, LCC enrichment and purification of the linkage can be used to enhance signals. Selectively removing the bulk of carbohydrates in an LCC sample can also solve the problem of fast relaxation times of T_2. In the case of polymeric lignin bound to polymeric carbohydrates, the macromolecule tumble very slowly in solution and T_2 becomes too short for long range correlation as required in an HMBC experiment or for efficient spin-lock needed in TOCSY program. In other words, both HMBC and TOCSY require long mixing times but if T_2 is shorter than the mixing time, d6 (delay for the evolution of long-range couplings) and d9 (TOCSY mixing time), respectively, there will be no signal.

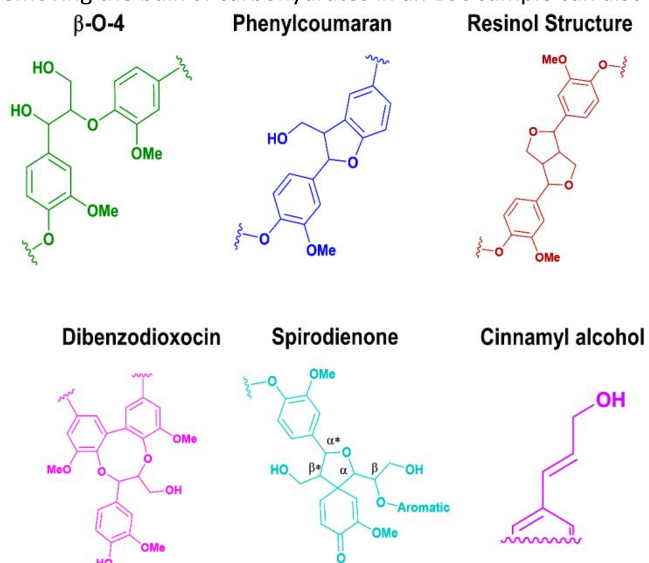


Figure 6: Main inter-unit linkages and terminal end in lignin.

Lignin-Carbohydrate (LC bond) analysis by 2D NMR

NMR studies on LCC model compounds have established spectral regimes assigned to LC linkages that have been extremely important for studies on LCC substrates isolated from wood.[136,137,138,139,140,141,142,143] However, a number of critical issues need to be resolved to upgrade LC linkage studies to unequivocal status. In this section, we will critically address the state of the art of lignin-carbohydrate linkage studies in which NMR characterization has been used as an analytical tool.

Model compound studies. In general, model studies have been beneficial in mapping resonance regimes associated with LC linkages. These regimes can only be approximate since the spectral features of the models are too simplistic when compared with those of true biomass substrates consisting of macromolecular lignin and carbohydrate signals, were possible overlap and neighbouring group effects could dramatically affect the data interpretation.

- Benzyl ethers LCC models. Toikka and Brunow synthesized and studied model substrates of benzyl ethers and their acetylated analogues.[139,140] More specifically, benzyl ether to primary alcohol such as C_5 arabinose and C_6 galactose as well as to secondary alcohol C_3 in galactose, C_2 and C_3 on xylose (Table 1) were thoroughly characterized, including the

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

effects of stereochemistry (erythro- and threo forms) on chemical shifts, by 1D ^{1}H and ^{13}C NMR.139,140 Moreover, synthesis of benzyl ethers to methyl β-D-glucoside has also been investigated by similar techniques providing ^{1}H and ^{13}C NMR assignment in $CDCl_3$.140

- Phenyl glycosides LCC models. A comprehensive NMR study of various phenyglycosides models was recently published.[112] The HSQC study clearly showed the broad distribution in chemical shifts of phenyl glycoside linkages (i.e., C/H correlation of the anomeric carbon C_1 coupled to the 4-O-linkage on monolignol) depending on the sugar involved (mannose, xylose, glucose, galactose), the monolignol type (i.e., Guaiacyl, Syringyl or Coumaryl units) and aliphatic side chain structure (conjugated, oxidized or reduced) of monolignol. Accordingly, the phenyl glycosides linkage signal, viewed from the anomeric carbon on sugar units, were assigned within the $^{13}C/^{1}H$ regime of 97-104 ppm/4.7-5.4 ppm for HSQC experiments performed using DMSO-*d6* as solvent (Table 1). More advanced studies applying HMQC, HMBC and COSY techniques have also been done for model compound analogs to phenyl glycosides involving p-glucocoumaryl alcohol, coniferin, and syringin.[142]

Table 1: ^{13}C and ^{1}H assignments of LC linkages observed from model compound studies.

	Benzyl ether139,140			
	Cα/Hα (ppm) in Acetone-*d6*			
	βO4 erythro1	βO4 erythro2	βO4 threo1	βO4 threo2
To C5 Arabinose	82.4/4.63	82.3/4.65	82.7/4.68	82.3/4.69
To C6 Galactose	82.9/4.61	82.8/4.61	82.9/4.66	83.0/4.66
To C3 Galactose	76.3/4.94	-	77.1/4.94	-
To C2 Xylose	82.6/5.06	80.8/5.13	82.7/5.12	80.6/5.21
To C3 Xylose	81.9/5.13	81.8/5.09	82.2/5.21	82.6/5.25
To C4 Xylose	82.9/4.86	79.5/4.82	83.4/4.98	80.2/4.83

	Phenyl glycoside[112]		
	C1/H1 (ppm) in DMSO-*d6*		
	Coniferyl	Sinapyl	p-Coumaryl
To C1 Glucose	100.3/4.83	102.9/4.82	100.6/4.79
To C1 Galactose	101/4.79	103.7/4.74	101.2/4.75
To C1 Xylose	101.3/4.81	102.8/4.88	101.2/4.78
To C1 Mannose	99.0/4.93	101.7/4.81	97.9/5.07

	γ–ester[143]	
	Cγ/Hγ (ppm) in Acetone-*d6*:D2O (9:1)	
	βO4 erythro	βO4 threo
To 4OMeGluA	64.88/4.43*	64.89/4.04,4.37

4OMeGlu=4-O-methyl-d-glucuronic acid; * Signal of $H\gamma_1$ and $H\gamma_2$ overlaps

- Lignin-Carbohydrate (LC) Esters: LC esters model compounds representative of the hypothesized linkage between lignin and 4-O-methylglucuronic acid present in glucuronoxylan have been synthesized and studied by 1D NMR techniques.[143] It was shown that both benzyl esters and γ-esters (coupled through Cγ of lignin) were present (Table 1). Initially, the benzyl ester was formed and then migrated to the γ carbon, through a transition state involving a 6-member ring. This reaction outcome is further discussed in the next section of this review.

Mechanisms of lignin-carbohydrate bond formation

The proposed mechanism of formation of lignin-carbohydrate ethers and esters is captured in Figure 7. Lignin polymerization begins with the radical coupling of mildly oxidizable monolignols and is predominated by couplings involving β-radicals, yielding mainly β–O-4 units but also significant amounts of β–β and β–5 couplings.[144] These couplings result in the formation of a quinone methide intermediate, an electrophile, which is subjected to nucleophilic attack from available nucleophile (Figure 7). When radicals couple to form

Please do not adjust margins

β–β and β–5, internal trapping of the electrophilic site by closely proximal phenolic hydroxyls is the predominant reaction and leads to the formation of resinol and coumaran structures respectively (Figure 7). When they couple to form β–O-4 on the other hand, several nucleophile candidates are present, but NMR studies support the dominance of water addition to yield α-hydroxylated β-O-4. The addition of carboxylic or aliphatic hydroxyls in carbohydrates would yield the formation of benzylic-ester and ether LCC, respectively.

Evtuguin and co-workers[111] incorporated ^{13}C enriched coniferin and syringin into the newly formed xylem of eucalyptus. ^{13}C enrichment was specifically done on the benzylic- and β-carbons of the monolignol since the former was the assumed coupling site and the latter the neighboring linkage (see Scheme 1 for assumed LCC structure). This approach is unique as no extraneous enzymes initiating the polymerization of lignin were added. The formed lignin was then studied after ball milling of the xylem tissue by solid state ^{13}C CP/MAS NMR, and in the liquid state following extraction with DMSO-*d6*. The solution state analysis included 1D and 2D techniques such as HSQC and HMBC. Benzyl ether LC linkage was suggested from the HSQC study but not confirmed by the HMBC. On the other hand, γ–ester LCC was substantiated by the HMBC studies showing the required evidence of correlations between two spin systems, i.e., the lignin spin system and a second spin system through ester bridge (Scheme 1). However, details on the second spin system, i.e., the atom connectivity, were not shown by HMBC.

