

Special feature: tutorial

Received: 4 October 2013

Revised: 31 October 2014

Accepted: 3 November 2014

Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI 10.1002/jms.3541

A critique on the structural analysis of lignins and application of novel tandem mass spectrometric strategies to determine lignin sequencing

Joseph Banoub,^{a,b*} Guo-Hua Delmas Jr.,^c Nicolas Joly,^d Grahame Mackenzie,^e Nadja Cachet,^c Bouchra Benjelloun-Mlayah^c and Michel Delmas^{c,f*}



This review is devoted to the application of MS using soft ionization methods with a special emphasis on electrospray ionization, atmospheric pressure photoionization and matrix-assisted laser desorption/ionization MS and tandem MS (MS/MS) for the elucidation of the chemical structure of native and modified lignins. We describe and critically evaluate how these soft ionization methods have contributed to the present-day knowledge of the structure of lignins.

Herein, we will introduce new nomenclature concerning the chemical state of lignins, namely, virgin released lignins (VRLs) and processed modified lignins (PML). VRLs are obtained by liberation of lignins through degradation of vegetable matter by either chemical hydrolysis and/or enzymatic hydrolysis. PMLs are produced by subjecting the VRL to a series of further chemical transformations and purifications that are likely to alter their original chemical structures.

We are proposing that native lignin polymers, present in the lignocellulosic biomass, are not made of macromolecules linked to cellulose fibres as has been frequently reported. Instead, we propose that the lignins are composed of vast series of linear related oligomers, having different lengths that are covalently linked in a criss-cross pattern to cellulose and hemicellulose fibres forming the network of vegetal matter. Consequently, structural elucidation of VRLs, which presumably have not been purified and processed by any other type of additional chemical treatment and purification, may reflect the structure of the native lignin.

In this review, we present an introduction to a MS/MS top-down concept of lignin sequencing and how this technique may be used to address the challenge of characterizing the structure of VRLs.

Finally, we offer the case that although lignins have been reported to have very high or high molecular weights, they might not exist on the basis that such polymers have never been identified by the mild ionizing techniques used in modern MS. Copyright © 2015 John Wiley & Sons, Ltd.

Additional supporting information may be found in the online version of this article at the publisher's web site.

Keywords: lignins; GC-MS; APPI; ESI; MALDI-MS

Introduction

The plant cell wall (CW) of all vascular plants and woods is composed of cellulose, hemicelluloses and lignins, which are formed in large quantities by means of solar energy. The CW structure is composed of primary and secondary CW parts. Both parts are composed of cellulose microfibrils (9–25%) and an interpenetrating matrix of hemicelluloses (25–50%), pectins (10–35%) and proteins (10%).^[1] Cellulose is the main constituent of all skeletal constituents in plants^[2] and forms the framework of the CWs, whereas hemicelluloses cross-link non-cellulosic and cellulosic polymers.^[3,4] Cellulose is composed of approximately 8×10^3 D-glucopyranose residues linked by β -D-(1 → 4)-glycosidic bonds.^[5] Hydrogen bonds hold about 40 of these glycan chains together to form a cellulose microfibril. The cellulose microfibril arrangements in the primary wall are random. Cellulose microfibrils are linked to hemicellulosic polysaccharides that are mainly composed of xylans, mannos, galactans or combinations thereof, which are attached to the cellulose fibres by hydrogen and covalent bonding.^[3,5]

In general, secondary walls are derived from the primary walls by thickening and inclusion of lignin into the CW matrices. The

* Correspondence to: Joseph Banoub, Department of Chemistry, Memorial University of Newfoundland, St John's, Newfoundland, A1C 5X1, Canada. E-mail: banoubj@fldo-mpo.gc.ca

* Correspondence to: Michel Delmas, Compagnie Industrielle de la Matière Végétale (CIMV), 109 Rue Jean Bart – Diapason A, 31670 Labège, France. E-mail: m.delmas@cimv.fr

a Department of Chemistry, Memorial University of Newfoundland, St John's, Newfoundland, A1C 5X1, Canada

b Science Branch, Special Projects, Fisheries and Oceans Canada, St John's, NL, A1C 5X1, Canada

c Compagnie Industrielle de la Matière Végétale (CIMV), 109 Rue Jean Bart – Diapason A, 31670, Labège, France

d Centre Technologique des Agroressources, IUT de Béthune, Université d'Artois, 1230 rue de l'Université, BP 819, 62408, Béthune Cedex, France

e Department of Chemistry, University of Hull, Hull HU6 7RX, UK

f Inp-Ensiacet, Laboratoire de Génie Chimique, Université de Toulouse, 4 allée Emile Monso, 31432, Toulouse Cedex 4, France

secondary plant CW contains hemicelluloses (10–40%) that are embedded in lignin (25%). Historically, this type of arrangement has been compared with steel rods embedded in concrete to form prestressed concrete.^[6] Cellulose and hemicelluloses appear to be more structurally organized in the secondary CW than in that of the primary CW. Lignin occurs widely in the middle lamellae and secondary CWs of higher plants and plays a key role in constructive tissues as a building material, giving it its strength and rigidity and resistance to environmental factors.^[1] The diverse and complex nature of lignin monomers and hemicellulosic moieties in lignohemicellulosic complexes make stereotypical conceptualization of secondary wall structures for all plants extremely difficult.^[3,5]

History of lignin

Lignin was discovered by Anselme Payen (1839) and was first described as the encrusting material in wood. Researchers were puzzled by the nature of this very abundant material.^[7] Although lignin was found to have higher carbon content than carbohydrates, the chemical nature remained obscure for a long time. Accordingly, for almost the last century, the structure of lignin was described as a complex polymer composed of irregular branched units. After cellulose, lignin is the second most abundant biopolymer on Earth. It is found in all vascular plants, mostly between the cells, as well as within the cells and in the CWs. The term lignin is derived from the Latin word for wood: *lignum*. The polyphenolic structure of lignin is well known for its role in woody biomass to give resistance to biological and chemical degradation. This is due to their hydrophobic nature and insolubility in aqueous systems preventing access of degrading chemicals and organisms.

It is well accepted in our era that lignin is a very promising massive renewable organic resource. The worldwide annual production of lignin, as a side product of wood processing industries, exceeds 50 million tons.^[8]

Lignins in the CWs of vascular plants are intimately mixed with the carbohydrate components and form a complex glycolignin network. In recent years, it was shown that the chemical structure of the various lignin biomolecules depends on the botanical origin and chemical composition of the vegetal fibres.^[9] It has also been proposed that the several different chemical, enzymatic and mechanical extraction methods are accountable for the major structural divergence that occurs after extraction and isolation. As a result, it appears that the only logic step to determine the natural structure of lignin is to isolate it from the vegetal matrix, without causing any structural change.^[10] Therefore, preparing pure samples of unchanged lignin is not an easy endeavour with the consequence that structural determination of lignin is perhaps more challenging than with other biopolymers.^[11–13]

Proposed structure of lignin

During the last century, it was proposed that lignin existed as a cross-linked amorphous macromolecular material composed of phenylpropanoid monomers known as lignols. Lignin is relatively hydrophobic and aromatic in nature and consists of several types of substructures, which possess the same phenylpropanoid skeleton, but differ in the degree of oxygen substitution on the phenyl ring. The monolignol H-structure consists of a 4-hydroxyphenyl ring. The monolignol guaiacyl G-structure contains one hydroxyl and one methoxyl group, whereas the monolignol syringyl S-

structure contains two methoxyl groups and one hydroxyl group (Fig. 1).^[8]

The H, G and S constituent units are linked covalently, forming ether, ester and carbon–carbon bonds, which repeat in an apparently random manner and provide great complexity.^[14] In general, the site of attachment of the covalent linkages can vary between two lignols. It is generally accepted that lignins are composed of H, G and S units attached by β -O-4', β -5', β - β' , β -1', 5-5' and 5-O-4' linkages that are relatively resistant towards degradation. These are designated according to the atomic centres in the radicals that are coupled together during the final step of lignin biosynthesis.^[15] The equivalent substructures in the presumed overall structure of the biopolymer consist of alkyl-aryl ethers, phenylcoumarans, resins, dibenzodioxocins, biphenyls, tetrahydrofuran-spiro-cyclohexadienones and diaryl ethers (Fig. 2).^[16]

In softwood lignin, the G structure is dominant. Hardwood lignin normally contains a mixture of S and G lignols, with S being in the majority. The H lignols predominate in lignin found in grasses.^[17]

It has been shown that during lignin extraction, the β -O-4' dilignol linkage was the most commonly and easily cleaved bond, whereas other dilignol linkages were more chemically recalcitrant. In addition, it is believed that the relative proportion of monomers dictates the relative abundance of interunit linkages present. For example, lignins rich in G units contain more recalcitrant β -5', β - β' and β -O-4' linkages, whereas lignins enriched with S units are less cross-linked and less recalcitrant to extraction. For this reason, lignin composition is classically described by the relative abundance and ratio of H, G and S units.^[18–20] As a result, our knowledge of lignin chemical structure is less precise than our knowledge of other natural and synthetic polymers.

The typical molecular mass of isolated lignin has been estimated to be in the range 1000–20 000 u.^[8] There have been many reports in the literature suggesting that the degree of polymerization in natural lignin is difficult to measure because it is assumed that it invariably fragments during extraction.^[8]

It was speculated that lignin may exist as one single molecule in its native environment and that its structure may never cease to grow.^[21] As a result, the chemical structure of lignin was described as a highly cross-linked and three-dimensional biopolymer (Fig. 3).^[22]

In contrast, it was proposed that isolated lignin has cross-linking frequencies of less than 1 in 19 monomer units.^[23] In addition, it was anticipated that isolated lignin could exist as lamella-like

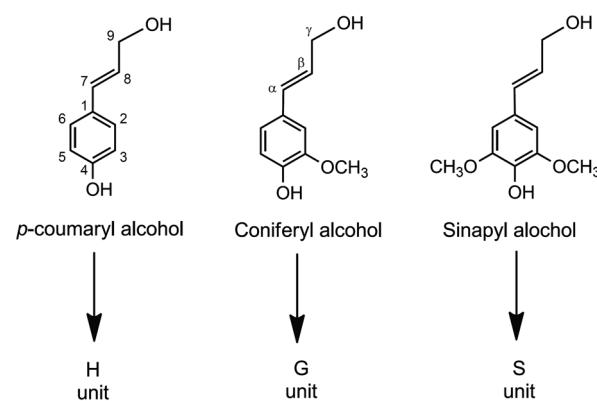


Figure 1. Monomeric structures of lignin. Reproduced with permission from W.O.S. Doherty et al. (2011). Copyright 2011 by Elsevier Ltd. All rights reserved.

Chemical structure of native and modified lignins

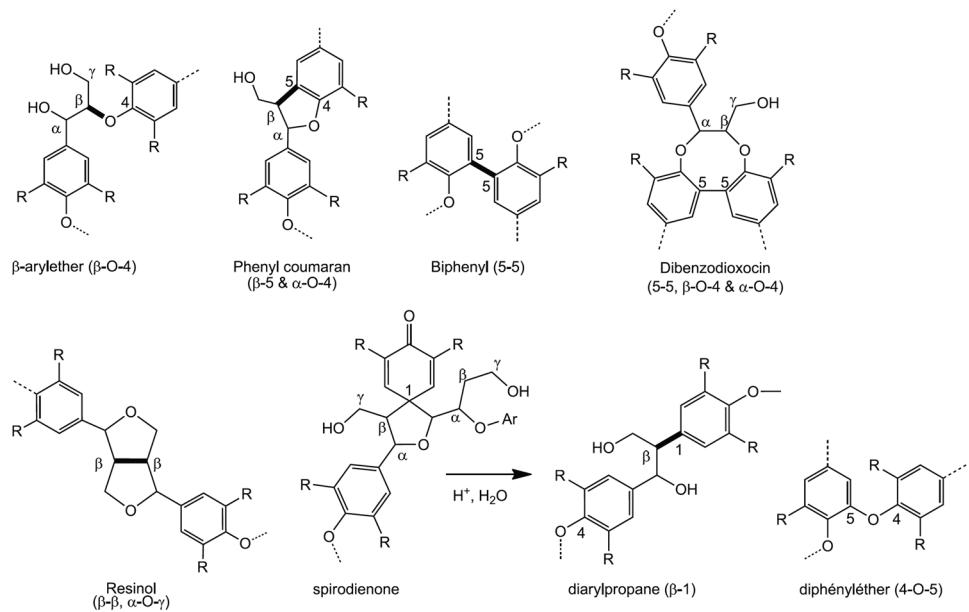


Figure 2. Main linkages in native lignins ($R=H$ in *p*-hydroxyphenyl units or OCH_3 in guaiacyl and syringyl units). Reproduced with permission from C. Lapierre (2010). Copyright 2010 by Taylor and Francis Group, LLC. All rights reserved.

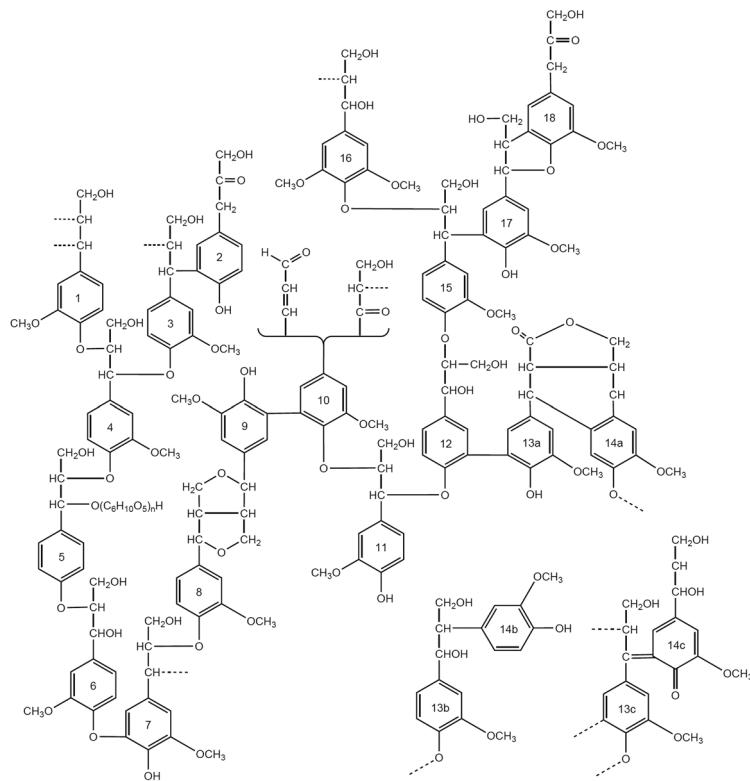


Figure 3. Proposed structural model of spruce lignin by K. Freudenberg. Reproduced with permission from L. B. Davin *et al.* (2005). Copyright 2005 by Elsevier Ltd. All rights reserved.

sheets, which lacked cross-linking.^[24,25] Yet, it was also suggested that the *in situ* molecular structure of lignin in the species of Japanese cedar (*Cryptomeria japonica*), camellia (*Camellia japonica*) and ginkgo (*Ginkgo biloba*) behaved much as if it was made up of linear macromolecules.^[26] These newest views have been reported by many researchers (Fig. 4).^[11,27–29]

Lignans and neolignans

The lignans encompass a newly classified group of natural plant products that are related biochemically to phenylalanine metabolism and composed of two phenylpropanoid units.^[30] They are characterized by the two phenylpropanoid units (C-9) being linked to each other *via* the C-8 to C-8' bonds (refer to Fig. 5 for

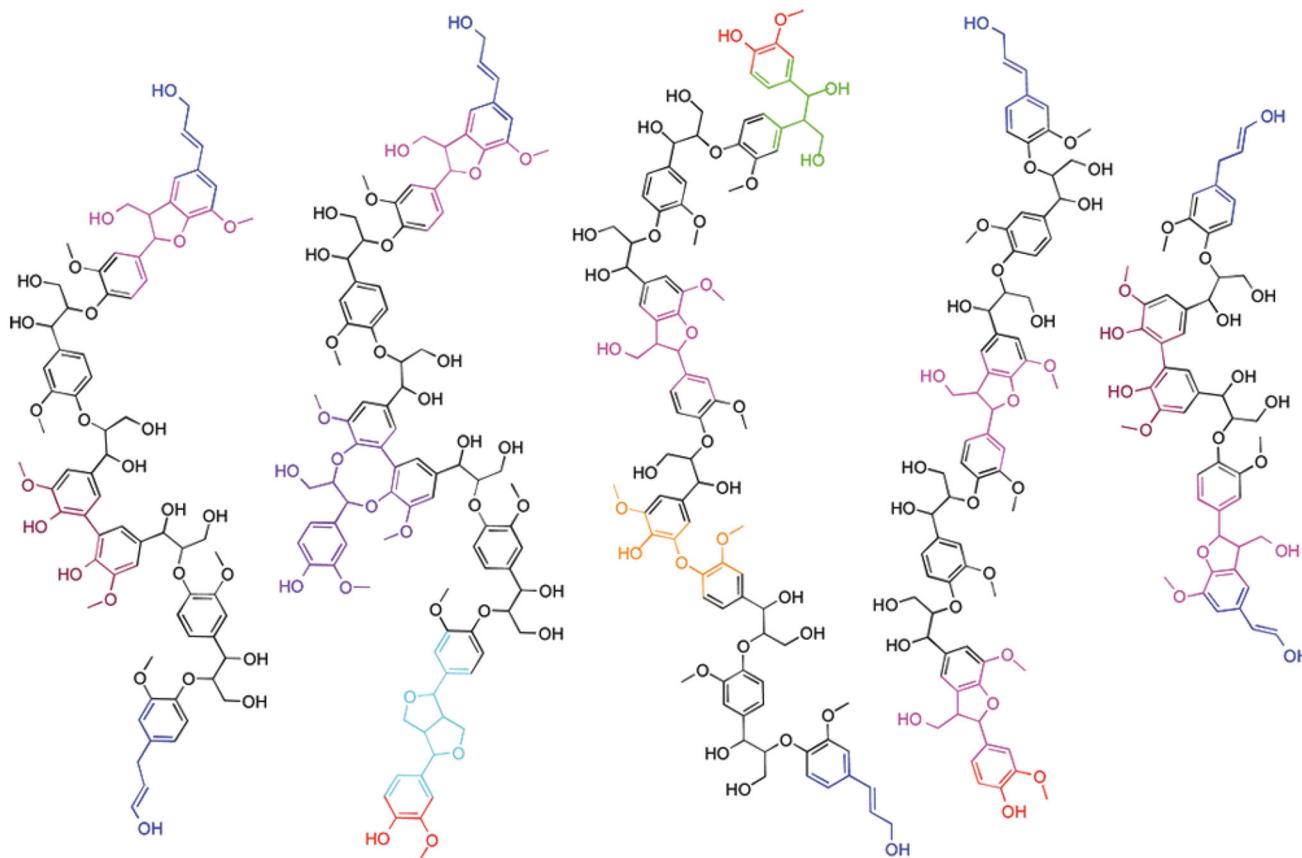


Figure 4. Schematic representation of softwood milled wood lignin. The structures reported are indicative of the occurrence of each interunit bonding and do not strictly reflect the statistical frequency. β -O-4' is labelled in black, β -5' pink, DBDO purple, 5'-5' amaranth, β - β' azure, β -O-5' orange, terminal aliphatic chains blue, and terminal phenolic OH red. Reproduced with permission from C. Crestini et al. (2011). Copyright 2011 American Chemical Society. All rights reserved.

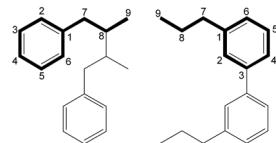


Figure 5. Basic chemical structure of lignans and neolignans. Reproduced with permission from G.P. Moss (2000). Copyright 2000 by IUPAC International. All rights reserved.

numbering).^[31] According to IUPAC nomenclature, in the absence of the C-8 to C-8' bonds between the two phenylpropanoid units, the dimers formed are known as neolignans.^[31]

As a result of their wide distribution in the plant kingdom, lignans have attracted a great interest for their pharmacological properties. They exhibit anti-platelet, antiviral, anti-tumour, anti-depressant and pesticidal activities (Fig. 5).^[32–36]

Within the neolignans, three subgroups are distinguished: the oxyneolignans, the sesquineolignans and the dineolignans. The oxyneolignans have their C-9 units linked via an ether bond. The sesquineolignans are composed of three C-9 units and the dineolignans of four C-9 units.^[31] Lignans have been identified in many plants,^[37] whereas neolignans are thought to be less abundant and mainly present in coniferous trees.^[38] Lignans have been found almost exclusively in edible plant sources.^[39,40]

The unsolvable paradigms of lignin

Two types of lignin analyses are recognized. Usually, non-destructive analytical methods employing topochemical exploration are used to assess the presence and distribution of lignin in the plants.^[41] In contrast, chemical structural analysis of the lignin polymer is mostly performed by destructive analysis methods. In these methods, the isolated lignin is depolymerized to produce small fragments that provide partial structural information of the original native structure.^[42]

More than a century of lignin research has produced an enormous amount of experimental results that were, to a large extent, dissimilar and difficult to reproduce. The probable reason for these irregularities between the several experimental results may well be due to the different extraction methods used in attempts to isolate an identical lignin product each time from the same plant.

From the chemical point of view, it is well recognized that the isolation of lignin in its unaltered form is highly unlikely because of the relatively harsh extraction conditions used. However, there is still much debate on whether any lignin extract adequately represents the native lignin structure. Therefore, the harsh extraction condition required to release lignin from lignocellulosic cellular material results in the degradation of the lignin polymeric structure itself. For example, acid and alkaline depolymerization of lignin causes cleavage of ester bonds and some ether bonds; however, the reactivity of the released fragments may lead to more complex rearranged condensed polymeric structures as shown in Fig. 6.^[43]

Chemical structure of native and modified lignins

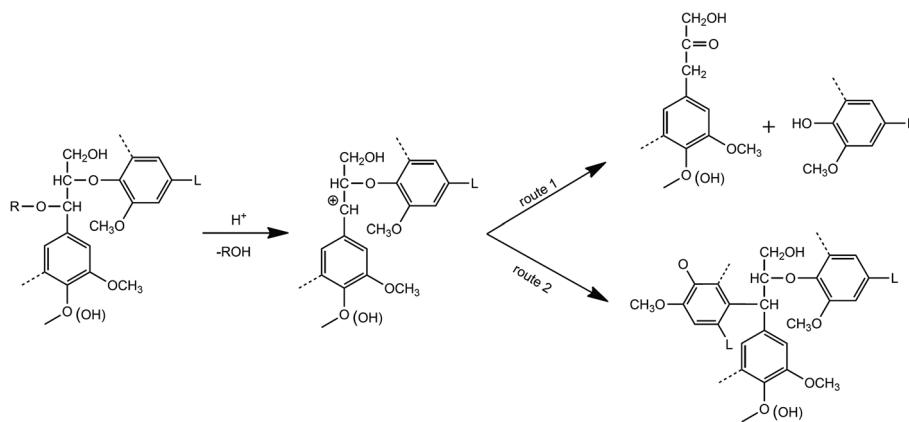


Figure 6. Acid catalysed depolymerization of lignin causing the cleavage of ester bonds and ether bonds (route 1) and reassociation to form complex rearranged condensed polymeric structures (route2). Reproduced with permission from R. J. A. Gosselink (2011). Copyright 2011 by Wageningen Universiteit, The Netherlands. All rights reserved.

Moreover, the continued use of inadequate and outdated analytical methodology to characterize both lignin and lignin-like polymers further exacerbates this situation.^[43] Needless to say, there were some opinionated claims that implied that structural analysis of lignin should be based on pure samples.^[43]

Other notable enigmas concerning the lignin biosynthesis worthy of mention are as follows. The mechanism by which the biosynthetic formed monolignols migrate to the CW is not known.^[44] Although, all steps of lignin biosynthesis have been identified, little is known about the number of genes encoding each enzymatic step.^[45]

The most established biosynthetic polymerization route known as the combinatorial process postulates that the oxidizing enzymes that produce the lignol radicals, which combine to form dilignols, trilignols and oligolignols, depend on nothing except the chemical control that dictates the coupling and by association all the resulted polymeric structures.^[14,46,47] This polymerization is initiated by the oxidative radical ionization of phenols, followed by combinatorial radical coupling. Decades of research are consistent with this model.^[48–51]

Yet, a conflict exists between this traditional biochemical polymerization of lignin,^[1,15,46,48–50,52–55] with another so far unconfirmed model that declares that the polymerization that forms lignin must be controlled by proteins (dirigent proteins).^[22] Replication of the primary lignin chain by template polymerization implies that lignins may have a regular repeating unit.^[52–54]

Nevertheless, this postulate of template polymerization was chided by the following statement: it will be astronomically improbable to find chemically ordered region in the lignins, and the search for regularity is futile.^[44] In addition, it was suggested that because lignin polymerization is combinatorial, the sequencing of lignin is meaningless, except for individual molecules.^[44] Nevertheless, understanding the propensity of how monolignols couple will afford insight into the factors that determine lignin structure.^[44]

In general, we can conclude that the structural analysis of lignin is a very tedious task, especially as lignins are heterogeneous mixture of oligomers. In reality, structural analysis permit generally to estimate the average frequency of the main units and the main bond types in the polymer, from a conglomerate of oligomers simply of the whole plant or in crudely fractionated major tissues.

Aims of this review

The present review is devoted to the application of MS to the elucidation of the chemical structure of lignin. Of importance, we have chosen pertinent examples to highlight the key role of the state-of-the-art MS methods that employ softer ionization modes to analyse the structure of native and modified types of lignin.

We provide an overview and critique of the current understanding of lignin structure taking in account of the various extraction methodologies that have been employed (refer to the supplementary data). In addition, we emphasize how these various aspects have contributed to the current knowledge of the structure of lignins.

Despite the many workers being engaged in the study of lignins over the years,^[18,56] the many aspects of the chemistry, biosynthesis and molecular biology have remained not fully elucidated.^[1,15,46,48–55] It is noteworthy that lignins contain a range of chemical functional groups, many of which may be the direct result of the extraction method. The main functional groups of lignins are hydroxyl (aromatic and aliphatic), methoxyl, carbonyl and carboxyl. Furthermore, the solubility of the lignins is affected by the relative proportions of such functional groups. In addition, most lignins are quite soluble in alkaline solution as a result of the ionization of phenolic and carboxyl functional groups.

In this review, we have introduced a new nomenclature concerning the chemical state of the lignin, namely, the virgin released lignins (VRLs) and the processed modified lignins (PML). The VRLs are obtained by liberation of lignins by denaturing vegetable matter using chemical hydrolysis and enzymatic hydrolysis methods. Furthermore, we present the case that only the structures of the VRL types represent a close reflection of the unique native structural features present in the covalently attached glycolignin network. This is explained by the fact that the VRL types are not further transformed into other PML types by any other type of additional extraction (e.g. alkali) or purification procedure. It should be understood that when we use the term lignins in plural, it is mainly because we have presumed that these oligomers are released in different sizes and molecular weights.

Recently, a new term lignome has been introduced in the literature, which is used to include lignin structures biosynthesized and those obtained following extraction.^[57] The term lignome was defined by K. Morreel *et al.* as the ensemble of all phenolics for which

the biosynthesis is co-regulated by lignin polymerization and includes the oligolignol (small lignin polymers) pool.^[57]

We wish to advise the reader that the field is too extensive for a single comprehensive review; hence, we apologize in advance for the omissions.

Main momentum for lignin research in the 21st century

The development of methods for the conversion of biomass into biofuels and other high value organic molecules is becoming increasingly important because of the efforts to find renewable alternatives to fossil fuels.^[58–60] Because lignins are mainly composed of aromatic rings, they have the potential to serve as an alternative source of aromatic materials to those derived from petroleum.^[58–60]

Recently, various controlled degradation methods were explored to convert lignins into smaller molecules, and new catalytic pathways were sought to remove oxygen excess from these molecules to increase their energy density and synthetic value.^[61] However, a major challenge that remains is the structural elucidation of these high volume molecules.