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

Green Chemistry

Critical Review

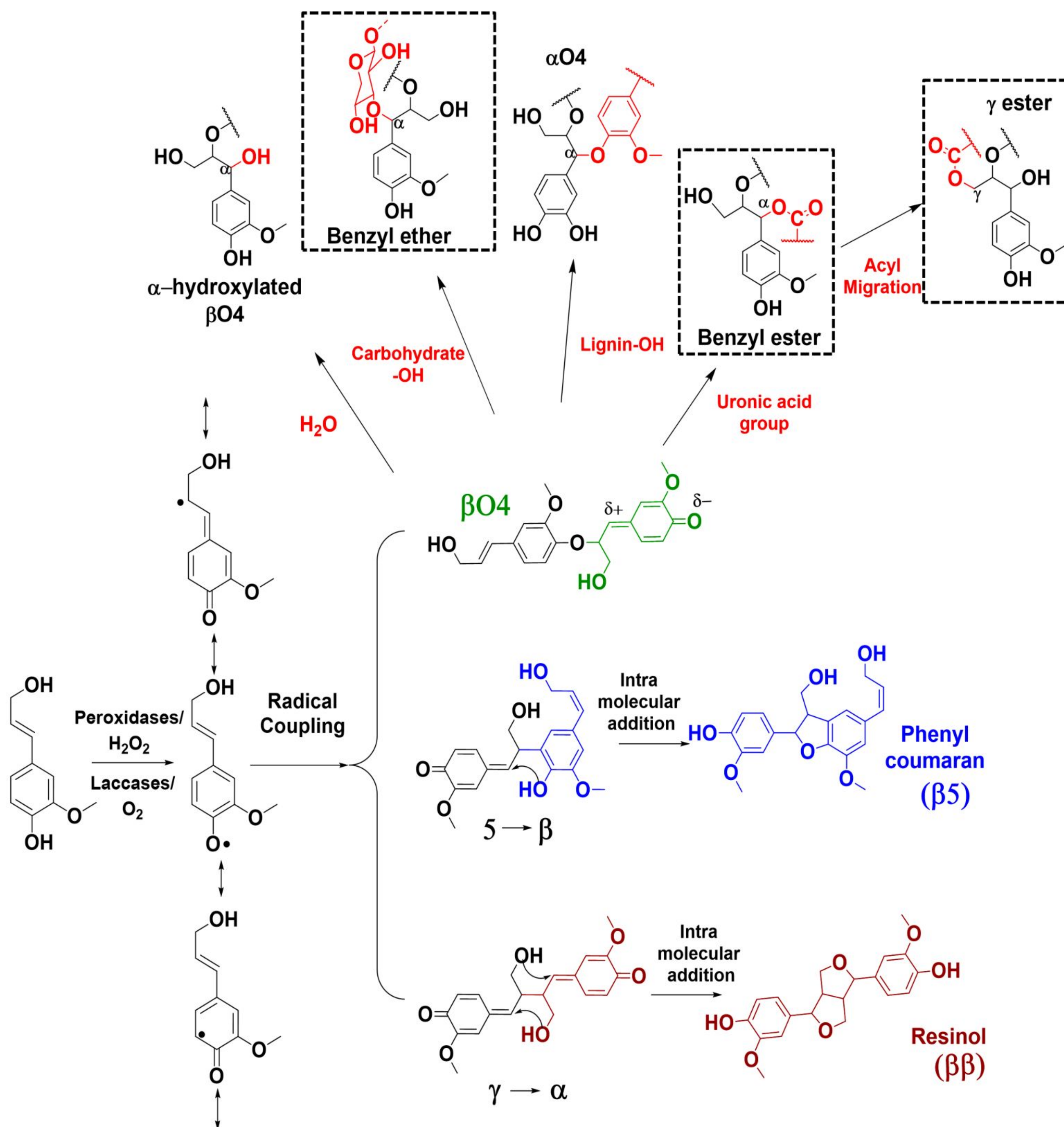


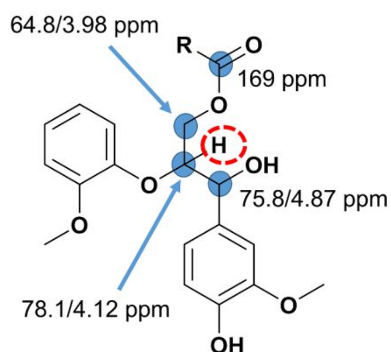
Figure 7: Scheme showing proposed mechanisms of formation of LC ethers and esters.

Please do not adjust margins

Please do not adjust margins

Green Chemistry

Critical Review



Scheme 1: Multiple bond correlation showing γ-esters. Proton in β is shown to correlate with the α and γ carbons in the lignin spin system and with the carbonyl carbon in the ester bridge. Schematic illustration of the assignments in accordance with reference 111.

Formation of phenyl glycosides has barely been discussed in the literature. One speculative mechanism is the acid-catalyzed addition of a phenolic hydroxyl to the reducing end of a carbohydrate moiety by the well-known chemistry of hemiacetal formation (Figure 8, up). Acetal formation is catalyzed by slightly acidic conditions, which would be consistent with what has been reported for pH in plant cell walls (pH~5).[145] Besides, acidic conditions could exist locally in the proximity of acidic polysaccharides like glucuronoxylan and pectins. Phenyl glycosides may also form through transglycosylation of a phenolic end group to the reducing end of a polysaccharide catalyzed by one of the multiple types of transglycosylating enzymes (Figure 8, bottom). These are known to cleave and relegate hemicellulose and also cellulose chains during plant development.[17,146,147]

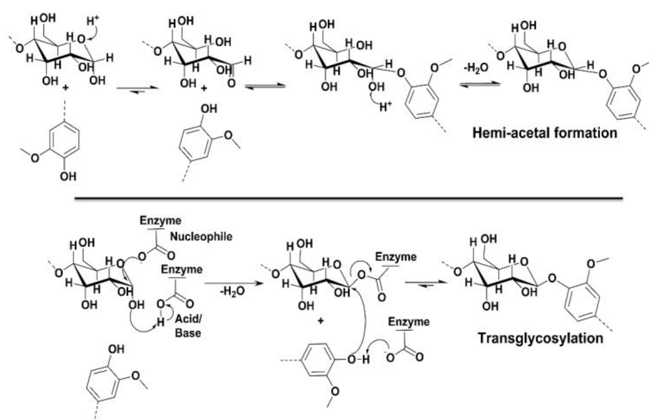


Figure 8: Hypothetical scheme of phenyl glycoside formation by hemiacetal formation (up) or by transglycosylation of a phenolic end group to carbohydrate reducing end catalyzed by one of the multiple types of transglycosylating enzymes (bottom).

NMR studies on biomimetic- synthesized LCCs: Original work

Esters are formed by the specific addition of the acidic groups to quinone methide at Cα in quinone methide intermediate, yet, NMR analysis of isolates have almost exclusively suggested Cγ linked esters. It has been suggested that a uronosyl migration from Cα to Cγ occurs through a six-member ring transition state.[111,143] A valid question is if the migration occurs during isolation procedure or during lignin polymerization.

To address this, we present new data from our attempts to biomimic LC ester formation. Furthermore, such a study would also shed light on nativity of LC bonds as ball milling is circumvented. In one case, galacturonic acid (GalA) was included in the biomimetic synthesis of lignin starting with coniferyl alcohol and horseradish peroxidase as the oxidant. In another case, glucuronoxylan (Xylan) was used still with the same monolignol and oxidant. The reference consisted of only monolignol and oxidant, which produces synthetic lignin, commonly referred to as dehydrogenation polymer (DHP). The products were studied by HSQC (Figure 9). Depending on whether an α-ester or γ-ester was formed, a remarkable deshielding caused by the presence of an electron withdrawing group should cause downfield shifts in the Cα/Hα or Cγ/Hγ. In the case of the reference (Figure 9a), the expected chemical shift of Cα/Hα in Cα-hydroxylated β-aryl ether structure which appears at 71.2/4.76 ppm (Signal β–O-4_{α}) is observed. The HSQC spectrum of the lignin-glucuronic acid derivative on the otherhand clearly shows a new signal at 74.8/5.92 ppm assigned to Cα/Hα in Cα-esterified β-aryl ether structure (Signal BEst, Figure 9c). The Cα-hydroxylated β-aryl ether structure however still dominated (Figure 9c). Unlike the Cγ/Hγ correlation in Cγ–esterified β-aryl ether structure, which appears in a heavily overlapped region, the Cα/Hα in Cα-esterified β-aryl ether structure is easily discernible in the spectra. In the case of the lignin-glucuronoxylan derivative (Figure 9d), the signal at 74.8/5.92 ppm was again observed (Signal BEst), but a new suspect signal for γ-esters also appeared at 63.0/4.3 ppm in the overlapped methylene region (Signal GE). Therefore, it seems that if the previously proposed ester migration occurs, the reason could be to mitigate sterical constraints of bulky molecules such as polymeric glucuronoxylan. Such constraints would support the migration to less hindered Cγ. Indeed, no such migration was observed in the case of the lignin-galacturonic acid derivative (Figure 9c). The galacturonic acid molecule is much smaller. Due to the weak signals, we were unable to confirm the LC ester linkage unequivocally by HMBC or HSQC-TOCSY for this set of experiments. However, our recent studies showed that when stronger carboxylate nucleophiles were used in a similar biomimic approach, the esterification efficiencies were

This journal is © The Royal Society of Chemistry 20xx

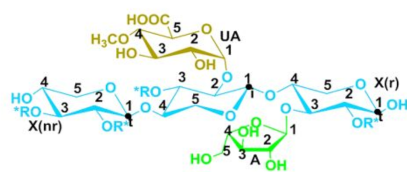
Please do not adjust margins

Please do not adjust margins

ROYAL SOCIETY OF CHEMISTRY

Green Chemistry

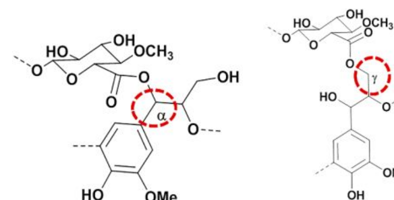
Critical Review



A=Arabinose;; X= Xylose; U=4-O-Methyl Glucuronic Acid
R*=H or Acetyl group; t=terminal; i=internal;
r=reducing; nr=non reducing end:

Abbreviation

β04 = Benzyl aryl ether
β5 = Phenyl coumaran
ββ= Resinol structure
SD = Spirodienone
CA = Cinnamyl alcohol
DBDO = Dibenzodioxocin
$-OCH_3$ = Methoxy