Note: Additional information to support this review

Supplementary information pertaining to the methods of extraction of lignins, for the determination of their molecular weights and for their standard degradations, is available free of charge via the Internet at <http://pubs.jms.org>.

Lignin structure determination by conventional MS methods using ionization sources under reduced vacuum

Mass spectrometry offers advantages in terms of speed, specificity and sensitivity and has been demonstrated to be a very powerful technique in the structural elucidation of lignin. Electron ionization MS and chemical ionization MS (CI-MS) have been used for the study and characterization of derivatized lignol monomer constituents, obtained by either reductive cleavage or by pyrolysis. In this case, only the monomeric and, to a small extent, dimeric products have been identified by comparison of gas chromatographic retention times and mass spectra using authentic samples.^[62–65]

In the last few decades, hardly any reviews have been published dealing with the application of MS to the structural elucidation of chemically degraded lignins. These MS methods have dealt with thermal degradation of lignins including the well-known pyrolysis method, which continues to be a method of choice for many practicing lignin chemists.^[66]

Reale *et al.* have described the use of MS, equipped with different ionization sources, for the study of lignin degradation products and the authors have also attempted to characterize the whole lignin macromolecule.^[47] The different ionization methods described in their review article cover the following methods: CI-MS; photoionization MS, single-photon ionization MS, molecular beam MS, fast-atom bombardment MS and resonance-enhanced multiphoton ionization MS. For this reason, these methods will not be discussed here, and readers are encouraged to read their article to obtain a comprehensive literature survey on these subjects.^[47]

Likewise, the recent development of new ionization techniques such as atmospheric pressure ionization, which includes electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI) and atmospheric pressure photoionization (APPI)

together with matrix-assisted laser desorption/ionization (MALDI) MS, has provided new possibilities to analyse lignin oligomers (LOs).

Atmospheric pressure ionization MS methods as powerful tools for the identification of synthetic and degraded lignins

The following section is devoted to the application of MS using soft ionization methods with a special emphasis on ESI-MS, APCI-MS and APPI-MS and tandem MS (MS/MS) for the elucidation of the chemical structure of synthetic, native (VRL) and modified lignins (PML).

The continued development of liquid chromatography (LC) MS techniques has made possible the analysis of trace amounts of analytes from complicated matrices. With LC, the analytes of interest can be separated from each other and reliably identified due to the sensitivity and specificity of MS. Thus, LC-MS has become an exceptional tool for many applications, to include the analysis of proteins, pharmaceuticals in biological fluids and toxic substances in environmental samples.^[67] ESI and APCI have thus far been the most popular interfaces for the coupling of LC with MS.^[68] ESI is selective for species that exist as ions in solution and species with acidic or basic functionalities capable of donating or stabilizing excess electrolytes (e.g. H⁺, Na⁺, Cl⁻).^[68] It has been found that ion abundances can be determined by a competition of solutes for charge at the surface of an electrospray droplet,^[69] resulting in ion abundances (peak height) that are not simply a function of the analyte concentration. This is the reason why minor constituents with high charge affinity can dominate the ion current. Such interferences are known as matrix effects and make it impossible to calibrate peak height as a function of the dissolved analyte concentration by ESI-MS.^[70] For that reason, quantitative evaluation of ESI-MS data for unknown mixtures such as dissolved organic matter in natural waters is extremely difficult.^[71] Another well-known flaw of ESI is its high susceptibility to ion suppressions, leading to data inaccuracies, elevated method detection limits, required matrix-matched calibration standards and tedious sample cleanup procedures.^[70,72]

In APCI, the liquid sample is first evaporated and then ionization of the analytes takes place with a charged plasma formed by using a corona discharge.^[70] APCI may be used to analyse less polar analytes rather than with ESI; however, completely non-polar analytes cannot be ionized by either of the two ionization methods.

A newer ionization technique, namely the APPI, has been developed, which is complementary to ESI. Definitely, APPI is a very soft photoionization process, which does not cause extensive fragmentation of covalent bonds. The ionization process in APPI is instigated by a krypton discharge lamp emitting 10-eV photons. These photons ionize compounds, which have ionization energies below 10 eV. Compounds appropriate for this technique can include larger molecules; however, gases and solvents are not ionized by APPI, which greatly minimize the background interference. Moreover, in contrast to ESI and APCI, APPI ionization of the analytes depends on the ionization energy rather than their proton affinity, which results in the ionization of molecules of low polarity.^[73]

As a result, APPI extends the analytical window for the simultaneous analysis of nonpolar and polar species.^[74] Yet, APPI has proved to be the one of the best API source for LC-MS analysis of small molecules. Moreover, APPI is much less susceptible to matrix effects than ESI^[75,76] and offers four to five orders of dynamic linear

ranges.^[3,4] In addition, APPI requires simplified sample cleanup procedures, less dilution and reanalysis for high-level analytes. Also, it produces more accurate results and lessens or prevents the need to use matrix-matched calibration standards for quantitative analysis.^[76,77] When using APPI, it was established that there was no ion suppression due to charge competition among solutes because both solvents and salts are not ionized. Accordingly, elimination of chemical noise is successful. This latter feature was shown by Hockaday *et al.* to be important in FT-ICR-MS analysis, in which the space-charge effects greatly influence the sensitivity and resolution.^[71] Finally, the APPI ionization source is virtually field-free resulting in the improvement of the efficiency of ion transport to the MS inlet.

Atmospheric pressure photoionization MS can be used to analyse a multitude of different classes of compounds, which include the following: aliphatic lipids,^[72,78] isoprenoid lipids,^[79] condensed tannins,^[80] hydrophobic peptides, polycyclic aromatic hydrocarbons,^[77,81–84] pesticides,^[85,86] mycotoxins,^[85] crude oil fractions^[83,87] and lignins.^[88–90]

APCI-MS/MS-based sequencing of oligolignols

Recently, a series of VRL oligomers (i.e. dilignols, trilignols and tetralignols) were detected in poplar (*Populus* spp.), which were extracted with methanol catalyst, and in the PML degraded products of tobacco (*Nicotiana tabacum*) xylem, which was extracted by the Bjorkman method with 1,4-dioxane : water (96 : 4) (refer to the supplementary material), using LC coupled to APCI quadrupole ion trap (QIT) MS.^[91–93] Their structures were resolved using mass spectral information followed by further authentication by chemical synthesis.

To help the systematic characterization of poplar (*Populus* spp.) lignome, Morreel and coworkers studied the gas-phase fragmentation behaviour of the major dilignol, trilignol and oligolignols containing different bonding types using APCI-MS/MS in the negative ion mode with a QIT instrument.^[94] They also used high-resolution ESI-FT-ICR-MS (– ion mode) for confirmation of the empirical formulae of identified ions. With such advances, the authors were able to devise a method that allows the sequencing of individual LOs and detection of new units and linkages. Morreel and coworkers identified 134 oligolignols present in the xylem of poplar (*Populus tremula* and *Populus tremuloides*), ranging from dimers up to hexamers, and more than half oligomers possessed unknown linkages types.

The authors described the characteristic collision-induced dissociation (CID) MS/MS gas-phase fragmentations of the precursor deprotonated molecules obtained from various LOs containing each of the main linkage units (i.e. β -O-4', β -5' and β - β') (Fig. 7).

It was observed that the ESI-MS analysis of lignin prefers the negative ion mode, which promotes the formation of an easily deprotonated molecule. Similarly, CID-MS/MS of the deprotonated molecule allows the study of the different gas-phase fragmentation routes of the precursor ions, which allows the establishing of its chemical structure. Likewise, it is generally accepted that the CID-MS/MS of the protonated molecules obeys the even-electron fragmentation rules and the dissociation that can occur *via* either heterolytic or homolytic cleavage and might involve ion-neutral complexes.^[95]

Application of MS/MS to the precursor of deprotonated phenylcoumaran anions obtained from the β -aryl ether $[G(t_8-O-4')G]$ afforded the product ion $[M-H-H_2O]^-$ which eliminated water

(18 u). This latter product ion loses formaldehyde (48 u) to afford the product ion $[M-H-H_2O-CH_2O]^-$.

Upon CID-MS/MS of the precursor of deprotonated phenylcoumaran anions derived from $[G(8-5')G]$, elimination of formaldehyde loss (30 u) afforded the product ion assigned as $[M-H-H_2O-CH_2O]^-$. This latter product ion lost a water molecule to afford the $[M-H-H_2O]^-$ product ion, which was attributed to the primary alcohol function. An additional loss of formaldehyde (loss of 30 u) was noted to generate the $[M-H-CH_2O]^-$ product ion. It should be noted that as no primary hydroxyl group is present in the case of $[G(8-8')G]$ resins, no water loss was observed, although a formaldehyde loss $[M-H-CH_2O]^-$ (–30 u) is always present.^[96]

Finally, it was shown that the CID-spectra of the precursor anions isolated from the β -aryl ethers and benzodioxanes were distinguished by loss of water (18 u) and formaldehyde (30 u) and in some cases by the consecutive elimination of both formaldehyde and water (48 u). It should be noted that the authors described the first product ion corresponding to the 48-u loss as a major peak in the spectra of β -aryl ethers, but such a peak was hardly visible in spectra of benzodioxanes. In addition, the CID-spectra of resins showed peaks corresponding to the loss of a methyl radical (–15 u) and formic acid (46 u).^[94]

Furthermore, beside the extremely complex CID-mechanisms described, it is interesting to note that the authors have used nuclear magnetic resonance (NMR) terminology in attempt to explain their MS/MS fragmentations, notably for the presence of a β -5' linkage unit as evidenced by the presence of a characteristic *m/z* triplet in the CID-spectrum, whereas any oligomers containing the β - β' linkage unit would be characterized by the resinol *m/z* quartet. In a very similar study, the same research group described fully the CID-fragmentations of all the synthetic standard structures that were employed.^[97]

Nevertheless, when the authors studied the CID-MS/MS of the deprotonated dimers, they found that the charge-remote fragmentations were more widespread than in the positive ionization mode.^[20] Consequently, the authors studied specifically the CID gas-phase fragmentation pathways of the β -O-4', β -5' and β - β' deprotonated dimers as they represented the major bonding structures encountered in the lignome.^[95–99]

It has been well documented in the literature that the charge-driven CID fragmentations are initiated from the most acidic site of the precursor ion,^[98] whereas the charge-remote reactions are actually responsible for the fragmentations, yet the former type will occur whenever possible.^[98,99] As a result, a charge-remote fragmentation is to be considered only when no charge-driven pathway is possible. Accordingly, Morrell and coworkers presumed that all of their CID-MS/MS dissociations were initiated by charge-driven mechanisms. These were started by the conversion of the phenoxide anion to a quinone methide, with the simultaneous elimination of the remaining seven-position linked portion.^[94] This last moiety was eliminated as a neutral fragment, via an anion/neutral complex-mediated proton abstraction, promoting water loss from the β -aryl ether linkage unit. As well, Morrell and coworkers proposed that phenoxide anions could also trigger a secondary CID pathway, which could be mediated by an anion/neutral complex by a charge remote fragmentation mechanism (Fig. 8).^[94]

As well, the authors studied also the MS/MS analyses of a large series of the following oligolignols, $G(t\beta-O-4')G(t\beta-O-4')G$, $G(t\beta-O-4')S(\beta-5')G$, $S(t\beta-O-4')S(\beta-5')G$, $S(t\beta-O-4')G(\beta-5')G$, $G(t\beta-O-4)S(\beta-5')G$, $G(e\beta-O-4')S(\beta-5')G$, $G(t\beta-O-4')S(\beta-\beta')G$, $G(t\beta-O-4')S(\beta-\beta')S$, $G(e\beta-O-4')S(\beta-\beta')S$ and $S(t\beta-O-4')S(\beta-\beta')S$.^[91,94]

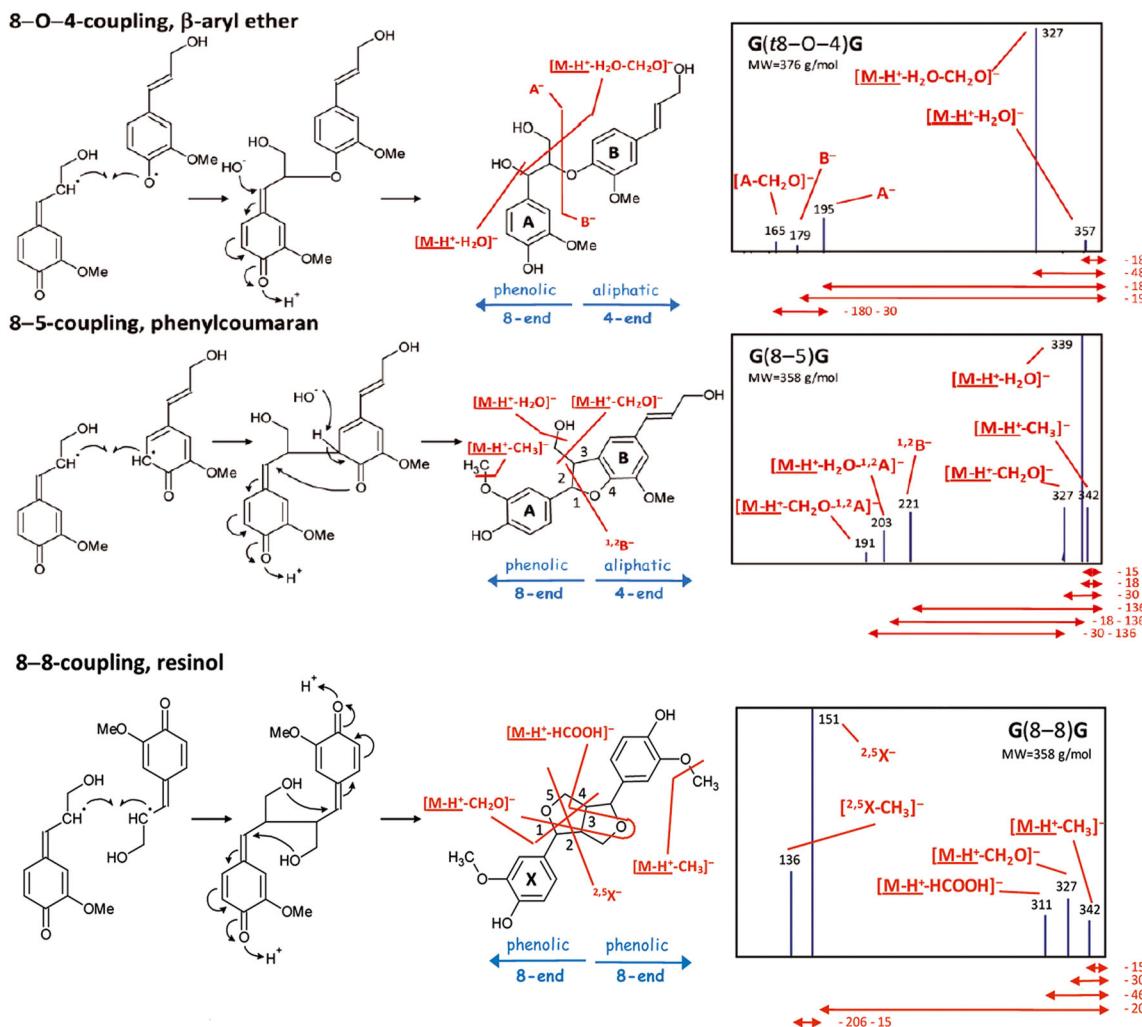
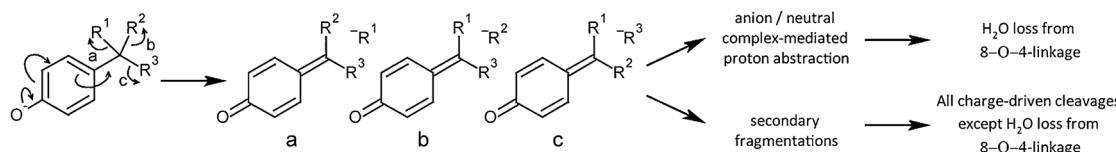


Figure 7. Radical–radical coupling during lignin polymerization. Reproduced with permission from K. Morreel et al. (2010). Copyright 2010 by American Society of Plant Biologists. All rights reserved.

Phenoxyde / quinone methide conversion



C⁻ ion formation of G(8-O-4)S(8-8)G

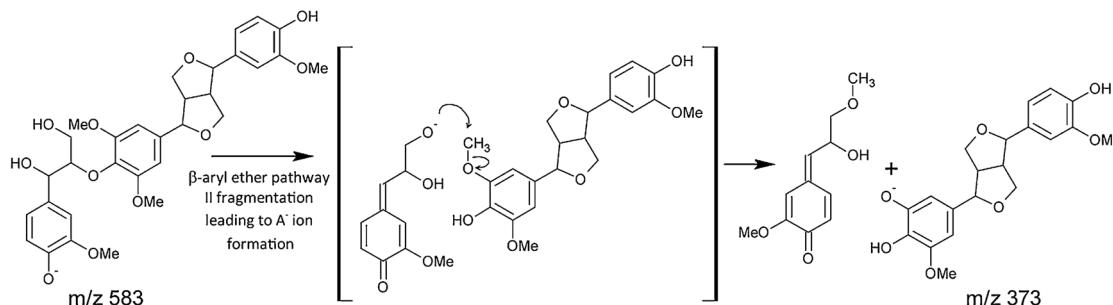


Figure 8. CID pathways. Phenoxyde anion/quinone methide conversion with charge migration that was shown to induce all charge driven fragmentation pathways of the lignin linkage types A new β -aryl ether-associated pathway II fragmentation was observed upon CID of X(β - β)X-containing trilignols, leading to the formation of the C2 ion at m/z 373. Reproduced with permission from K. Morreel et al. (2010). Copyright 2010 by American Society of Plant Biologists. All rights reserved.

Chemical structure of native and modified lignins

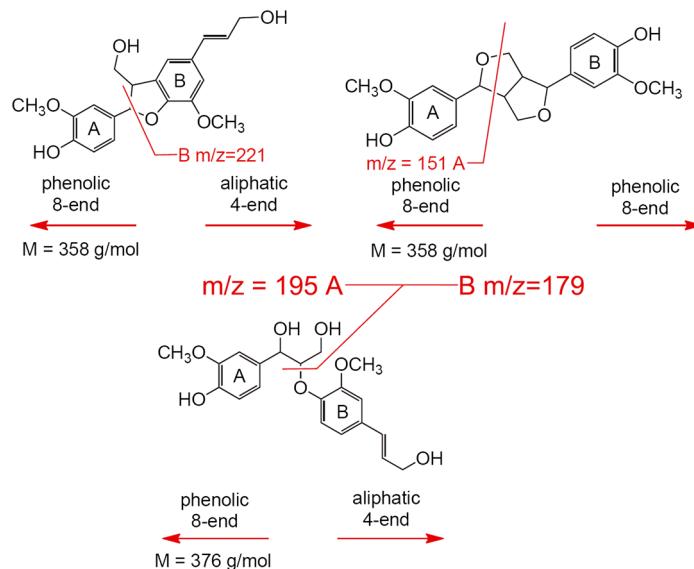


Figure 9. Dimers of G units showing the three main links. Their characteristic fragmentation patterns are indicated by red lines. (I) The β -O-4', β -aryl ether link yields fragment ions A – (*m/z* 195) and B – (*m/z* 179). (II) The β -5', phenylcoumaran link yields fragment ion B – (*m/z* 221). (III) The β - β' , resinol link yields fragment ion A – (*m/z* 151). Reproduced with permission from E. Kiyota *et al.* (2012). Copyright 2012 by American Chemical Society. All rights reserved.

It should be noted that the fragmentation schemes described were modelled on the MS/MS nomenclature fragmentation of proteins, oligonucleotides and carbohydrate polymers.^[100–105] However, it is imperative to mention that the product ions identification system will only work when the analysed lignin derivative possesses an eight-end aromatic residue and the other terminus possesses an aliphatic four-end aromatic residue. Unfortunately, this is not the case with the majority of all isolated technical lignins, which have both the eight-end and four-end terminated with aromatic groups.

As a final point, it is noteworthy to mention that these authors conjured that they were the first to use MS/MS for lignin sequencing. However, it is of note that APCI-MS, ESI-MS and APPI-MS in conjunction with CID-MS/MS for the sequencing and establishing lignin structures were reported earlier.^[9,64,88,106,107]

Ultrahigh-performance liquid chromatography ESI-MS and MS/MS analysis of sugar cane soluble lignin

Kiyota *et al.* synthesized *in vitro* the LOs starting from the individual monomers by a procedure described by de Angelis *et al.* using two different peroxidase enzymes and three different reaction times.^[65,108] The reaction conditions were tested to determine which would produce the longest chains and greatest variety of oligomers. Twenty-five compounds were identified, which were mostly dimers, trimers and one tetramer. The retention times of these series of compounds were also determined. The resulting compounds were analysed by ultrahigh-performance liquid chromatography (UPLC) ESI-MS and CID-MS/MS (negative ion mode). Their work resulted in the structures of 25 identified compounds being determined and used to build a LO data library.^[108]

The authors compared the CID-fragmentation pattern of the deprotonated isobaric dimers $G(\beta\text{-}5')G$ and $G(\beta\text{-}\beta')G$ (*m/z* 357), and $G(\beta\text{-O-}4')G$ (*m/z* 375), to data obtained from the literature.^[94]

The CID-MS/MS of the deprotonated molecules of $G(\beta\text{-}5')G$ linked dimer afforded a product ion at *m/z* 221, which is characteristic of the four-aliphatic end (B). The CID-MS/MS of the isobaric

deprotonated $G(\beta\text{-}\beta')G$ dimer afforded the product ion of *m/z* 151, which is characteristic of the eight-phenolic end (A).

Obviously, the mass of the $G(\beta\text{-O-}4')G$ linked dimer increases by 18 u because the incorporation of a molecule of water and the CID-MS/MS of the deprotonated $G(\beta\text{-O-}4')G$ molecules gives two diagnostic product ions at *m/z* 179 (B) for the four-aliphatic end and at *m/z* 195 for the eight-phenolic end (A) (Fig. 9).

In contrast, the CID-MS/MS of the S-G dimer deprotonated anions followed parallel gas-phase fragmentation comparable to the G-G dimers, except that we notice that the formed product ions had a difference of 30 u between S and G monomers, which of course was due to the presence of one additional methoxy group in the S monomer. Therefore, MS/MS of the deprotonated $S(\beta\text{-}5')G$ linked dimer afforded the product ion at *m/z* 221, which was characteristic for G residue present on the four-aliphatic end. The precursor ion scan of the $S(\beta\text{-O-}4')G$ dimer deprotonated molecule at *m/z* 405 afforded the product ion at *m/z* 225, which indicated that the S unit was on the eight-phenolic end.

It is of noteworthy to mention that the deprotonated $S(\beta\text{-}\beta')S$ at *m/z* 417 possesses an 8-8 link, whereas, the $S(\beta\text{-O-}4')S$ dimer contains 18 u more than the $S(\beta\text{-}\beta')S$ dimer due to water incorporation. Subsequently, the MS/MS of the deprotonated $S(\beta\text{-}\beta')S$ dimer molecule at *m/z* 435 afforded the product ions at *m/z* 225 and 209, which were characteristic for this structure.

Kiyota *et al.* also identified the following deprotonated molecules by comparison of their MS/MS data with those obtained from the literature: $G(\beta\text{-O-}4')G(\beta\text{-O-}4')G$ at *m/z* 571, $G(\beta\text{-O-}4')S(\beta\text{-}5')G$ at *m/z* 583, $S(\beta\text{-O-}4')G(\beta\text{-}5')G$ at *m/z* 583, $S(\beta\text{-O-}4')S(\beta\text{-}5')G$ at *m/z* 613, $G(\beta\text{-O-}4')S(\beta\text{-}\beta')S$ at *m/z* 613 and $S(\beta\text{-O-}4')S(\beta\text{-}\beta')S$ at *m/z* 643. Consequently, the CID-MS/MS spectra of the trimers afforded product ions that were characteristic of the monomers composed of the aromatic part A, whereas the BC product ions were characteristic of the remaining dimer ions, allowing the elucidation of the complete structure. The following are examples of characterization of trimers performed using MS/MS analysis (Fig. 10).^[108]

It is interesting to note that these authors chose not to use the complex CID nomenclature proposed by Morreel *et al.* described in section APCI-MS/MS-based Sequencing of Oligolignols,^[94] but

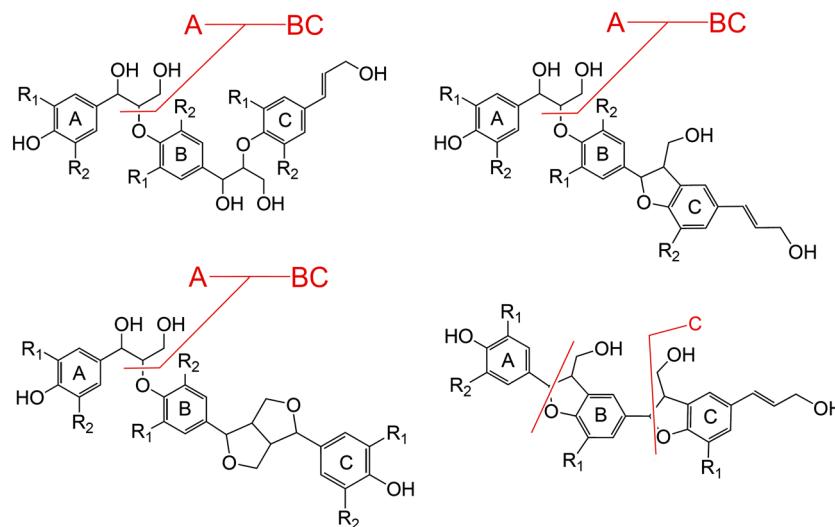


Figure 10. The general structures of the trimers identified by UPLC-MS ($R_1 = \text{CH}_3$, $R_2 = \text{H}$ or CH_3). The ($\beta\text{-O-}4$) link present in these structures results in fragment ions A monomer plus a molecule of water and the BC dimer. The deprotonated fragment ion BC^- mass has the same m/z as the dimers composed of S and G. Structure with C fragmentation can only have two $\beta\text{-O-}4$ links as there is no increase in the mass of the monomers and fragment ion C is typical of the phenylcoumaran link. Reproduced with permission from E. Kiyota et al. (2012). Copyright 2012 by American Chemical Society. All rights reserved.

rather, a very simple and comprehensible system. The work of Kiyota *et al.*^[108] illustrated the simplicity of the CID-MS/MS process.

ESI-MS/MS and APCI-MS/MS characterization of lignans isolated from sesame seeds

During acidic extraction of glucosylated lignans (sesaminol glucosides), hydrolysis occurs to release the aglycones sesamin and sesamolin, which are found in either the oil fraction and/or in the defatted fraction of sesame seeds.^[109–111] It has been demonstrated that the exact lignan composition of sesame seeds changes with the developmental stage of the seeds.^[112–117]

Struijs has shown that the extracted sesaminol glucosides from the defatted sesame meal, when purified by chromatography, yielded a series of well-resolved peaks.^[118] The identities of this series of separated compounds were confirmed by ESI-MS, APCI-MS and CID-MS/MS.^[119]

The ESI-MS (+ ion mode) of sesaminol produced only traces of the $[\text{M} + \text{H}]^+$ protonated molecule. However, when the ESI-MS was measured in the negative ion mode, it afforded appropriately the deprotonated molecule at m/z 369.1. However, in contrast to that shown by ESI-MS (+ ion mode), APCI-MS (+ ion mode) afforded the major protonated molecule $[\text{M} + \text{H}]^+$ at m/z 371.1 and the fragment ions $[\text{M} + \text{H}-\text{H}_2\text{O}]^+$ at m/z 353.1.