BEst = Benzyl ester

GE = γ ester

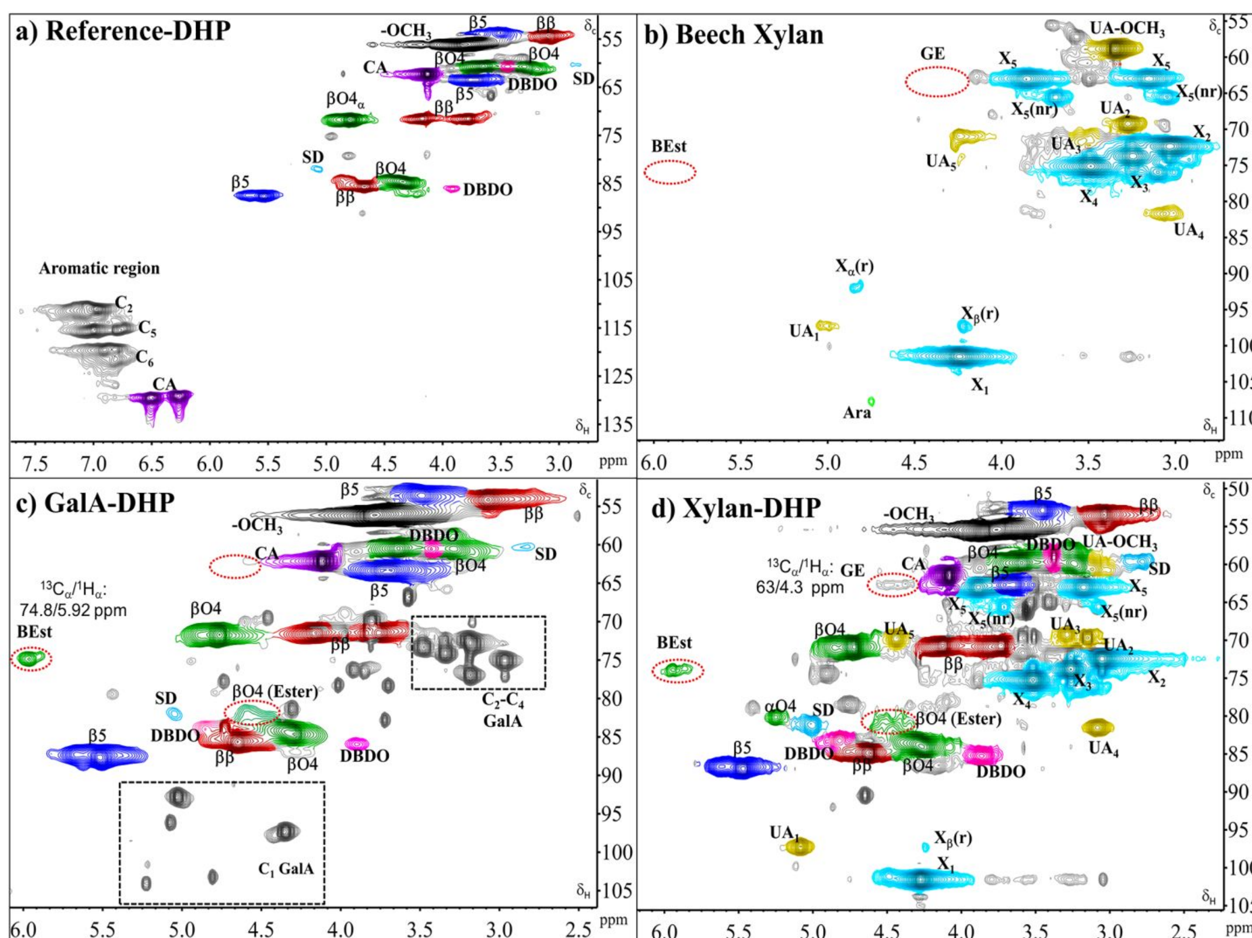


Figure 9: Expanded HSQC spectrum in DMSO-*d6* of the a) Reference-synthetic lignin (DHP) b) glucuronoxylan from beech c) Synthetic lignin produced in the presence of galacturonate (GalA-DHP) and d) Synthetic lignin produced in the presence of beech glucuronoxylan (Xylan-DHP). Presence and absence of characteristic LCCs are marked by red dotted circles. Colors used are expressed in Figure 6.

improved, and the signal at 74.8/5.92 ppm was confirmed by both HMBC and HSQC-TOCSY to be Cα/Hα in Cα-esterified β-aryl ether structures.[148]

Ferulate Ester in grass

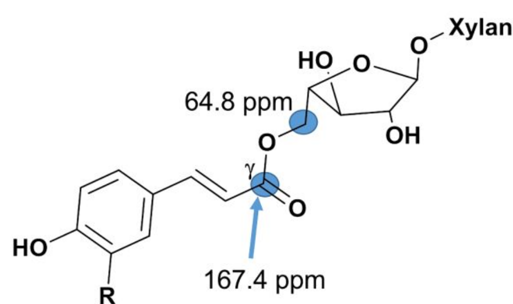
Many angiosperms such as *Graminea* and *Caryophylales* contain esterified ferulic acid (FA) and to some extent, *p*-coumaric (pCa) structures covalently bound to cell wall

Please do not adjust margins

polymers such as polysaccharides. The linkage to polysaccharide is highly specific.[33] The synthesis of feruloylarabinoxylan (FA-AX), whose substitutions of FA-Ara is regulated by the activity of Arap mutase, occurs within the Golgi in conjunction with polysaccharides synthesis. Thereafter FA-AX is transported to the extracellular space and shuttled
into the wall matrix.[149] The analysis of the FA-Ara linkage of the monolignol-sugar linkage was established by gas chromatography combined with ^{1}H and ^{13}C NMR after pre-treatment of the entire cell wall cell wall, or the constituting polysaccharides with purified endoxylanases.[33] This linkage analysis was also applied to barley straw,[32] sugar cane bagasse,[27] and Zea shoots,[28] yielding feruloylated and to some extent p-coumaroylated structures at C5 of arabinose that was in turn attached to xylan, the major hemicellulosic polysaccharide in grass.[30,150] In later approaches, the hydrolysis with mild TFA hydrolysis (0.05-0.1 M aqueous TFA, 100 °C, 1−3 h) selectively cleaves only glycosidic bonds but not the ester linkages between arabinose and feruloylate or coumarate.[151,152] An HMBC experiment performed on representative model compounds has clearly shown the correlation between C9 in lignin and one of the H5 protons of arabinose creating unambiguous assignments based on short- and long-range heteronuclear ^{13}C/^{1}H shift-correlated experiments (Scheme 2).[153,154]
In grasses, the presence of feruloylarabinoxylan provides a convenient and reliable tool of cross-linking polysaccharides.[155] From model studies using feruloylated arabinose (FA-Ara) and plant peroxidases, ferulate esters were easily oxidized by the peroxidase. Then dimerization occurs via phenoxy radicals leading to the formation of dehydrodiferulates such as 5–5, β–5, β–β, β–O-4, a 4-O-5 structure predicted from free-radical coupling chemistry reactions.[156,157,158,159] Active incorporation of ferulate polysaccharide esters into ryegrass lignins was supported by the work of Ralph and co-workers in which HMBC NMR study was applied to uniformly ^{13}C-labeled ryegrass.[160] The HMBC spectrum showed coupling in β position of ferulate with both coniferyl and sinapyl alcohol. Also, ferulates were bio-mimetically incorporated into synthetic lignin (dehydrogenation polymer, DHP) showing that β−5, β−β and β–O-4 structure could form. Recently,[161] hydroxycinnamate acylation patterns in arabinoxylans was evaluated by a more robust method consisting of mild acidolysis in dioxane, methanol, and aqueous HCl 2M (60:30:10, in volume) mixture. The quantification was performed by using HPLC-UV, HPLC-MS and GC-MS detection. The main advantage of this new method is the high yield of dioxane lignin (60−70% of total lignin) and the mildness of the hydrolysis condition leading to minimal structural changes in the native lignins.[162]

Ball milled wood substrates

The steep improvement of sensitivity in 2D NMR provides, upon solubilization, a broad fingerprint of the structural feature of constituting biopolymers as well as their

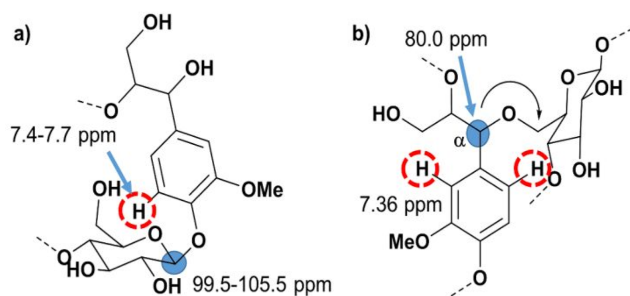


Scheme 2: Schematic illustration of the ^{13}C assignment of FA-Ara in acetone-*d6* given in reference 155 and confirmed by HMBC.

connectivity. Adopting HSQC analysis on ball milled substrates obtained from hardwoods and softwoods, the first assignments of lignin-carbohydrate bonds, specifically phenyl glycoside (PG), 44,[163] benzyl ether (BE)44,47 and γ-ester (GE) LC linkages44,47 were reported about two decades ago. The first HSQC assignment for phenyl glycosides on wood substrates was reported for *Eucalyptus globulus*.[163] Later on, in a study
performed on Loblolly pine wood, Balakshin et al.44 assigned a group of signals at ^{13}C/^{1}H: 107.1–99.5/ 5.36–4.94 ppm to various anomeric carbons in carbohydrate involved in phenyl glycoside bonds with lignin, in accordance with previous work.[142] In the same work,44 signals at ^{13}C/^{1}H: 82.5–80.0/4.7–4.3 ppm were assigned to benzyl ethers accordingly with model studies.[140]
Lastly, although in an overlapped HSQC region, signals at ^{13}C/^{1}H: 62–65/4.0–4.5 ppm were assigned to γ-esters. The authors investigated an LCC fraction isolated from ball milled Loblolly pine (see Figure 2, LCC-AcOH). It was claimed that the HMQC assignment of these structures was supported by HMBC technique.
This claim, however, is not fully supported by the analysis for the following reasons;

- The signals at ^{13}C/^{1}H: 105.5–99.5/7.69–7.42 ppm in HMBC spectrum that was assigned to phenyl glycosides based on the assumption that the correlation was between aromatic proton attached to C5 of lignin and the C_1 of the carbohydrate. The number of bonds between the said proton and the C_1 of the carbohydrate is four (Scheme 3a). This is an unusual occurrence in HMBC which is typically a multiple bond correlation up to 3 bonds. In addition, the chemical shift on the proton axis of 7.69–7.42 ppm that was assigned to H5 on the aromatic ring is significantly more downfield than the expected chemical shift for H_5 (6.6-6.7 ppm).[132,133]
- Regarding benzyl ether assignment of the HMBC spectrum, the analysis would have been more compelling if the signal of benzylic carbon in α-ether (Cα: 80.0 ppm) could show a long-range correlation to the next spin system of carbohydrates rather than to the signal of aromatic itself at H_2 and/or H_6 at 7.36 ppm (Scheme 3b). Besides, several signals from lignin side chains appear in the HSQC region

between 80-85/3.8-4.8 ppm which would show similar correlations as those reported44 and shown in Scheme 3b.

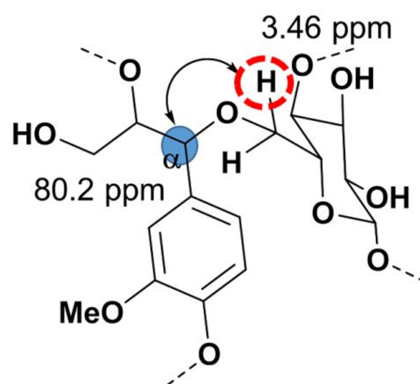


Scheme 3: Graphical representation of a) long distance correlation (4 bonds) between C1 in sugar units with H5 in lignin of phenyl glycosides and b) HMBC correlation of benzyl ether where Cα correlates with H2/H6. Both assignments are discussed in reference 44.