On the other hand, it is interesting to note that when APCI-MS was used in the negative ion mode, it afforded the abundant deprotonated molecule $[\text{M}-\text{H}]^-$ at m/z 369.1 identical to that obtained by ESI-MS. For this reason, APCI permitted the characterization of the dibenzylbutadiene lignans containing the hydroxyl groups.^[110,111] Similarly, it was shown that the CID-MS/MS fragmentations of the precursor deprotonated molecule isolated from sesaminol were similar when using either APCI-MS or ESI-MS in the negative ion mode.

Moreover, APCI-CID-MS/MS (+ ion mode) of the protonated molecules $[\text{M} + \text{H}]^+$ at m/z 371.1 gave the product ions at m/z 353.1, 335.1, 323.1, 305.1 and 135.0. On the other hand, CID-MS/MS (– ion mode) of the deprotonated molecules at m/z

369 afforded a series of product ions at m/z , 339.1, 219.1, 191.0 and 150.0.^[120] The tentative fragmentation pathways for sesaminol in the negative and positive ionization modes are presented in Fig. 11.^[105,121,122]

Finally, by using the same methodology, it was also established that the protonated sesamolin contained the methylenedioxy-bridged furanofuran structure and lacked phenolic hydroxyl groups.

APCI-MS/MS and ESI-MS/MS characterization of PML and model lignin compounds

Haupert *et al.* was able to identify the most suitable atmospheric pressure ionization method(s) for the tandem mass spectrometric characterization of complex mixtures related to PML degradation products. They used a set of lignin-degraded products as model compounds, which were subjected to positive and negative ion mode APCI and ESI under traditional as well as novel conditions using LQIT-MS.^[123] They found that when using APCI-MS (+ ion mode) only guaiacol and eugenol formed an abundance of protonated molecules.

However, in contrast, it was found that vanillyl alcohol and coniferyl alcohol formed exclusively the $[\text{M} + \text{H}-\text{H}_2\text{O}]^+$ fragment ions. The loss of water from these protonated molecules was facilitated by virtue of their presence near to allylic and benzylic hydroxyl locations. This was also rationalized by the fact that after protonation, the precursor protonated molecule loses water to generate the resonance-stabilized carbocations (Table 1).

Obviously, guaiacol and eugenol do not contain these functionalities; however, the dimer, guaiacylglycerol- β -guaiacylether, contains a benzylic hydroxyl group. Consequently, this dimer easily loses water to form the $[\text{M} + \text{H}-\text{H}_2\text{O}]^+$ ion. Nevertheless, when using APCI-MS, the two most abundant fragment ions are $[\text{M} + \text{H}-2\text{H}_2\text{O}]^+$ and/or $[\text{M} + \text{H}-\text{H}_2\text{O}-\text{CH}_2\text{O}]^+$ from formaldehyde (Table 1).

When the model compounds were studied by traditional negative mode APCI ionization (methanol and water as solvents),

Chemical structure of native and modified lignins

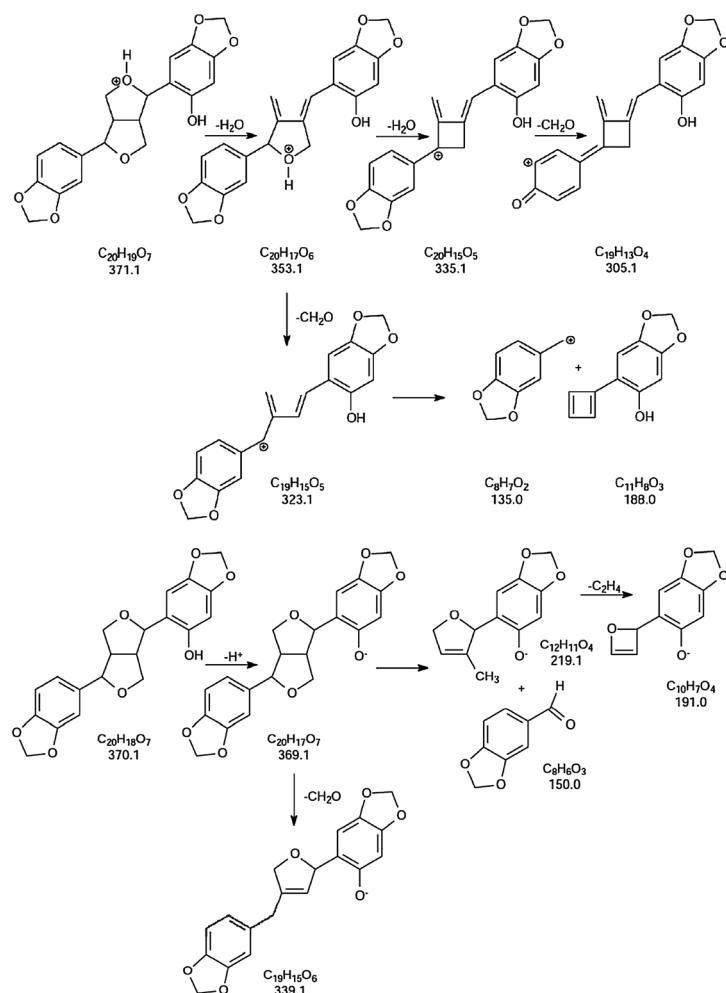


Figure 11. Proposed fragmentation of sesaminol in the negative ionization mode and in the positive mode. Reproduced with permission from K. Struijs (2008). Copyright 2008 by Wageningen Universiteit, The Netherlands. All rights reserved.

they gave ions with poor abundances. The lack in sensitivity and the weak observed fragmentation make this a poor ionization method for the characterization of lignin degradation products.^[123]

Using the model compounds, Haupert *et al.* have shown that it was possible to use ESI-MS (+ion mode) doped with sodium chloride, to generate abundant $[M + Na]^+$ sodiated molecules without any fragmentation. This approach allowed the direct determination of the molecular weights of the lignin degradation products in mixtures.^[123] However, as expected, no structural information could be obtained by examining the CID-MS/MS of the sodiated molecules using either triple quadrupole or QIT tandem mass spectrometric instruments because the sodiated molecules are known not to fragment during MS/MS analysis.

When Haupert *et al.* used ESI-MS (–ion mode) with the dopant sodium hydroxide, they obtained abundant deprotonated molecules $[M - H]^-$ without any fragmentation. The ESI-MS (–ion mode) of a synthetic mixture is showed in Fig. 12.

The CID-MSⁿ analysis of this series of precursor deprotonated molecules established the connectivities between precursor and product ions; these were obtained by measuring multiple consecutive ion isolation and CID-MSⁿ (up to MS⁷). This methodology has since been found to improve significantly the information that can be obtained by mass spectrometric analysis for lignin degradation products and lignome (Table 2).^[123]

Characterization of VRL degradation products by HPLC-ESI-MSⁿ

Owen *et al.* developed a high-performance liquid chromatography/multiple-stage MS/MS method with a commercial linear QIT/Fourier-transform ion cyclotron resonance mass spectrometer (HPLC-ESI-LQIT-FT-ICR-MSⁿ) as a tool for the analysis of complex mixtures obtained from PML degradation products.^[124] In their study, the HPLC separation method of the lignin degradation products was followed by ESI-MS analysis doped with NaOH (–ion mode). This method was established by Hauper *et al.* to be the most appropriate for the identification of lignin-related model compounds.^[123]

Owen *et al.* showed that for the analysis of artificial and real mixtures, data-dependent scans allow the instrument to select automatically the most abundant three ions from the ion source.^[124] This allowed the authors to acquire separate MS acquisitions simultaneously for the same ions with two different mass analysers. Thus, the LQIT-FT-ICR spectrometer operated in the higher duty-cycle LQIT when performing tandem mass spectral acquisitions for precursor ion selection, whereas the lower duty-cycle FT-ICR carried out the high-resolution measurements for elemental composition determination for the same ions.^[124] Accordingly, the most abundant product ion formed in the MS² experiments, was subjected to a further stage of ion isolation and MS³ fragmentation.^[124]

Table 1. Symbols, retention times, characteristic mass fragments and limits of detection for CuO oxidation products. Reproduced with permission from K. Kaiser et al. (2012). Copyright 2012 by American Chemical Society. All rights reserved

Compound (MW)	APCI (+) (<i>m/z</i>) Relative abundance	APCI (–) (<i>m/z</i>) Relative abundance	ESI (+) (<i>m/z</i>) Relative abundance	ESI (–) (<i>m/z</i>) Relative abundance
Guaiacol (124)	[M + H] ⁺ (125) 100%	Poor signal [M-H] ⁻ (123) 100%	[M + H] ⁺ (125) 100%	Poor signal [M-H] ⁻ (123) 100%
Vanillyl alcohol (154)	[M + H-H ₂ O] ⁺ (137) 100%	Poor signal [M-H] ⁻ (153) 100%	[M + H-H ₂ O] ⁺ (137) 100%	[M-H] ⁻ (153) 100%
Eugenol (164)	[M + H] ⁺ (125) 100%	Poor signal [M-H] ⁻ (163) 100%	No Signal	Poor signal [M-H] ⁻ (163) 100%
Coniferyl alcohol (180)	[M + H-H ₂ O] ⁺ (163) 100%	Poor signal [M-H] ⁻ (148) 15%	[M + H-H ₂ O] ⁺ (163) 100%	[M-H] ⁻ (163) 100%
Guaiacylglycerol- β -guaiacylether (320)	[M + H-H ₂ O] ⁺ (303) 40% [M + H-2H ₂ O] ⁺ (285) 95% [M + H-H ₂ O-CH ₂ O] ⁺ (273) 97% <i>m/z</i> 285-C ₆ H ₅ OCH ₃ (<i>m/z</i> 179) 20% <i>m/z</i> 285-HOC ₆ H ₂ OCH ₃ (<i>m/z</i> 163) 25% <i>m/z</i> 285-HOC ₆ H ₄ OCH ₃ (<i>m/z</i> 161) 20% <i>m/z</i> 233-HOC ₆ H ₄ OCH ₃ (<i>m/z</i> 149) 30% <i>m/z</i> 285-HOC ₆ H ₃ (OCH ₃)CCH (<i>m/z</i> 137) 16%	[2M + Na] ⁺ (663) 15% [M + Na] ⁺ (343) 100%	Poor signal [M-H] ⁻ (319) 47%	[2M - H] ⁻ (639) 27% [M-H] ⁻ (319) 100%

Chemical structure of native and modified lignins

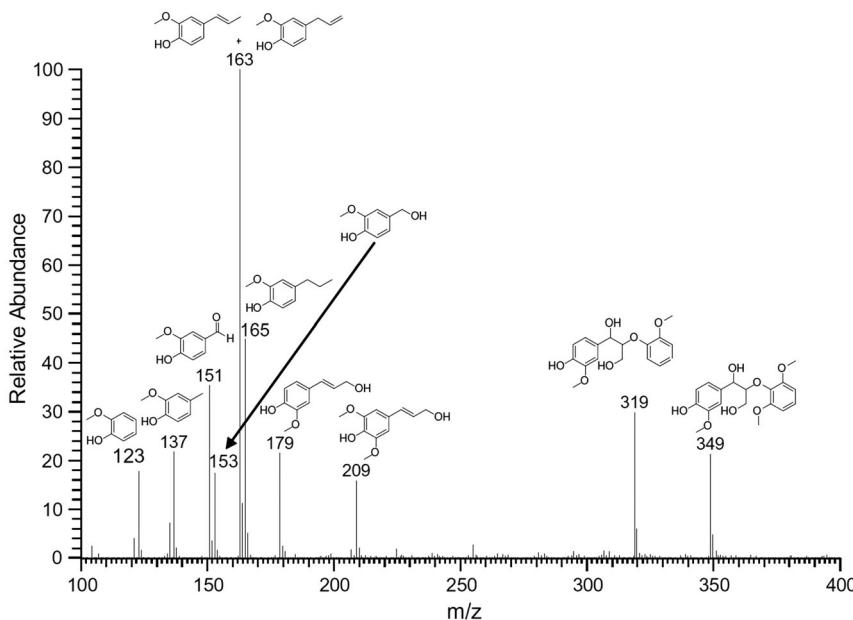


Figure 12. Negative mode ESI mass spectrum obtained for an equimolar mixture of guaiacol (MW 124 u), 2-methoxy-4-methylphenol (MW = 138 u), vanillin (MW = 152 u), vanillyl alcohol (MW = 154 u), eugenol (MW = 164 u), isoeugenol (MW = 164 u), 2-methoxy-4-propylphenol (MW = 166 u), coniferyl alcohol (MW = 180 u), sinapyl alcohol (MW = 210 u), guaiacylglycerol- β -guaiacylether (MW = 320 u) and guaiacylglycerol- β -syringylether (MW = 350 u) dissolved in 50/50 (v/v) methanol/water and doped with sodium hydroxide. Reproduced with permission from L. J. Haupert *et al.* (2012). Copyright 2012 by Elsevier. All rights reserved.

The authors used degradation products, pure standard solutions as well as a mixture of 12 model compounds of lignin degradation products (Table 3), which were separated by reversed-phase HPLC using an acetonitrile and water gradient elution followed by high-resolution MS³ analysis. The authors used three different HPLC columns to separate a mixture of 12 lignin-related model compounds. Using water and acetonitrile gradient buffered with ammonium formate, full baseline separation was attained for all components of the mixture using both the Zorbax SB-C18 and Zorbax SB-Phenyl columns. It was noted that a poorer separation was attained using Kinetex PFP column.^[124]

All of the model compounds studied exclusively by ESI-MS (– ion mode) gave the abundant deprotonated ion [M-H]⁻. Furthermore, the authors have shown that MS/MS analysis of all model compounds containing methoxyl groups attached to the aromatic ring were initiated *via* a homolytic cleavage, by the loss of a methyl radical. Additionally, all ions containing allylic or benzylic hydroxyl groups (vanillyl alcohol, coniferyl alcohol, and sinapyl alcohol) fragmented primarily by loss of a water radical.^[124] In addition, the authors indicated that the deprotonated guaiacylglycerol- β -guaiacyl ether was shown to mainly lose water followed by loss of formaldehyde.^[96]

To explore whether additional structural information could be obtained, the authors performed CID-MS³ experiments on the most abundant product ion formed in the MS² experiment. For some ions that lack a weakly bound group, no further fragmentation were observed. However, most of the secondary product ions fragmented to yield structurally informative product ions in the MS² experiments, demonstrating that multiple-stage MS/MS is useful in the structural characterization of unknown lignin degradation products.

Owen *et al.* showed that the deprotonated sinapyl alcohol containing two methoxyl groups could lose sequentially one methyl radical in the MS² experiment followed by another methyl radical in the MS³ experiment. These losses permitted the counting of

the methoxyl substituents in the analyte. In the MS² experiment of the deprotonated 2-methoxy-4-propylphenol, the product ion formed by loss of the methyl radical was further subjected to MS³. Very interestingly, this product ion fragmented by eliminating an ethyl radical by homolytic cleavage of the benzylic bond, to reveal the presence of a propyl substituent.^[124] An analogous benzylic bond cleavage was also observed for deprotonated eugenol. The MS² of deprotonated eugenol afforded a product ion by loss of the methyl radical. MS³ of this product ion was subjected to the consecutive cleavage of the benzylic carbon–hydrogen bond and loss of a hydrogen atom. In contrast, MS² of the deprotonated isoeugenol lost a methyl radical to afford a product ion that, when subjected to an MS³ experiment, did not fragment any further because it lacked a weakly bound substituent. This difference in these two MS³ fragmentations allowed the differentiation between these structural isomers.^[124]

In addition, the authors successfully separated by HPLC, the diastereomeric pair (RS/SR and RR/SS) of guaiacylglycerol- β -syringyl ethers. These were shown to yield the same CID-fragmentation product ions from the deprotonated diastereomeric pairs.

To assess the ability of this method to identify unknown components, a VRL organosolv oak lignin extract was analysed by HPLC-ESI-MS. Though the organosolv lignin sample is a very complex mixture, the HPLC was able to separate many analytes, including isobaric and isomeric molecules.^[124]

ESI-MS study of PML Kraft black liquor

Pinto *et al.* attempted to characterize the lignin structural changes during Kraft pulping (PML), using ¹H-NMR and ¹³C-NMR techniques and to estimate the changes in the lignin molecular weight using ESI-MS and gel permeation chromatography (GPC).^[125,126] The authors used four Kraft cooking processes that were designated as representative of the different delignification phases as follows: initial cooking (19.0% delignification), initial-to-bulk transition cooking

Table 2. CID-MSⁿ analysis for some of the negative ions listed in Fig 25. Reproduced with permission from L.J. Haupert et al. (2012). Copyright 2012 by Elsevier. All rights reserved

Compound (m/z of $[M - H]^-$)	MS ² fragmentations (product ions' m/z) relative abundance	MS ³ fragmentations (product ions' m/z) relative abundance	MS ⁴ fragmentations (product ions' m/z) relative abundance	MS ⁵ fragmentations (product ions' m/z) relative abundance	MS ⁶ fragmentations (product ions' m/z) relative abundance	MS ⁷ fragmentations (product ions' m/z) relative abundance
Vanillyl alcohol (153) 	153 – CH ₃ (138) 52% 153 – H ₂ O (135) 100%	138 – H ₂ O (120) 100% 135 – CH ₃ (120) 100%	120 + H ₂ O (138) 20% 120 – CO (92) 100% 120 – CO (92) 100%	No further fragmentation observed		
Eugenol (163) 	163 – CH ₃ (148) 100%	148 – H (147) 100%	No further fragmentation observed			
Isoeugenol (163) 	163 – CH ₃ (148) 100%		No further fragmentation observed			
Sinapyl alcohol (209) 	209 – CH ₃ (194) 100% 209 – H ₂ O (191) 35% 209 – H ₂ O-CH ₃ (176) 12%	194 – CH ₃ (179) 100% 191 – CH ₃ (176) 100% 176 – CH ₃ (161) 100%	179 – CO (151) 100% 176 – CH ₃ (161) 100% 161 – CO (133) 100% 161-2CO (105) 75%	151 – H ₂ O (133) 100% 151 – CH ₂ O (121) 100% 161 – CO (133) 100% 133 – CO (105) 100% 133 – 2CO (77) 19%	121 – CO (93) 100% 133 – CO (105) 100% 133 – 2CO (77) 47% 105 – CO (77) 100%	105 – CO (77) 100%

Table 3. Twelve model compounds of lignin degradation products. Reproduced with the permission of B.C. Owen *et al.* (2012) Copyright 2012 by American Chemical Society. All rights reserved

Model compound (m/z of $[M-H]^+$; determined elemental composition)	Measured exact m/z (error in $\pm m\text{Th}$ from the expected exact mass)
[A] Guaiacol (123; $C_7H_7O_2$)	123.0458 (1.2)
[B] 2-Methoxy-4-methylphenol (137; $C_8H_9O_2$)	137.0615 (1.3)
[C] Vanillin (151; $C_8H_7O_3$)	151.0408 (1.3)
[D] Vanillyl alcohol (153; $C_8H_9O_3$)	153.0564 (1.2)
[E] Eugenol (163; $C_{10}H_{11}O_2$)	163.0771 (1.2)
[F] Isoeugenol (163; $C_{10}H_{11}O_2$)	163.0771 (1.2)
[G] 2-Methoxy-4-propylphenol (165; $C_{10}H_{13}O_2$)	165.0927 (1.2)
[H] Coniferyl alcohol (179; $C_{10}H_{11}O_3$)	179.0717 (0.9)
[I] Sinapyl alcohol (209; $C_{11}H_{13}O_4$)	209.0819 (0.5)
[J] Guaiacylglycerol- β -guaiacylether (319; $C_{17}H_{19}O_6$)	319.1186 (0.4)
[K + L] Guaiacylglycerol- β -syringylether stereoisomers (349; $C_{18}H_{21}O_7$)	349.1292 (0.5)

(40.0% delignification), bulk cooking (76.0% delignification) and residual phases cooking (94.4% delignification). The dioxane lignin isolated from *E. globulus* (referred as EDL) represented the lignin in the starting wood material (0% delignification).^[125,126]

Accordingly, Pinto *et al.* reported the ESI-MS spectrum of Kraft black liquor lignin (BL) obtained from the residual phase of delignification and showed the presence of both lower and higher m/z ions, which were attributed to lignin oligomeric substructure ions containing fragments bounded by oligosaccharides.^[126]

To estimate the molecular weights of components in the cooked Kraft BL, the obtained residue was purified by selective precipitation in dioxane to yield two fractions: BLp (about 70% yield, sugars content 1.5%) and BLs (about 30% yield, sugars content 18.9%). The BLp fraction showed an ESI-MS containing lower ions around m/z 800, whereas the BLs fraction showed an MS with a higher mass distribution at m/z 1100. Such a difference is probably due to the BL fraction containing lignin residues attached to oligosaccharides (Fig. 13).

The expansion of the MS of BL showed series of fragment ions corresponding to LOs having m/z differences of 202–206 u, which is the approximate molecular mass of the phenylpropane unit (C-9) (Fig. 14). In addition, Pinto *et al.* observed that the additional ion at m/z 339 was assigned respectively to a disaccharide composed of 4-O-methyl-D-glucuronic acid and D-xylose residues (GlcpA → Xylp). This disaccharide was previously detected in the composition of *E. globulus* xylan.^[125]

Nonetheless, it is noteworthy that the authors actually deduced a low mass distribution value for such a complex ESI-MS. The absence of higher masses was a consequence of the presence of the lignin degradation products formed by this complex mixture of LOs.

Elucidation of the PML *Eucalyptus globulus Labill* lignins by ESI-MS

Evtuguin *et al.* studied the chemical structure of PML isolated from *Eucalyptus globulus Labill*.^[106,127]

The authors found that the *Eucalyptus globulus* lignins fraction was mainly composed of S/G lignols with an extremely high proportion of syringyl (S) units (82–86 mol %) and a minor proportion of p-hydrophenyl propane (H) units (roughly 2–3 mol %). They also described the presence of unknown C-6 substituted and β -O-5' type syringyl substructures representing about 65% of lignin condensed structures. ESI-MS analysis revealed a wide molecular weight distribution within the lignin with a mass distribution around 2500 u.^[106]

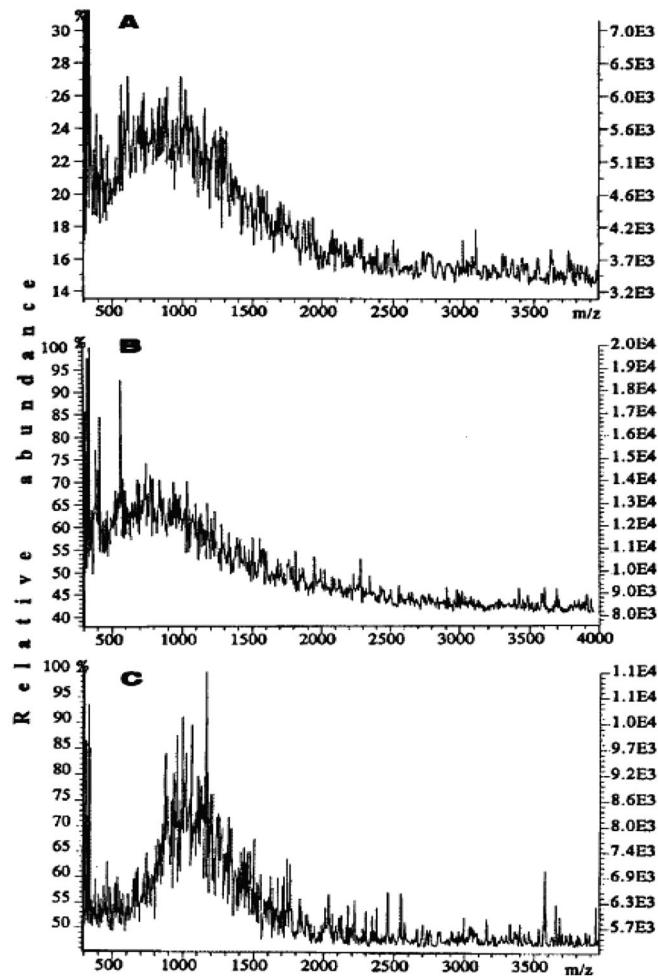


Figure 13. Negative mode ESI-MS spectrum of BL (A), BLp (B) and BLs (C) isolated from black liquor after 94.4% delignification of eucalypt wood. Reproduced with the permission of P.C. Pinto *et al.* (2002). Copyright 2002 by Marcel Dekker, Inc. All rights reserved.

In a different study, Evtuguin and Amado attempted to study the structures of PML *E. globulus* lignins and its low molecular weight fraction (EDLF), by ESI-MS and CID-MS/MS using a quadrupole time-of-flight (QTOF) MS instrument.^[107] In this study each of the lignin components were isolated by means of GPC as described previously by the same group.^[106] In addition, for the qualitative

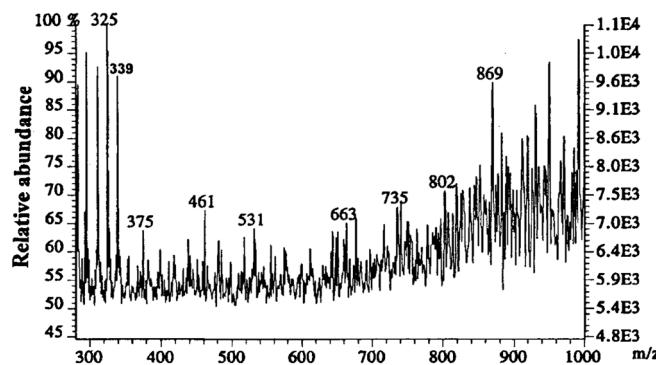


Figure 14. Negative mode ESI-MS spectrum (expanded mass range from m/z 300 to 1000) of BLs isolated from black liquor after 94.4% delignification of eucalypt wood. Reproduced with the permission of P.C. Pinto et al. (2002). Copyright 2002 by Marcel Dekker, Inc. All rights reserved.

structural elucidation of their obtained ions, the authors synthesized standards models composed of lignin dimeric model of β -arylglycerol, pinoresinol, syringaresinol and dehydrodiconiferyl alcohol according to established procedures.^[128]

The ESI-MS of the EDLF fraction showed a complex mixture of ions, which were assigned as the deprotonated LOs $[M-H]^-$ over a mass range at m/z 150–1500 (Fig. 15). The group of signals at m/z 150–210 represents monomeric degradation products obtained during acidolytic isolation of lignins. In addition, the authors were able to elucidate the CID-MS/MS of the major deprotonated molecules at m/z 319.4, 349.4, 357.4, 417.4, 643.1 and 838.8. (Fig. 15) They interpreted the CID-fragmentation patterns, which helped to elucidate the structures of these deprotonated molecules, which were assigned and compared with dimeric model compounds. Finally, they concluded that their ESI-MS data revealed a significant abundance of LOs composed of linear fragments of β -O-4' linked syringyl and guaiacyl units as well as syringaresinol.^[107]

Because the average molecular weight of the monomeric guaiacyl phenyl propane unit (G) is 190 u and that of the syringyl phenyl propane unit (S) is 220 u, the series of signals centred around m/z 500–1400, 640, 870, 1070 and 1300 were assigned as lignin dimer, trimer, tetramer, pentamer and hexamer units,

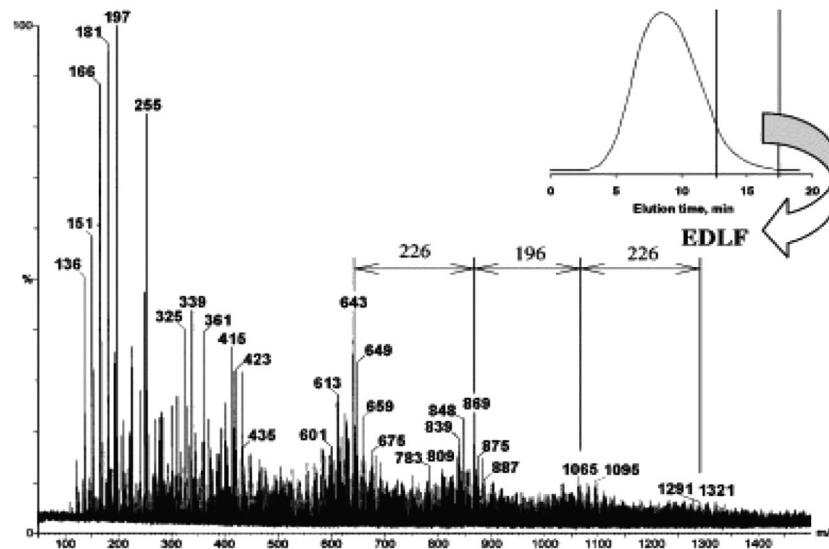


Figure 15. The ESI-MS of EDLF. Reproduced with permission from D.V. Evtuguin et al. (2003). Copyright 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. All rights reserved.