- The existence of GE esters was proposed from a few HMBC correlations signals between C6 in uronic acid and protons in γ of lignin at $^{13}C/^{1}H$: 166.5–169.4/4.30–4.35 ppm. However, when observing the HSQC of the same sample, several signals appear in the region 4.3-4.5 ppm on the proton axis, making it impossible to assign to the γ protons unambiguously. Similar to the work by Evtuguin et al.[111] (summarized in Scheme 1), if the long-range correlation between the carbonyl carbon and protons on adjacent carbons, i.e., C5 or C4 on the uronic acid, was shown, a stronger justification for LC esters could be concluded.

Recent work by Watanabe and coworkers[114] presented the first unequivocal direct NMR evidence of benzyl ether in wood isolates. More specifically, the benzyl ether-linkage between the C6 hydroxyl of mannose unit in glucomannan and Cα position of lignin was shown. They adopted an NMR analytical toolbox consisting of HSQC, edited TOCSY, HMBC and, most importantly, 3D TOCSY-HSQC; a combination enabling complete assignment of connectivity of the LC linkage and all atoms in close proximity to it. In their work, an LCC fraction from ball milled Japanese red pine (LCC-WE, Figure 2) was further purified.

The strategy involved improving signal intensities by enrichment of LC bonds through glycohydrolase digestion of the carbohydrates units, followed by usage of polyvinyl gel, hydrophobic lignin-like gel to further separate carbohydrates from enriched LCC fraction. This approach tackled the problems of both low frequency of LC bonds and size of macromolecule analyzed. The latter is a crucial issue during NMR analysis especially when investigating macromolecules with fast relaxation times (T2) which render the observation of long-range correlation signals between lignin and carbohydrates challenging to achieve. Undebatable is the HMBC correlation signal of the Cα (δ_C = 80.2 ppm) and the proton of C6 in mannose (δ_H = 3.48 ppm) which was also seen symmetrically on proton axis, with the signal at δH_α of 4.50 ppm correlating with carbohydrate carbon C6 (Scheme 4). The quantification of the LC linkage was however not mentioned.



Scheme 4: a) Schematic representation of the symmetrical long-distance correlation between Cα in lignin and proton of C6 in mannose in accordance with reference 114.

The enrichment also involves a tedious and yet necessary procedure.

To summarize this section, the unequivocal assignments of the LC linkages are important since several researchers have adopted these assignments. The work approach adopted by the Watanabe group is promising and is foreseen to advance LC linkage studies.

Problems associated with signal overlap

One of the main issues of γ-ester identification and quantification by HSQC is the heavy overlapping of signals falling in the region around $^{13}C/^{1}H$: 62–65/4.0–4.5 ppm. For instance, the presence of native acetyl esters in γ position of lignin ($^{13}C/^{1}H$ in γ: 62.5/3.82 and 4.22 ppm) has been shown.[34] Their presence can be confirmed by the appearance of aliphatic signals from methyl in acetyl groups at $^{13}C/^{1}H$ = 20/1.93, 2.01 ppm and by the presence of a $^{13}C/^{1}H$ cross signal in β at 80.8/4.58 ppm. Due to the above-mentioned overlap in the regime 62-65/4.0-4.5 ppm, the $^{13}C/^{1}H$ cross signal in β at 80.8/4.58 ppm serves as a diagnostic peak of ester in the γ–position of β-O-4 lignin structure. Also, ferulate and coumarate lignin structure are known to be present in native non-wood lignins whereas benzoates structures have been found in Populus.[34] Therefore, before HSQC analysis of LC esters, it is crucial to investigate and verify the presence of other non-carbohydrate related esters, e.g., acetyl or ferulate in the original of substrate.47

Benzyl ether LC bond (BE_1) occurs between Cα of lignin side chain and primary alcohol of the carbohydrate units, C6 in hexoses or C5 in arabinose whereas BE_2 is benzyl ether towards secondary alcohol (C2 or C3) of the sugar units (Figure 1).[47] BE_2, appearing at $^{13}C/^{1}H$ in α at 80–81/4.9–5.1 ppm, overlaps with signal at $^{13}C/^{1}H$ in α* at 81.2/5.1 ppm in spirodienone sub-structure of lignin (see Figure 6 for lignin sub-structures). However, the amount of spirodienone structures can be quantified from the signal of $^{13}C/^{1}H$ in β* at 79.4/4.11 ppm.[47,118,134] Hence comparing the volume integrals of those two cross peaks, it is possible to detect and

This journal is © The Royal Society of Chemistry 20xx

semi- quantify the presence of BE_2. However, the differences in T2 relaxations could affect the signal intensities making this approach obscure. The need for diagnostic multiple bond correlation NMR experiments is beneficial so as to increase the resolution of the analysis. Accordingly, 2D techniques such as HMBC and 3D NMR provide more solid proof of linkages. This approach forms the scientific core of first unequivocal proof of LC benzyl ethers[114] Indeed, such experiments would solve the inability to distinguish the signals of benzyl ether between lignin aliphatic chains structures in close proximity, such as the signal in β of γ acylated β-aryl ethers appearing at $^{13}C/^{1}H$: 80.8/4.58 ppm, from the newly proven BE_1 LC type whose resonance cross peak of $C\alpha/H\alpha$ is at 80.2/4.52 ppm.[114]

Benzyl ethers have also been identified in residual lignins of pine kraft pulps[164] by HMQC technique based on the $C\alpha/H\alpha$ correlation at 81/4.6-4.8 ppm. Other studies on softwood kraft lignins applying HSQC have also assigned this particular signal the benzyl ethers.[59,60] Recent interest in lignin for material applications have however prompted the necessity for deeper understanding of the structure of technical lignins. One such study of kraft lignin using a combination of NMR techniques assigned the aforementioned signal of benzyl ethers to epiresinol structures.[165] The presence and amounts of LC ethers in kraft lignin and residual pulp lignin remains therefore an open question.

HSQC signals of PG ($^{13}C/^{1}H$: 4.85/101.3 ppm and $^{13}C/^{1}H$: 4.88/102.6 ppm) involving guaiacyl and syringyl units, respectively[112] may overlap with the internal anomeric signals from rhamnose and galacturonic acid in pectins at $^{13}C/^{1}H$: 100.8/5.02 ppm, and the anomeric carbon in esterified glucuronic acid at $^{13}C/^{1}H$: 100.1-101/4.6-4.7.[44,47] This makes the identification of PG challenging especially for pectin enriched substrates. A table summarizing the notorious overlaps compromising previous LC linkage identifications and quantitation is shown in Table 2.

With a resolution in the sub-nanometer scale, solid-state NMR (ss-NMR) for analysis of interaction between plant cell wall polymers has recently gained ground. The advantage here is that the polymers are unperturbed in analogue to liquid state NMR analysis. Accordingly, interactions between pectins, hemicelluloses and cellulose microfibrils have been studied. By adopting multi-dimensional ^{13}C correlation solid-state NMR, close interactions between pectin and cellulose, specifically involving rhamnogalacturonan and homogalacturonan, were proposed in Arabidopsis.[166] In other work, ^{13}C magic-angle spinning (MAS) were supported by in silico modelling of chemical shifts to reveal that xylan in Arabidopsis had two conformations: a threefold and twofold helical screw. The latter was proposed to bind to cellulose while the former did not. It was proposed the twofold screw gives the segment a flattened conformation with ability to interact with cellulose surfaces.[167]

Indeed, in a recent study adopting 2D solid-state ^{13}C–^{13}C correlation INADEQUATE (Incredible Natural Abundance DoublE QUAntum Transfer Experiment) NMR technique on ^{13}C-labeled maize, switchgrass and Arabidopsis stems, the three-fold screw xylan with a non-flattened conformation was shown to interact with lignin through electrostatic forces.[168] More specifically, hydroxyl groups in xylan were shown to interact with methoxyl groups in lignin. Furthermore, such interactions were shown to be more favorable for syringyl units in lignin than guaiacyl units, which in turn interacted better than p-hydroxyphenol units.[168] It was concluded that electrostatic forces governed interactions between lignin and xylan and that covalent interaction, if present, were of minimal importance.

Table 2: $^{13}C/^{1}H$ HSQC assignments in DMSO-*d6* of LC linkages detected in ball milled woody substrates with possible overlapping signals.

LC linkages	^{13}C-^{1}H Assignment	Overlapping signals and ^{13}C-^{1}H Assignment
Benzyl ether 1 (BE_1) **To C_6 in sugars**	**Cα/Hα** 80.2/4.52 ppm[114] 80.0–82.5/4.3–4.7 ppm44 80-81/4.5-4.7 ppm[70]	γ acylated βO4 Cβ/Hβ: 80.8/4.58 ppm[70] Epiresinol Cβ/Hβ: 81.2/4.77 ppm[165]
Benzyl Ether 2 (BE_2) **To C_2 /C_3 xylan**	**Cα/Hα** 80.0/5.08 ppm[70] 81.7/5.04 ppm[71] 81.3/5.07 ppm[73]	Spirodienone Cα∗/Hα∗: 81.2/5.1 ppm[47,]52·118·134
Phenyl glycosides **To C_1 sugar**	**C1/H1** 107.1–99.5/ 5.36–4.94 ppm44 99-104/4.8-5.2 ppm[70,71] 100.4/5.02 ppm[72] 100.6-101.5/4.65-4.79[73]	Rha to GalA[47,]69 C1 to C2: 100.8/5.02 ppm GluA[47,]69 C_1/H_1 : 100.1-101/4.6-4.7 ppm
γ-esters **To 4OMeGlu in Xylan**	**Cγ/Hγ** 62–65/4.0–4.5 ppm44	Acetylated lignin Cγ/Hγ : 62.7/3.83 and 4.30 ppm44·154

4OMeGlu=4-O-methyl-D-glucuronic acid. α∗ = signal in α∗ (see Figure 6).

On the other hand, one could view such electrostatic forces as the first step in bringing the lignin and xylan in the close proximity required for covalent bonding. All in all, the advent of solid-state NMR studies of interactions in the sub-nanometer scale is a positive development in connection to resolving the intricacies surrounding LCC studies.

A future outlook on LCC analysis studies

At this point when several biorefinery concepts are being developed to address environmental and sustainability issues, studies on lignin-carbohydrate bonds are essential due to their hypothesized contribution to recalcitrance. Emerging technical fractionation processes for lignocellulosic biomass to obtain the constitutional polymers in high yield and purity would benefit from such essential fundamental studies. As discussed in this review, progress in related analytics of the lignin-carbohydrate linkage has been slow. This is due to challenges in characterizing heterogeneous biomass substrate in combination with the low frequency of lignin-carbohydrate linkages. Therefore, without getting sound analytics, a continued sluggish development is foreseen. The superior analytical technique is NMR, and this needs to be coupled with novel sample preparation which enriches the lignin-carbohydrate bond and the issue of fast relaxation times of T_2 allowing multiple bond correlations analysis such as HMBC. The enrichment should preferably be mild and highly selective. The use of enzymes is a promising approach as recently demonstrated.[114] Scalable production of pure enzymes is becoming more common, and the use of enzyme "cocktails" with tailored selectivity is a future consideration for lignin-carbohydrate linkage studies. Alternatively, mild chemical methods with high selectivity could be used to achieve such enrichment.