respectively.^[107] Another dominant ion series in the same ESI-MS was assigned as trimer at m/z 643, tetramer at m/z 868, pentamer at m/z 1065, and hexamer at m/z 1291. These ions appear to differ from each other by 226 and 196 u, which can be assigned to S and G units, respectively, in the composition of oligomers.^[107]

The structure of the most abundant LO was the ion at m/z 643, which was assigned as LO643 and was studied in more detail. The comparison between the CID-MS/MS analysis of LO643 at m/z 643 with that of a synthetic dimeric lignin model precursor ion at m/z 417 showed common fragmentation product ions (Fig. 16).^[107] So, CID-MS/MS of the precursor ion LO643 at m/z 643 afforded the diagnostic product ion at m/z 417 by consecutive elimination of 226 u by the loss of the terminal β -O-4'-linked syringyl propane unit containing the benzylic hydroxyl group. Also, it was suggested that the mechanism of this CID-fragmentation was driven by the homolytic splitting of the alkyl-O-aryl bond, to afford the product ions at m/z 417 and 225. As a result, Evtuguin and Amado

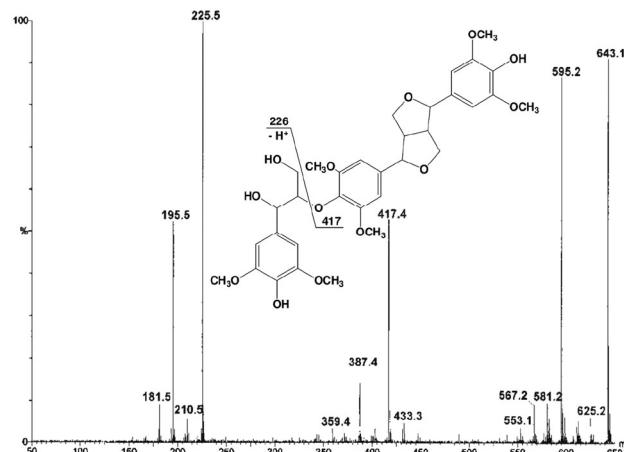
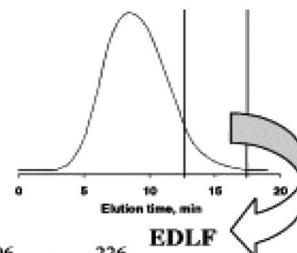


Figure 16. The product ion scan of the precursor ion extracted from the tetramer represented at m/z 643. Reproduced with permission from D.V. Evtuguin et al. (2003). Copyright 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. All rights reserved.



Chemical structure of native and modified lignins

concluded that the structure of LO643 was a trimer β -O-4-linked S unit with syringaresinol [S-(β -O-4')-S-(β - β')-S] (Fig. 16).^[107]

The same approach was used to elucidate the ion at m/z 839, which was attributed to mixture of isobaric tetramer composed of G-(β -O-4')-S-(β -O-4')-S-(β - β')-S and S-(β -O-4')-S-(β -O-4')-S-(β - β')-G. The CID-MS/MS of the precursor deprotonated molecule at m/z 839 showed the presence of two product ions at m/z 643 and m/z 613, and the authors concluded that such an assignment confirmed both structures assigned to this isobaric ion at m/z 839 (Fig. 17).^[107]

Similarly, CID-MS/MS of the ion at m/z 1065 was tentatively assigned to two alternative structures: S-(β -O-4')-S-(β -O-4')-S-

(β - β')-S-(β -O-4')-G and G-(β -O-4')-S-(β -O-4')-S-(β - β')-S.

In addition, the authors have shown that CID-MS/MS analysis of the ions obtained from lignin dimers (particularly 8-O-4', 8-8' and 8-5' linked), formed diagnostic product ions from which the specific gas-phase fragmentation patterns can be deduced.

In conclusion, the authors proposed that CID-MS/MS fragmentation of the precursor ion of LOs occurred predominantly by the cleavage of ether linkages. They also proposed that such fundamental information about lignin primary structure would form a sound basis to enable new structure elucidation studies to be made more reliably and rapidly by employing ESI-MS.

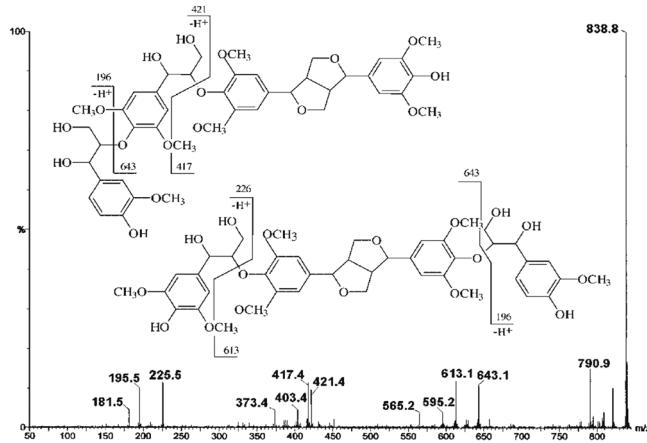


Figure 17. The product ion scan of the precursor ion extracted from the tetramer represented at m/z 839. Reproduced with permission from D.V. Evtuguin *et al.* (2003). Copyright 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. All rights reserved.

ESI-FT-ICR-MSⁿ structural analysis of PML thioacidolysis degradation lignin products

Onnerud *et al.* investigated the structures of the pinoresinol thioacidolysis degradation products by ESI-FT-ICR-MSⁿ from lignin model compounds and PML wood lignin.^[129]

The authors used the thioacidolysis ($\text{BF}_3/\text{C}_2\text{H}_5\text{SH}$) treatment for the degradation of pinoresinol and PML, which resulted in the conversion of uncondensed monomeric units in lignin into simple diastereomeric mixtures of 1,2,3-trithioethane phenylpropanoid monomers.^[130]

The authors mentioned that two major peaks were detected for thioacidolysed MWL, and these were identical when compared with the thioacidolysed pinoresinol and acetylated pinoresinol. Thioacidolysis is known to produce a large fraction of the monomeric C-6-C-3' adduct containing three thioethyl groups in the side chain as the major component.^[130]

The authors used ESI in both positive and negative ionization modes for the identification of the obtained products. In

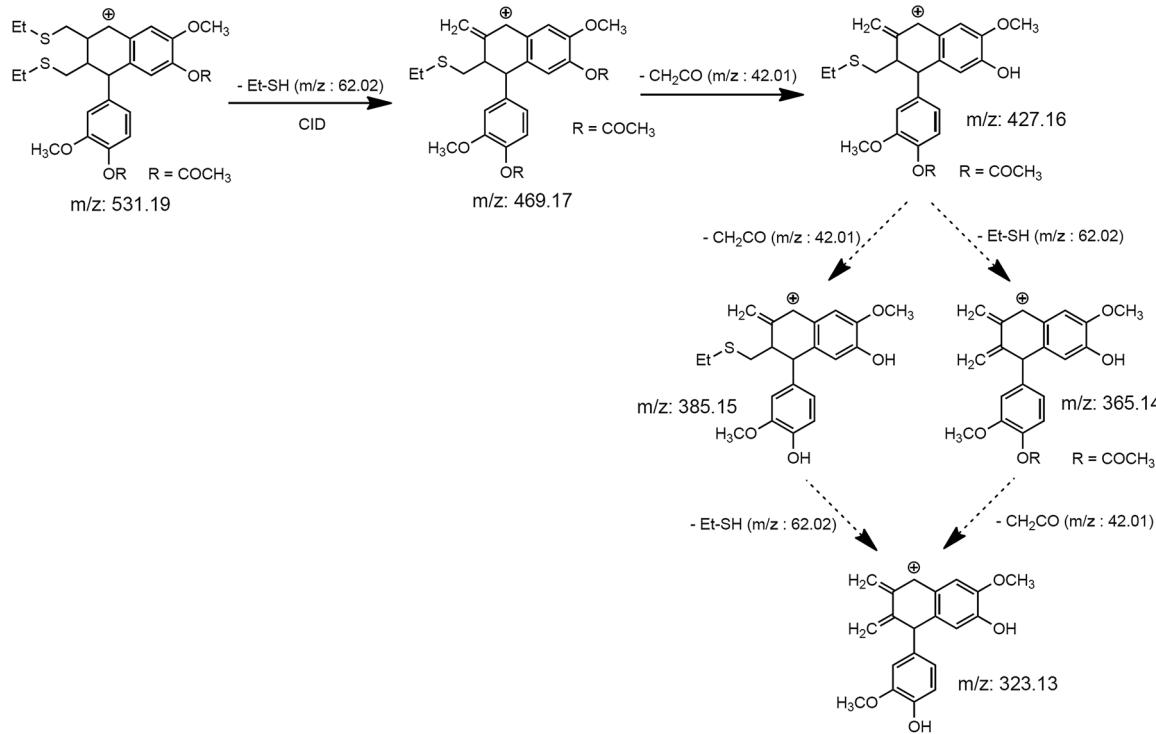


Figure 18. The CID-fragmentation pattern of the thioacidolysed acetylated pinoresinol at m/z 531.19. Reproduced with permission from H. Onnerud *et al.* (2003). Copyright 2003 by Walter de Gruyter GmbH & Co. KG. All rights reserved.

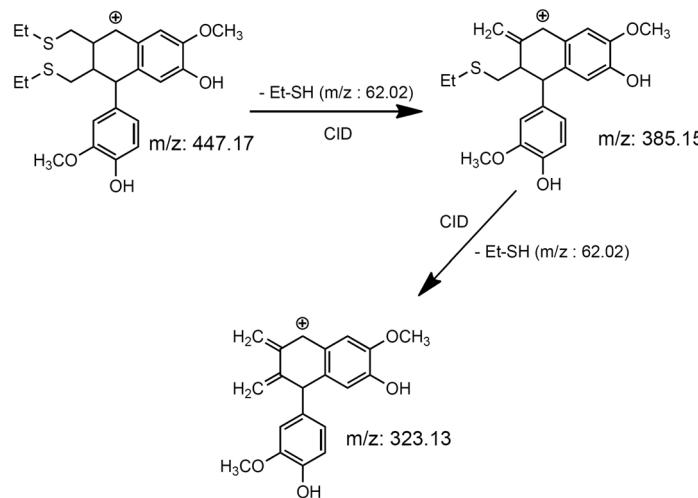


Figure 19. The CID-fragmentation pattern of thioacidolysed pinoresinol 1 without acetylation. Reproduced with permission from H. Onnerud *et al.* (2003). Copyright 2003 by Walter de Gruyter GmbH & Co. KG. All rights reserved.

addition, they used tandem MSⁿ to elucidate the structure of thioacidolysed pinoresinol before and after acetylating the products.

The ESI-MS (+ ion mode) of the thioacidolysed acetylated pinoresinol and the thioacidolysed pinoresinol did not afford the expected protonated molecules but instead produced the pseudo-molecular ions obtained by elimination of the ethane thiol CH₃CH₂SH (−62 u) assigned as the acetylated dithioether [M_{Ac}−SCH₂CH₃]⁺ and the non-acetylated dithioether [M−SCH₂CH₃]⁺, respectively, at *m/z* 531.19 and *m/z* 447.17.^[129]

The acetylated dithioether [M_{Ac}−SCH₂CH₃]⁺ precursor ion at *m/z* 531.19 was analysed by CID-MS/MS, which showed two major product ions at *m/z* 469.17 and 427.17 that were attributed respectively to the loss of the second CH₃CH₂SH group and a molecule of ketene. Consequently, these two product ions were assigned as [M_{Ac}−CH₃CH₂S−CH₃CH₂SH]⁺ at *m/z* 469.17 and [M_{Ac}−SCH₂CH₃−CH₂CO]⁺ at *m/z* 427.17, respectively. The latter product ion at *m/z* 427.16 also eliminated individually a molecule of formaldehyde and thioethanol to afford respectively the ions at *m/z* 385.15 and 365.14 (Fig. 18).