Qualitatively, the NMR spectroscopic studies for lignin-carbohydrate bond analysis needs to overcome signal overlap issues and show connectivity between the two spin systems that extends at least a couple of atoms away from the linkage. Hence a variety of 2D NMR studies are required. More specifically, HSQC/HMQC, HSQC-TOCSY, and HMBC are needed. Also, 3D ^{1}H–^{13}C TOCSY-HSQC NMR could extend the span of atom connectivity. On the quantitative analysis, challenges remain as standard 2D HSQC NMR methods are semi-quantitative and tailor-made pulse sequences are needed to render the measurements quantitative. However, quantitative ^{13}C NMR could be applied together with HSQC to obtain more accurate values. Such methods exist for lignin[116] and could be optimized for lignin-carbohydrate linkages.[47]

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

A part of this study was supported by the Knut and Alice Wallenberg Foundation which is gratefully acknowledged for financial support to the Wallenberg Wood Science Center.
In addition, this study was partially supported by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. This study was supported and performed as part of the Center for BioEnergy Innovation (CBI). CBI is the U.S. Department of Energy Bioenergy Research Centers supported by the Office of Biological and Environmental Research in the DOE Office of Science. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Neither the United States Government nor any agency thereof, nor any of their

Please do not adjust margins

employees, make any warranty, expressed or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. The publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

Notes and reference

Please do not adjust margins

1 A. J. Ragauskas, C. K. Williams, B. H. Davison, G. Britovsek, J. Cairney, C. A. Eckert, W. J. Frederick Jr, J. P. Hallett, D. J. Leak, C. L. Liotta, R. Jonathan, J. R. Mielenz, R. Murphy, R. Templer, and T. Tschaplinski. The Path Forward for Biofuels and Biomaterials. *Science*, 2006, **311**(5760), 484–489.
2 X. Meng, A. Parikh, N. Nagane, B. Seemala, R. Kumar, C. M. Cai, Y. Pu, C. E. Wyman, and A. J. Ragauskas. Chemical transformations of poplar lignin during co-solvent enhanced lignocellulosic fractionation process. ACS Sustainable Chemistry & Engineering, 2018, **6**, 7, 8711–8718.
3 J. Erdman. *Annalen der Chemie und Pharmacie*, 1866, **138**, 1.
4 F. E. Brauns and H. Seiler. Reactions with homogenized spruce wood. *Tappi,* 1952, **35**, 67-71.
5 T. Koshijima and T. Watanabe. *Association Between Lignin and Carbohydrates in Wood and Other Plant Tissues.* Springer Series in Wood Science. Springer, Berlin, Heidelberg, 2003.
6 A. Björkman. Studies on finely divided wood. Part 1. Extraction of lignin with neutral solvents. *Svensk. Papperstidn.* 1956, **59**, 477-485.
7 M. Lawoko, G. Henriksson and G. Gellerstedt. Characterization of lignin-carbohydrate complexes from spruce sulphite pulp. *Holzforchung*, 2006, **60**, 162–165.
8 M. Lawoko, R. Berggren, F. Berthold, G. Henriksson, and G. Gellerstedt. Changes in the Lignin-Carbohydrate complex in softwood kraft pulp during kraft- and oxygen delignification. Lignin-polysaccharide networks. *Holzforchung,* 2004, **58**, 6, 603–610.
9 G. Gellerstedt and E.-L. Lindfors. Structural changes in lignin during kraft pulping. *Holzforschung*, 1984, **38**, 151-158.
10 T. Yamasaki, S. Hosoya, C.-L. Chen, J.S. Gratzl and H.-m. Chang. Characterization of residual lignin in kraft pulps. 1st International Symposium on Wood and Pulping Chemistry. Stockholm, Sweden. Proceeding, **2**, 34-42.
11 B. Hortling, M. Ranua, and J. Sundquist. Investigation of residual lignin in chemical pulps. *Nordic Pulp Pap Res. J.*, 1990, **5**, 33-37.
12 M. Tenkanen, T. Tamminen and B. Hortling Investigation of lignin-carbohydrate complexes in kraft pulps by selective enzymatic treatments. *Appl. Microbiol. Biotechnol.,* 1999, **51**, 241-248.
13 J. A. Smelstorius. Chemical composition of wood of Australian-grown Pinus radiata. III. Lignin-polysaccharide complexes. *Holzforschung,* 1974, **28**, 3, 99-101.
14 R. Kondo, T. Sako, T. Iimori, and H. Imamura. Formation of glycosidic lignin-carbohydrate complex in the dehydrogenative polymerization of coniferyl alcohol. *Mokuzai Gakkaishi,* 1990, **36**, 332-338.
15 J. P. Joseleau and R. Kesaraoui. Glycosidic bonds between lignin and carbohydrates. *Holzforschung,* 1986, **40**, 163-168.
16 B. Kosikova, D. Joniak, and J. Skamla. Lignin-carbohydrate bonds in beech wood. *J. Cellul. Chem. Technology*, 1972, **6**, 579-588.
17 F. Yaku, Y. Yamada and T. Koshijima. Lignin-carbohydrate complex. Part II. Enzymatic degradation of acidic polysaccharide in Björkman LCC. *Holzforschung,* 1976, **30**, 148-156.
18 K. Freudenberg and G. Grion. Contribution to the mechanism of formation of lignin and of the lignin-carbohydrate bond. *Chem. Ber.,* 1959, **92**, 1355-1363.
19 Ö. Eriksson and B.O. Lindgren. About the linkage between lignin and hemicelluloses in wood. *Svensk Papperstidn.,* 1977, 59-63.
20 B. Kosikova, D. Joniak, and L. Kosikova. On the properties of benzyl ether bonds in lignin-saccharidic complex isolated from spruce. *Holzforschung,* 1979, **33**, 11-14.
21 F. Yaku, R. Tanaka, and T. Koshijima. Lignin carbohydrate

This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

complex. Part IV. Lignin as side chain of the carbohydrate in Björkman LCC. *Holzforschung,* 1981, **35**, 4, 177-181.
22 T. Koshijima, T. Watanabe and J. Azuma. Existence of benzylated carbohydrate moiety in lignin-carbohydrate complex from pine wood. *Chemistry Letters,* 1984, **13**, 10, 1737-1740.
23 K. Freudenberg and J. M. Harkin. Models for the linkage of lignin to carbohydrates. *Chem. Ber.,* 1960, **93**, 2814-2819.
24 T. Watanabe and T. Koshijima. Evidence for an ester linkage between lignin and glucuronic acid in lignin-carbohydrate complexes by DDQ-oxidation. *Agric. Biol. Chem.,* 1988, **52**, 11, 2953-2955.
25 K. Lundquist, R. Simonson, and K. Tingsvik. Studies on lignin carbohydrates linkages in milled wood lignin preparations. *Svensk Papperstidn.*, 1983, **86**, 6, R44-R47.
26 J. R. Obst. Frequency and alkali resistance of lignin-carbohydrate bonds in wood. *Tappi,* 1982, **65**, 4, 109-112.
27 A. Kato, J. Azuma, and T. Koshijima. A new feruloylated trisaccharide from bagasse. *Chemistry Letters 12*, 1983, **1**, 137-140.
28 Y. Kato and D. J. Nevins. Isolation and identification of O-(5-O-feruloyl-α-L-arabinofuranosyl)-1 (→ 3)-O-β-D-xylopyranosyl-(1→ 4)-D-xylopyranose as a component of Zea shoot cell-walls. *Carbohydr. Res.*, 1985, **137**, 139-150.
29 A. Kato, J. Azuma, and T. Koshijima. Isolation and identification of a new feruloylated tetrasaccharide from bagasse lignin-carbohydrate complex containing phenolic acid. *Agr. and Biol. Chem.,* 1987, **51**, 6, 1691-1693.
30 T. Ishii and T. Hiroi. Isolation and characterization of feruloylated arabinoxylan oligosaccharides from bamboo shoot cell walls. *Carbohydr. Res.*, 1990, **196**, 175–183
31 M. M. Smith and R. D. Hartley. Occurrence and nature of ferulic acid substitution of cell-wall polysaccharides in graminaceous plants. *Carbohydr. Res*, 1983, **118**, 65-80.
32 I. Mueller-Harvey, R. D. Hartley, P. J. Harris, and E. H. Curzon. Linkage of p-coumaroyl and feruloyl groups to cell-wall polysaccharides of barley straw. *Carbohydr. Res*, 1986, **148**, 1, 71-85.
33 E. Yamamoto, G. H. Bokelman and N. G. Lewis. *Phenylpropanoid metabolism in cell walls: an overview*. In ACS Symposium series-American Chemical Society (USA). 1989.
34 J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P.F Schatz, J.M Marita, R. Hatfield, S.A Ralph, J.H. Christensen, and W. Boerjan. Lignins: natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochem Rev*., 2004, **3**, 29–60
35 H. I. Bolker. A lignin carbohydrate bond as revealed by infra-red spectroscopy. *Nature*, 1963, **197**, 4866, 489.
36 Y. Xie, S. Yasuda, H. Wu, and H. Liu. Analysis of the structure of lignin-carbohydrate complexes by specific 13C tracer method. *J. Wood Sci*., 2000, **46**, 130–136.
37 K. M. Holtman, H. M. Chang, H. Jameel and J. F. Kadla. Quantitative 13C NMR characterization of milled wood lignins isolated by different milling techniques. *J. of Wood Chem. and Techn.*, 2006, **26**, 1, 21-34.
38 A. Björkman. Studies on finely divided wood. Part 3. Extraction of lignin-carbohydrate complexes with neutral solvents. *Svensk Papperstidn*. 1957, **60**, 243–251.
39 T. Koshijima, T. Taniguchi, T. Tanaka. Lignin carbohydrate complex. I. Influence of milling od wood upon Bjorkman LCC, *Holzforschung*, 1972, **26**, 211-217
40 Y. Lai, K. Sarkanen. Isolation and structural studies. In: Lignins; K.V. Sarkanen, C. Ludwing. Ed: Wiley-Interscience: New York, 1971; 165-240.
41 R. C. Sun, J. M. Fang, and J. Tomkinson. Fractional isolation and structural characterization of lignins from oil palm trunk and empty fruit bunch fibers. *Journal of wood Chem and Tech.,* 1999, **19**, 4, 335-356.
42 S. Yang, T. Q. Yuan, and R. C. Sun. Structural elucidation of whole lignin in cell walls of triploid of Populus tomentosa Carr. *ACS Sustainable Chem. Eng.,* 2016, **4**, 3, 1006-1015.
43 T. Watanabe, J. Azuma, and T. Koshijima. A convenient method for preparing lignin-carbohydrate complex from Pinus densiflora wood. *J. of the Japan Wood Res. Soc.*, 1987, Japan.
44 M. Y. Balakshin, E. A. Capanema, and H.-m. Chang. MWL fraction with a high concentration of lignin-carbohydrate linkages: Isolation and 2D NMR spectroscopic analysis. *Holzforschung*, 2007, **61**, 1–7.
45 M. Y. Balakshin, E.A. Capanema, and H.-m. Chang. Recent advances in isolation and analysis of lignins and lignin-carbohydrate complexes. In: Characterization of Lignocellulosics. Ed. Q. T. Hu. Blackwell, Oxford, 148–166, 2008.
46 P. P Yue, Y. J. Hu, G. Q. Fu, C. X. Sun, M. F. Li, F. Peng, and R. C. Sun. Structural Differences between the Lignin-Carbohydrate Complexes (LCCs) from 2-and 24-Month-Old Bamboo (Neosinocalamus affinis). *Intern. Journ. of mol. sciences*, 2018, **19**, 1, 1.
47 M. Y. Balakshin, E. A. Capanema, H. Gracz, H.-m. Chang and H. Jameel. Quantification of lignin-carbohydrate linkages with high-resolution NMR spectroscopy. *Planta*, 2011, **233**, 1097–1110.
48 F. Lu and J. Ralph. Non-degradative dissolution and acetylation of ball-milled plant cell walls: high resolution solution-state NMR. *Plant J.,* 2003, **35**, 535–544.
49 J. C. Pew. Properties of powdered wood and isolation of lignin by cellulolytic enzymes. *Tappi*, 1957, **40**, 553-558.
50 J. C. Pew, P. Weyna. Fine grinding, enzyme digestion, and the lignin-cellulose bonds in wood. *Tappi*, 1962, **45**, 247-256.
51 H.-m. Chang, E.B. Cowling, W. Brown, E. Adler, G. Miksche. Comparative Studies on Cellulolytic Enzyme Lignin and Milled Wood Lignin of Sweetgum and Spruce. *Holzforschung,* 1975, **29**, 153-159
52 B. C. Zhao, B. Y. Chen, S. Yang, T. Q. Yuan, A. Charlton, and R. C. Sun. Structural variation of lignin and lignin–carbohydrate complex in eucalyptus grandis× E. urophylla during its growth process. *ACS Sustainable Chem. Eng.,* 2016, **5**, 1, 1113-1122.
53 E. Capanema, M. Balakshin, R. Katahira, H-m. Chang and H. Jameel. How Well Do MWL and CEL Preparations Represent the Whole Hardwood Lignin? *Jour. of Wood Chem. and Techn.*, 2015, **35**, 17–26, 2015
54 M. Y. Balakshin, E. A. Capanema, and A. Berlin. Isolation and Analysis of Lignin–Carbohydrate Complexes Preparations with Traditional and Advanced Methods: A Review." In: Studies in Natural Products Chemistry. Ed: Elsevier, 2014, 42, 83-115
55 J. Li, G. Henriksson, and G. Gellerstedt. Lignin depolymerization/repolymerization and its critical role for delignifcation of aspen wood by steam explosion. *Bioresour. Techn.,* 2007, **98**, 3061–3068.
56 T. J. McDonough. The chemistry of organosolv lignin. *Tappi Journal.* 1993, **76**, 186–193
57 Ö. Eriksson, D. A. I. Goring, and B. O Lindgren. Structural studies on the chemical bonds between lignins and carbohydrates in Spruce Wood. *Wood Sci. Technology,* 1980, **14**, 267-279.
58 M. H. Sipponen, C. Lapierre, V. Mechin, and S. Baumberger. Isolation of structurally distinct lignin–carbohydrate fractions from maize stem by sequential alkaline extractions and

 This journal is © The Royal Society of Chemistry 20xx

Please do not adjust margins

endoglucanase treatment. *Bioresour. Technol.*, 2013, **133**, 522–8.
59 D. J. Nicholson, A.T. Leavitt, B. Stromberg, and R.C. Francis, Mechanistic differences between kraft and SODA-AQ pulping of hardwoods with regard to lignin–carbohydrate complexes (LCC), *Jounal of Wood Chem. and Technol.,* 2017, **37**, 4, 307-322.
60 C. Crestini, H. Lange, M. Sette, and D. S. Argyropoulos. On the structure of softwood kraft lignin. *Green Chem.*, 2017, **19**, 4104-4121.
61 N. Giummarella, G. Henriksson, L. Salmén, and M. Lawoko. On the effect of hemicellulose removal on cellulose lignin interactions. *Nordic Pulp & Paper Research Journal,* 2017, **32**, 4, 542-549.
62 T. Q. Yuan, S. N. Sun, F. Xu, and R. C. Sun. Characterization of lignin structures and lignin-carbohydrate complex (LCC) linkages by quantitative 13C and 2D HSQC NMR spectroscopy. *J. of Agric. Food Chem.*, 2011, **59**, 19, 10604-10614.
63 S. Wu. D. S. Argyropoulos. An improved method for isolating lignin in high yield and purity. *J. Pulp Paper Sci*., 2003, **29**, 235–240.
64 A. Guerra, I. Filpponen, L. A Lucia, D. S. Argyropoulos. Comparative evaluation of three lignin isolation protocols for various wood species. *J. Agric. Food Chem.*, 2006, **54**, 9696–9705.
65 M. Lawoko. Unveiling the structure and ultrastructure of lignin carbohydrate complexes in softwoods. *Int. J. Biol. Macromol.,* 2013, **62**, 705–713.
66 M. Lawoko, G. Henriksson and G. Gellerstedt. New method for quantitative preparation of lignin-carbohydrate complex from unbleached softwood kraft pulp: lignin–polysaccharide networks I. *Holzforschung*, 2003, **57**, 69–74.
67 M. Lawoko, G. Henriksson and G. Gellerstedt. Structural differences between the lignin-carbohydrate complexes present in wood and chemical pulps. *Biomacromolecules*, 2005, **6**, 6, 3467–3473.
68 J. Li, R. Martin-Sampedro, C. Pedrazzi, and G. Gellerstedt. Fractionation and characterization of lignin-carbohydrate complexes (LCCs) from eucalyptus fibers. *Holzforschung,* 2011, **65**, 43–50.
69 X. Du, G. Gellerstedt and J. Li. Universal fractionation of lignin-carbohydrate complexes (LCCs) from lignocellulosic biomass: an example using spruce wood. *Plant J.*, 2013, **74**, 2, 328–338.
70 J. C. del Rio, P. Prinsen, M. Cadena Edith, A. T. Martınez, A. Gutierrez and J. Rencoret. Lignin–carbohydrate complexes from sisal (Agave sisalana) and abaca (Musa textilis): chemical composition and structural modifications during the isolation process. *Planta*, 2016, **243**, 5, 1143–1158
71 N. Giummarella, L. Zhang, G. Henriksson, and M. Lawoko. Structural features of mildly fractionated lignin carbohydrate complexes (LCC) from spruce. *RSC Advances*, 2016, **6**, 48, 42120-42131
72 N. Giummarella and M. Lawoko. Structural basis for the formation and regulation of lignin-xylan bonds in Birch. *ACS Sustain. Chem. Eng*, 2016, **4**, 5319–5326
73 B. C. Zhao, J. D. Xu, B. Y. Chen, X. F. Cao, T. Q. Yuan, S. F. Wang, A. Charlton, R. C. Sun. Selective precipitation and characterization of lignin-carbohydrate complexes (LCCs) from Eucalyptus. *Planta*, 2018, **247**, 5, 1077-1087.
74 Z. Hu, T. F. Yeh, and H.-m. Chang, Y. Matsumoto and J. F. Kadla. Elucidation of the structure of cellulolytic enzyme lignin. *Holzforschung*, 2006, 60, 4, 389-397.
75 T. Ikeda, K. Holtman, J. F. Kadla, H.-m. Chang, and H. Jameel. Studies on the effect of ball milling on lignin structure using a modified DFRC method. *J. of Agric. Food Chem*, 2002, **50**, 1, 129-135.
76 A. Barakat, H. Winter, C. Rondeau-Mouro, B. Saake, B. Chabbert, B. Cathala. Studies of xylan interactions and cross-linking to synthetic lignins formed by bulk and end-wise polymerization: a model study of lignin carbohydrate complex formation. *Planta*, 2007, **226**, 1, 267-281.
77 B. Cathala, B. Chabbert, C. Joly, P. Dole, B. Monties. Synthesis, characterisation and water sorption properties of pectin-dehydrogenation polymer (lignin model compound) complex. *Phytochemistry*, 2001, **56**, 195-202.
78 N. Giummarella, and M. Lawoko. Structural insights on recalcitrance during hydrothermal hemicellulose extraction from wood. *ACS Sustainable Chemistry & Engineering,* 2017, **5**, 6, 5156-5165.
79 A. Martínez-Abad, N. Giummarella, M. Lawoko, and F. Vilaplana. Differences in extractability under subcritical water reveal interconnected hemicellulose and lignin recalcitrance in birch hardwoods. *Green Chemistry*, 2018, **20**, 2534-2546.
80 D. Tarasov, M. Leitch, and P. Fatehi. Flow through autohydrolysis of spruce wood chips and lignin carbohydrate complex formation. *Cellulose*, 2018, **25**, 2, 1377-1393.
81 M. S. Tunc, M. Lawoko, M., and A. R. van Heiningen. Understanding the limitations of removal of hemicelluloses during autohydrolysis of a mixture of southern hardwoods. *BioResources*, 2010, **5**, 1, 356-371.
82 S. D. Shinde, X. Meng, R. Kumar, and A. J. Ragauskas. Recent advances in understanding the pseudo-lignin formation in a lignocellulosic biorefinery. *Green Chemistry*, 2018, **20**, 10, 2192-2205.
83 M. Q. Zhu, J. L Wen, Y. Q. Su, Y, Q. Wei, and R. C. Sun. Effect of structural changes of lignin during the autohydrolysis and organosolv pretreatment on Eucommia ulmoides Oliver for an effective enzymatic hydrolysis. *Bioresource technology,* 2015, **185**, 378-385.
84 R. H. Narron, H. M. Chang, H. Jameel, and S. Park. Soluble Lignin Recovered from Biorefinery Pretreatment Hydrolyzate Characterized by Lignin–Carbohydrate Complexes. *ACS Sustainable Chem. Eng.*, 2017**, 5**, 11, 10763-10771.
85 Y. Zhang, S. Wang, W. Xu, F. Cheng, A. Pranovich, A. I. Smeds, S. Wilför and C. Xu. Valorization of lignin-carbohydrate complexes from hydrolysates of Norway spruce: Efficient separation, structural characterization, and antioxidant activity. *ACS Sustainable Chem. Eng*., 2019, **7**, 1, 1447-1456
86 K. S. Boes, R. H. Narron, S. Park, and N. R. Vinueza. Mass Spectrometry Exposes Undocumented Lignin Carbohydrate Complexes in Biorefinery Pretreatment Stream. *ACS Sustainable Chem. Eng.*, 2018, **6** ,8, 10654–10659.
87 Y. Cai, and R. B. Cole. Stabilization of Anionic Adducts in Negative Ion Electrospray Mass Spectrometry. *Anal. Chem.* 2002, **74**, 5, 985−991.
88 N. R. Vinueza, V. A. Gallardo, J. F. Klimek, N. C. Carpita, and H. I. Kenttämaa. Analysis of xyloglucans by ambient chloride attachment ionization tandem mass spectrometry. *Carbohydrate polymers*, 2013, **98**, 1, 1203-1213.
89 N. Takakashi and T. Koshijima. Ester linkages between lignin and glucuronoxylan in lignin-carbohydrate complex from beech (fagus crenata) wood. *Wood Sci. Technology*, 1988, **22**, 231-241.
90 E. A. Capanema, M.Yu. Balakshin, J.F. Kadla. A comprehensive approach for quantitative lignin characterization by nmr spectroscopy. *J. Agric. Food Chem.,* 2004, **52**, 7, 1850-1860.
91 I. J. Goldstein, G. W. Hay, B. A. Lewis and B. F. Smith. Controlled degradation of polysaccharides by periodate

oxidation, reduction, and hydrolysis. Methods in Carbohydrate Chemistry, 1965, **5**, 361-370.
92 G. W. Hay, B. A. Lewis, and F. Smith. *Methods in carbohydrate chemistry. Vol. V*. Edited by R. L. Whistler. Academic Press, Inc., New York. 1965, 357-361.
93 S. Hakamori. A rapid permethylation of glycolipid, and polysaccharide catalyzed by methylsulfinyl carbanion in dimethylsulfoxide. *J. Biochem.,* 1964, **55**, 205-208.
94 P. Prehm. Methylation of carbohydrates by methyl trifluoromethanesulfonate in trimethyl phosphate. *Carbohydr. Res.*, 1980, 78, 372-374.
95 J. L. Minor. Chemical linkage of polysaccharides to residual lignin in loblolly pine kraft pulps. *J. Wood Chem. Techn.,* 1986, **6**, 185-201.
96 T. Koshijima, F. Yaku and R. Tanaka. Fractionation of Björkman LCC from Pinus densiflora. *Appl. Polym. Symp.,* 1976, **28**, 1025-1039.
97 J. L. Minor. Chemical linkage of pine polysaccharides to lignin. *J. Wood Chem. Techn.,* 1982, **2**, 1, 1-16.
98 T. Iversen. Lignin-carbohydrate bonds in a lignin-carbohydrate complex isolated from spruce. *Wood Science and Tech,* 1985, **19**, 3, 243-251.
99 T. Iversen and S. Wännström. Lignin-carbohydrate bonds in a residual lignin isolated from pine kraft pulp. *Holzforschung-International Journal of the Biology, Chemistry, Physics, and Technology of Wood*, 1986, **40**, 1, 19-22.
100 T. Watanabe, T. Koshijima. Structural features of 2,3-dichloro-5,6 dicyano-1,4-benzoquinone-oxidized lignin acetate by carbon-13 nuclear magnetic resonance spectroscopy. *Mokuzai Gakkaishi*, 1989, **35**, 130-134.
101 M. Y. Balakshin, E.A. Capanema, in Proceedings of the 13th ISWFPC, Auckland, New Zealand, 2005, **2**, 53-360.
102 U. Westermark and K. Gustafsson. Molecular size distribution of wood polymers in birch kraft pulps. *Holzforschung*, 1994, **48**(Suppl.), 146-150.
103 M. Lawoko, and A.R. van Heiningen. Fractionation and characterization of completely dissolved ball milled hardwood. *Journal of Wood Chem. and Techn.,* 2011, **31**, 3, 183-203.
104 A. Rudin, V. Grinshpun, K.F. O'Driscoll. Long chain branching in polyethelene *J.Liquid Chrom. & Related Technol.*, 1984, **7**, 9, 1809–1821.
105 B. Zhang, G. Q. Fu, Y.S. Niu, F. Peng, C. L. Yao, and R.C. Sun. Variations of lignin–lignin and lignin–carbohydrate linkages from young Neosinocalamus affinis bamboo culms. *RSC Advances,* 2016, **6**, 19, 15478-15484.
106 S. N. Njamela, P. F. Vena, M. Brienzo, J. Görgens, and L. Tyhoda. The effect of hemicelluloses pre-extraction on the lignin-carbohydrate complex structure of sugarcane bagasse pulp. Tappsa Journal, 2013, **6,** 13-20.
107 Y. Liu, W. Chen, Q. Xia, B. Guo, Q. Wang, S. Liu, Y. Liu, J. Li, and H. Yu. Efficient Cleavage of Lignin–Carbohydrate Complexes and Ultrafast Extraction of Lignin Oligomers from Wood Biomass by Microwave-Assisted Treatment with Deep Eutectic Solvent. *ChemSusChem*, 2017, **10**, 8, 1692-1700.
108 L. Yao, H. Yang, J. Xu, N. Zhang, Y. Chen, and Y. Xie. Elucidation of the structure of lignin and lignin carbohydrate complex of Gramineae. *Cell. Chem. Technol*, 2015, **49**, 3-4, 259-266.
109 L. Yao, C. Chen, X. Zheng, Z. Pheng, H. Yang, and Y. Xie. Determination of Lignin-Carbohydrate Complexes Structure of Wheat Straw using Carbon-13 Isotope as a Tracer. *BioResources,* 2016, **11**, 3, 6692-6707.
110 R. Singh, S. Singh, K. Trimukhe, K. Pandare, K. Bastawade, D. Gokhale and A. Varma. Lignin-Carbohydrate complexes from sugarcane bagasse: Preparation, purification and characterization. *Carbohydrate Polym.,* 2005, **62**, 57– 66.
111 D. V. Evtuguin, B. J. Goodfellow, C. Pascoal Neto, and N. Terashima, Characterization of lignin-carbohydrate linkages in Eucalyptus globulus by 2D/3D NMR spectroscopy using specific carbon-13 labeling technique. Proceedings of the 13th ISWFPC, Auckland, New Zealand. 2005, **2**, 439–444.
112 Y. Miyagawa, T. Mizukami, H. Kamitakahara and T. Takano. Synthesis and fundamental HSQC NMR data of monolignol β-glycosides, dihydromonolignol β-glycosides and p-hydroxybenzaldehyde derivative β-glycosides for the analysis of phenyl glycoside type lignin-carbohydrate complexes (LCCs). *Holzforschung*, 2014, **68**, 7, 747–760.
113 F. Nylander, H. Sunner, L. Olsson, P. Christakopoulos and G. Westman. Synthesis and enzymatic hydrolysis of a diaryl benzyl ester model of a lignin-carbohydrate complex (LCC). *Holzforschung,* 2016, **70**, 5, 385–391.
114 H. Nishimura, A. Kamiya, T. Nagata, M. Katahira and T. Watanabe. Direct evidence for α ether linkage between lignin and carbohydrates in wood cell walls. *Scientific Report*, 2018, **8**, 6538.
115 H. Kim, J. Ralph, and T. Akiyama. Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d6. *BioEnergy Res.*, 2008, **1**, 56–66.
116 L. Zhang and G. Gellerstedt. Quantitative 2D HSQC NMR determination of polymer structures by selecting suitable internal standard references. *Magn. Reson. Chem.*, 2007, **45**, 37–45.
117 J. Rencoret, G. Marques, A. Gutierrez, L. Nieto, J. I. Santos, J. Jimenez-Barbero, A. T. Martinez and J. C. del Rio. HSQC-NMR analysis of lignin in woody (Eucalyptus globulus and Picea abies) and non-woody (Agave sisalana) ball-milled plant materials at the gel state. *Holzforschung*, 2009, **63**, 691–698.
118 E. A. Capanema, M. Y Balakshin, and J. F. Kadla. Quantitative characterization of a hardwood milled wood lignin by nuclear magnetic resonance spectroscopy. *J Agric Food Chem,* 2005, **53**, 9639–9649.
119 *A*. Teleman, M. Nordstrom, M. Tenkanen, A. Jacobs and O. Dahlman. Isolation and characterization of O-acetylated glucomannans from aspen and birch wood. *Carbohydr. Res.,* 2003, **338**, 6, 525–534.
120 A. Teleman, J. Lundqvist, F. Tjerneld, H. Stålbrand, and O. Dahlman. Characterization of acetylated 4-O-methylglucuronoxylan isolated from aspen employing 1H and 13C NMR spectroscopy. *Carbohydr. Res.*, 2000, **329**, 807–815.
121 G. Brunow, J. Sipilä and T. Mäkela. On the formation on non-cyclic benzyl ethers during lignin biosynthesis. *Holzforschung*, 1989, **43**, 55–59.
122 A. Granata and D. S. Argyropoulos. 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.,* 1995, **43**, 1538−1544.
123 D.J. Peterson, and N.M. Loening. QQ-HSQC: a quick, quantitative heteronuclear correlation experiment for NMR spectroscopy. *Magnetic Resonance in Chemistry,* 2007, **45**, 11, 937-941.
124 S. Heikkinen, M. M. Toikka, P. T. Karhunen, and I. A. Kilpeläinen. Quantitative 2D HSQC (Q-HSQC) via suppression of J-dependence of polarization transfer in NMR spectroscopy: application to wood lignin. *JACS*, 2003, **125**, 14, 4362-4367.
125 H. Koskela, I. Kilpeläinen, and S. Heikkinen. Some aspects of quantitative 2D NMR. *Journal of Magnetic Resonance,* 2005, 174, 2, 237-244.
126 J. L. Wen, S. L. Sun, B. L. Xue, and R. C. Sun. Recent advances

This journal is © The Royal Society of Chemistry 20xx

in characterization of lignin polymer by solution-state nuclear magnetic resonance (NMR) methodology. *Materials*, 2013, **6**, 1, 359-391.

127 K. Hu, W. M. Westler, and J. L. Markley. Simultaneous quantification and identification of individual chemicals in metabolite mixtures by two-dimensional extrapolated time-zero 1H− 13C HSQC (HSQC0). *JACS*, 2011, 133, 6, 1662-1665

128 M. Sette, H. Lange, and C. Crestini. Quantitative HSQC analyses of lignin: a practical comparison. *Computational and structural biotechnology journal,* 2013, 6, 7, e201303016.

129 C. Crestini, F. Melone, M. Sette, and R. Saladino. Milled wood lignin: a linear oligomer. *Biomacromolecules,* 2011, **12**, 11, 3928-3935.

130 L., T. M., Sirkka, L. Maunu, B. Hortling, M. Toikka, and I. Kilpeläinen. Analysis of technical lignins by two-and three-dimensional NMR spectroscopy. *Journal of agricultural and food chemistry*, 2003, **51**, 8, 2136-2143.

131 J. Ralph and F. Lu. Cryoprobe 3D NMR of acetylated ball-milled pine cell walls. *Organic & biomolecular chemistry*, 2004, **2**, 19, 2714-2715.

132 J. Ralph and L. L. Landucci. 5: NMR of Lignins. In: Lignin and lignans: advances in chemistry, 2011.

133 J. Ralph, J. M. Marita, S. A. Ralph, R. D. Hatfield, F. Lu, R. M. Ede, J. Peng, and L. L. Landucci. Solution state NMR of lignins. Advances in lignocellulosics characterization. Atlanta, Ga.: TAPPI Press, 1999, 55-108.

134 L. Zhang, G. Gellerstedt, J. Ralph and F. Lu. NMR studies on the occurrence of spirodienone structures in lignins. *J. of wood chem. and techn.,* 2006, **26**, 1, 65-79.

135 K. Pirkko, P. Rummakko, J. Sipilä, G. Brunow and I. Kilpeläinen. Dibenzodioxocins; a novel type of linkage in softwood lignins. *Tetrahedron Letters 36*, 1995, **1**, 169-170.

136 J. Ralph, R.D. Hatfield, S. Quideau, R.F. Helm, J.H. Grabber, H.-J.G. Jung. Pathway of p-Coumaric Acid Incorporation into Maize Lignin As Revealed by NMR. *J. Am. Chern. Soc*., 1994, **116**, 9448-9456.

137 H. Grabber, J. Ralph, R.D. Hatfield. Cross-Linking of Maize Walls by Ferulate Dimerization and Incorporation into Lignin. *J. Agric. Food Chern.,* 2002, **50**, 6008-6016.

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

139 M. Toikka, G. Brunow. Lignin–carbohydrate model compounds. Reactivity of methyl 3-O-(α-l-arabinofuranosyl)-β-d-xylopyranoside and methyl β-d-xylopyranoside towards a β-O-4-quinone methide. *J. Chem. Soc. Perkin Trans 1*, 1999, 1877–1883

140 T. Tokimatsu, T. Umezawa, M. Shimada. Synthesis of four diastereomeric lignin-carbohydrate complexes (LCC) model compounds composed of a β-O-4 lignin model linked to methyl β-D-glucoside. *Holzforschung*, 1996, **50**, 156–160

141 J. Sipilä and K. Lundquist. On the Acid-catalysed Alkylation of Lignins. *Holzforschung*, 1991, **45**, 3–7.

142 N. Terashima, S. A. Ralph, L.L. Landucci. New Facile Syntheses of Monolignols Glucosides; p- Glucocoumaryl Alcohol, Coniferin, and Syringin. *Holzforschung*, 1996, **50**, 151-155.

143 K. Li and R. F. Helm. Synthesis and Rearrangement Reactions of Ester-Linked Lignin-Carbohydrate Model Compounds. *J. Agric. Food Chem.*, 1995, **43**, 2098-2103.

144 W. Boerjan, J. Ralph, and Marie Baucher. Lignin biosynthesis. *Annual review of plant biology*, **2003**, 54, 1, 519-546.

145 H. H. Felle. pH: signal and messenger in plant cells. *Plant Biol.,* 2001, **3**, 577-591

146 S. C. Fry. Polysaccharide-modifying enzymes in the plant cell wall. *Ann. Rev. Plant Phys. Plant Mol. Biol.,* 2005, **46**, 497–520.

147 T. J. Simmons, K.E. Mohler, C. Holland, F. Goubet, L. Frankova, D.R. Houston, A.D. Hudson, F. Meulewaeter and S. C. Fry. Hetero-trans-β-glucanase, an enzyme unique to Equisetum plants, functionalizes cellulose. *Plant J.,* 2015, **83**, 753–769.

148 N. Giummarella, C. Gioia and Martin Lawoko. A one-pot biomimetic synthesis of selectively functionalized lignins from monomers: a green functionalization platform. *Green Chemistry*, 2018, **20**, 11, 2651-2662.

149 R. D. Hatfield, D. M. Rancour and J. M. Marita. Grass cell walls: a story of cross-linking. *Frontiers in plant science*, 2017, **7**, 2056.

150 T. Ishii. Structure and functions of feruloylated polysaccharides. *Plant Sci*., 1997, **127**, 111–127.

151 G. Wende and S.C. Fry. 2-O-β-D-Xylopyranosyl-(5-O-feruloyl)-L-arabinose, a widespread component of grass cell walls. *Phytochemistry*, 1997, **44**, 1019−1030.

152 L. Saulnier, J. Vigouroux and J. F. Thibault. Isolation and partial characterization of feruloylated oligosaccharides from maize bran. *Carbohydr. Res.* 1995, 272, 241−253.

153 J. Ralph and R. F. Helm. Lignin/hydroxycinnamic acid/polysaccharide complexes: Synthetic models for regiochemical characterization. in Forage cell wall structure and digestibility, 1993, 201-246.

154 R. D. Hatfield, R. F. Helm, and J. Ralph. Synthesis of methyl 5-O-trans-feruloyl-α-L-arabinofuranoside and its use as a substrate to assess feruloyl esterase activity. *Analytical biochemistry 194*, 1991, **1**, 25-33.

155 R. D. Hatfield, J. Ralph, and J. H. Grabber. Cell wall cross-linking by ferulates and diferulates in grasses. *J. of the Science of Food and Agriculture*, 1999, **79**, 3, 403-407.

156 J. Ralph, S. Quideau, J. H. Grabber and R. D Hatfield. Identification and synthesis of new ferulic acid dehydrodimers present in grass cell walls. *J Chem Soc Perkin Trans 1*. 1994, 3485–3498.

157 J. M. Harkin. *Lignin – a natural polymeric product of phenol oxidation, in Oxidative Coupling of Phenols*. Ed by Taylor WI and Battersby AR, Marcel Dekker, New York, 1967, 243–321.

158 J. Ralph, R.F. Helm, S. Quideau and R. D. Hatfield. Lignin – feruloyl ester cross-links in grasses. Part 1. Incorporation of feruloyl esters into coniferyl alcohol dehydrogenation polymers. *J Chem Soc Perkin Trans 1*, 1992, 2961–2969.

159 J. Ralph, M.T. Garcia-Conesa and G. Williamson. Simple preparation of 8-5-coupled diferulate. *J Agric Food Chem*, 1998, **46**, 2531–2532.

160 J. Ralph, J. H. Grabber, and R. D. Hatfield. Lignin-ferulate cross-links in grasses: active incorporation of ferulate polysaccharide esters into ryegrass lignins. *Carbohydrate research 275*, 1995, **1**, 167-178.

161 C. Lapierre, A. Voxeur, S. D. Karlen, R. F. Helm, and J. Ralph. Evaluation of Feruloylated and p-Coumaroylated Arabinosyl Units in Grass Arabinoxylans by Acidolysis in Dioxane/Methanol. *Journal of agricultural and food chemistry,* 2018, 6**6**, 21,5418-5424

162 R. Chazal, P. Robert, S. Durand, M. F. Devaux, L. Saulnier, C. Lapierre and F. Guillon. Investigating lignin key features in maize lignocelluloses using infrared spectroscopy. *Appl. Spectrosc.*, 2014, **68**, 1342−1347

163 M. Y. Balakshin, D.V. Evtuguin, C. Pascoal Neto, A.M.S. Silva, P. M. Domingues, F. M. L. Amado. Studies on lignin and lignin-carbohydrate complex by application of advanced

spectroscopic techniques. In: Proceedings of the 11th International Symposium on Wood Pulping Chemistry, Nice, France, 2001, 1, 103-110.
164 M. Y. Balakshin, E. A. Capanema, C. L. Chen and H. S. Gracz. Elucidation of the structures of residual and dissolved pine kraft lignins using an HMQC NMR technique. *Journal of agricultural and food chemistry*, 2003, **51**, 21, 6116-6127.
165 C. S. Lancefield, H. L. Wienk, R. Boelens, B. M. Weckhuysen, and P. C. Bruijnincx. Identification of a diagnostic structural motif reveals a new reaction intermediate and condensation pathway in kraft lignin formation. *Chemical science*, 2018, **9**, 30, 6348-6360.
166 T. Wang, Y. B. Park, D. J. Cosgrove, and M. Hong. Cellulose-Pectin Spatial Contacts Are Inherent to Never-Dried Arabidopsis thaliana Primary Cell Walls: Evidence from Solid-State NMR. *Plant Physiol.*, 2015, 168, 871-884.
167 T. J. Simmons, J. C. Mortimer, O. D. Bernardinelli, A. C. Pöppler, S. P. Brown, R. Dupree, and P. Dupree. (2016). Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nature communications*, 2016, **7**, 13902.
168 X. Kang, A. Kirui, M. C. Dickwella Widanage, F. Mentink-Vigier, D. J. Cosgrove and T. Wang. Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nature communication,* 2019, **10**, 347.

This journal is © The Royal Society of Chemistry 20xx