In contrast, MS/MS of the non-acetylated thioacidolysed precursor ion at *m/z* 447.17 afforded only two product ions [M−SCH₂CH₃]⁺ and [M−SCH₂CH₃−HSCH₂CH₃]⁺, respectively, at *m/z* 385.15 and 323.13 (Fig. 19).^[129]

On the other hand, it is important to note that the overall performance of the ESI-FT-ICR-MS of lignin did not provide unambiguous information about the detailed structure of the milled wood lignin given that many of the expected degradation products could not be identified.^[129]

In reality, when Onnerud *et al.* used FT-ICR-MS (+ ion mode) to analyse the total thioacidolysis extract of PML wood lignin, they obtained a very heterogeneous MS composed of a multitude of lignin ions up to about 1500 u, including a high abundance ion at *m/z* 791.28. The authors suggested that the assignment of this series of specific ions was complex. In addition, they indicated that the ion at *m/z* 791.28 was the most abundant ion (base peak) and difficult to assign. Indeed, CID-MS/MS analysis of this precursor ion did not yield any product ions under any CID conditions used. The authors indicated that this CID-MS/MS resilience of the precursor ion at *m/z* 791.28 was due to its very stable structure. Nevertheless, the authors indicated that the position of the thioethyl groups

found in this precursor ion was most probably different to that obtained from the thioacidolysed pinoresinol.^[129]

From the relative abundances of the different isotope peaks, Onnerud *et al.* calculated that the second isotopic peak (third peak) containing ³⁴S, [M−³⁴SCH₂CH₃]⁺, at *m/z* 793.28 was 0.278 compared with the base peak, the monoisotopic [M−³²SCH₂CH₃]⁺ at *m/z* 791.28. (Fig. 20) This value was compared with a theoretical value of the isotopic ion distribution measured using computer software, and it provided the number of sulfur atoms in the compound. For that reason, the molecular formula was calculated to be C₄₃H₅₁O₈S₃. The authors also calculated the double-bond equivalent (DBE) value for this molecular formula, which was found to be 18.5. This DBE value corresponded to a tetrameric structure containing four aromatic rings (that is 16 DBE), two double bonds (2 DBE) and the + charge (0.5 DBE). Accordingly, the authors calculated that the theoretical *m/z* would be 791.27. Onnerud *et al.* concluded that the non-acetylated thioacidolysed PML C₄₃H₅₁O₈S₃ at *m/z* 791.28 was a tetramer consisting of two coniferyl alcohol units, connected through a β-O-4' bond, with further links through a β-β' and β-1' structure (Fig. 20).^[129]

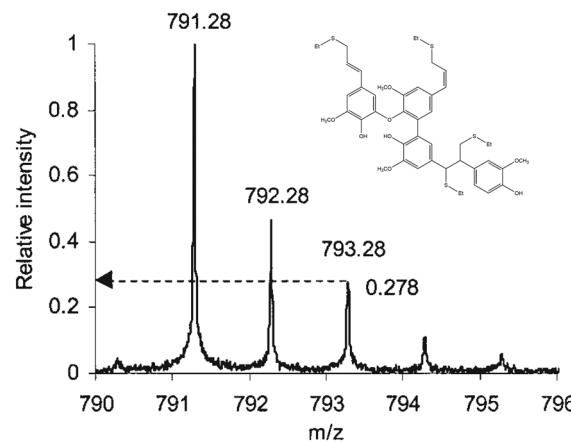


Figure 20. ESI-FT-ICR-MS (+ ion mode) of the total thioacidolysis extract of MWL. Reproduced with permission from H. Onnerud *et al.* (2003). Copyright 2003 by Walter de Gruyter GmbH & Co. KG. All rights reserved.

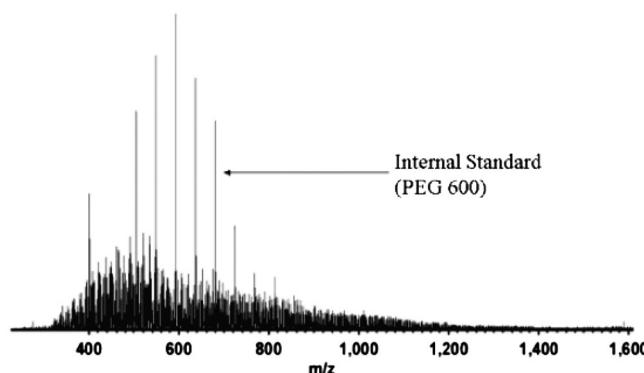


Figure 21. SRFA, positive-ion ESI-FTICR broadband mass spectrum, calibrated from a PEG 600 internal standard (1500 coadds). Reproduced with permission from A.C. Stenson *et al.* (2003). Copyright 2003 by American Chemical Society. All rights reserved.

ESI-FT-ICR-MS analysis of the degradation lignin products from fulvic acids

Dissolved organic matter is a class of compounds composed of degradation and condensation products of naturally occurring biomolecules such as humic acids, fulvic acids and industrial degraded modified lignin.^[89] The complexity and diversity of the various product mixtures in dissolved organic matter do not allow separation of a single component molecule by liquid-phase separation techniques.^[115] Nevertheless, Kujawinski *et al.* have shown that it was possible to volatilize humics with soft ionization techniques such as MALDI and ESI.^[90]

Stenson, Marshall and Cooper assigned the molecular formulas for 4626 individual Suwannee River fulvic acids. The described assignments were based on accurate mass measurements from ions generated by ultrahigh-resolution ESI-FT-ICR-MS.^[89,90] The formula assignments were made possible based on the mass accuracy of FT-ICR-MS at high field (9.4 T) and the regular mass spacing patterns found in fulvic acid mixtures.^[89,90] The authors sorted their observed 4626 ions individually according to the Kendrick mass defect

and the nominal mass series (z^* score), which permitted that all could be assigned to 1 of 266 distinct homologous series that differ in oxygen content and double-bond equivalence.^[90,131]

As well, Stenson, Marshall and Cooper performed MS/MS using infrared multiphoton dissociation for the identification of the fulvic acid molecules. Accordingly, these chemical formulas provided plausible structures to be proposed that were consistent with degraded lignin as a source of Suwannee River fulvic acids.^[89,90]

The lignin-based character of Suwannee river fulvic acids was further corroborated by the identification of predicted lignin degradation products.^[89,90] It is important to note that the authors concluded that QTOF-MS (moderate resolution) and double-focusing mass spectrometers are not able to fully resolve individual humic ions.^[90,132] They also indicated that it was the virtual absence of nitrogen in humic/fulvic acids that until now has prevented mass-based identification of these constituents.^[89,90]

To indicate the complexity and difficulties in measuring ESI-MS of degraded lignin (PML), the positive-ion ESI-FT-ICR broadband mass spectrum of Suwannee river fulvic acids showed that it contained more than 9800 peaks (Fig. 21).

With the use of ESI-MS, it was also found that all ions were singly charged, on the basis of the absence of peaks at intervals of $(1.0034/2)$ higher in mass than the monoisotopic ion.^[90,132,133] Very importantly, Stenson, Marshall and Cooper specified that the degraded lignin was the major constituents of Suwannee river fulvic acids. Accordingly, the assigned formulas and observed mass patterns obtained by ESI-MS were therefore assessed for uniformity with the expected formula for the degraded lignin.^[89] The most common lignin degradation pathways include depolymerization, demethylation, side-chain oxidation and aromatic ring cleavage (Fig. 22).

The authors showed that for any starting material, the degraded lignins was a complex mixture of molecules, consisting of C, H and O that differ in degrees of saturation. Also, it was proposed that different lignin degraded products could possess differences in bonding patterns, types of degradation pathways, number of CH_2 groups and O atoms.^[90,132]

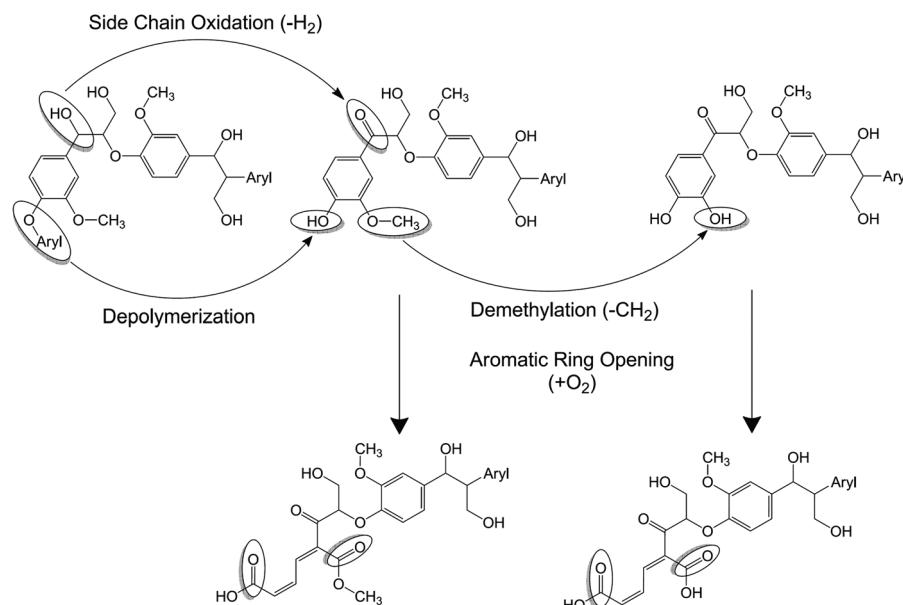


Figure 22. Schematic representation of common lignin biodegradation pathways. Reproduced with permission from A.C. Stenson *et al.* (2003). Copyright 2003 by American Chemical Society. All rights reserved.

Finally, the lignin-based character of Suwannee river fulvic acids was corroborated by detecting the expected lignin degradation products. It was also found that lignin degradation usually involves ring cleavage, resulting from the concurrent addition of two oxygen atoms to the chemical formula. Therefore, formulas differing by O_2 (e.g. one aromatic ring cleavage) and CH_4 versus O_2 (e.g. one aromatic ring cleavage after one demethylation and one side-chain oxidation) should be present among the assigned formulas.

LC-ESI-QIT-MSⁿ and FT-ICR-MSⁿ analyses of medicinal lignans

Huang and coworkers presented the ESI-MS analyses of a series of lignans known as schisandrin, schisantherin A, deoxyschisandrin and schisandrin B lignan and other reference compounds by LC-ESI-QIT-MSⁿ and high mass resolution MS experiments performed with an 7.0 T FT-ICR-MSⁿ (+ ion mode). The chemical structures schisandrin, schisantherin A, deoxyschisandrin and schisandrin B lignan are shown in Fig. 23.^[134] Huang *et al.* compared the retention time, UV spectra and ESI-MSⁿ data of the analysed compounds, with data obtained by the reference compounds and the literature. In the study, a total of 16 lignan constituents were unambiguously identified and characterized.

The precise masses measured by FT-ICR-MS of the protonated molecules of schisandrin at m/z 400.18922, schisantherin at m/z 371.14988, deoxyschisandrin at m/z 402.29431 and schisandrin B

at m/z 386.17295 and all other associated lignans are shown in Table 4. Such were used to identify lignan constituents isolated from a methanol extract of the *S. chinensis* (Turcz.) Baill.^[134]

The ESI-QIT-MS of schisandrin afforded the protonated molecule at m/z 433.2 and only one fragment ion assigned as $[M + H - H_2O]^+$ at m/z 415.1.

For simplicity purposes, only the multistage CID-MSⁿ fragmentation routes are shown of the protonated molecule of schisandrin $[M + H]^+$ at m/z 415.1, which can be used as a general model for the identification of unknown lignans. The various multistage CID-MSⁿ fragmentation routes are self-explanatory and are indicated in Scheme 1. This strategy formed the basis for the authors to successfully identify a series of 16 lignan compounds in the methanol extract of *S. Chinensis* (Turcz.).^[134]

De novo lignomics

In lignomics research, no explicit sequencing methods exist to establish the primary structure of complex and simple oligomers. Researchers are required to synthesize authentic compounds as standards to enable verification and comparison with the MS/MS scheme obtained by an unknown compound. It is well acknowledged that lignin polymer and/or oligomer are one of the last biopolymers to defy attempts to determine their full sequences and structures. Nevertheless, shorter oligomers of lignin can be extremely useful in providing the blueprints on how the full lignin

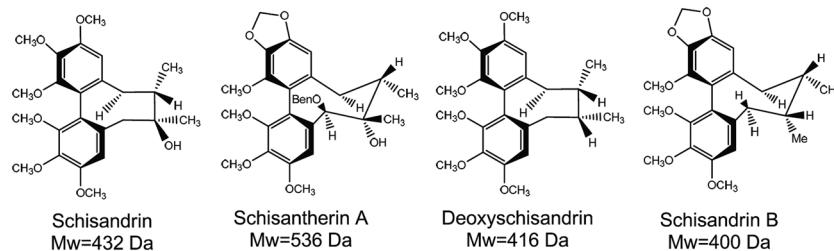
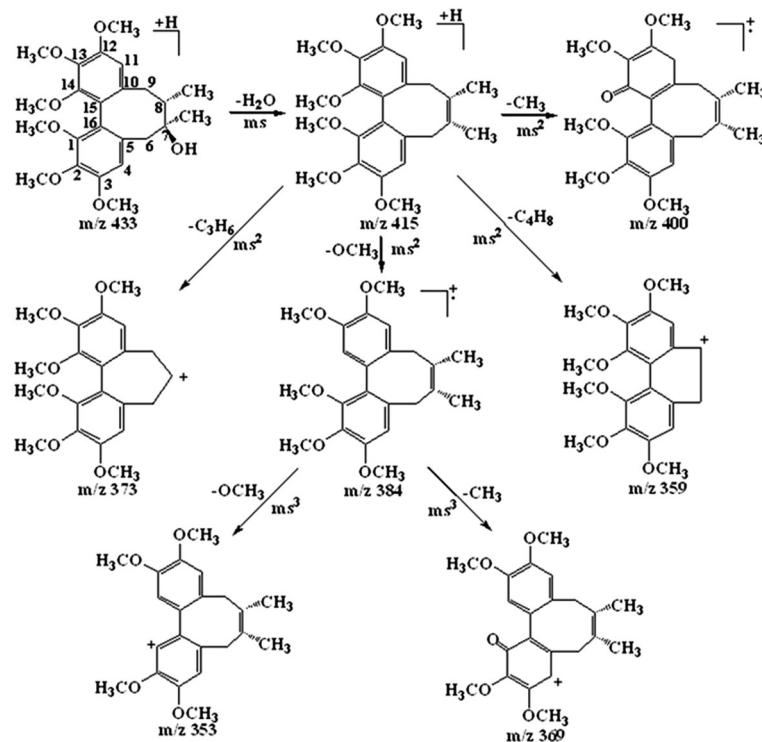


Figure 23. Lignan compounds in *S. chinensis* (Turcz.) Baill. Fruits. Reproduced with permission from X. Huang *et al.* (2007). Copyright 2007 by John Wiley & Sons, Inc. All rights reserved.

Table 4. Lignan compounds in *S. chinensis* (Turcz.) Baill. fruits. Reproduced with permission from X. Huang *et al.* (2007). Copyright 2007 by John Wiley & Sons, Inc. All rights reserved

Name	Proposed formula	Observed mass (Da)	Calculated mass (Da)	DBE	Error (ppm)	Proposed neutral loss
<i>Schisandrin</i>	$C_{23}H_{28}O_6^+$	400.18922	400.18804	10.0	2.9	CH_3
	$C_{23}H_{28}O_5^+$	384.19411	384.19313	10.0	2.6	CH_3O
	$C_{21}H_{25}O_6^+$	373.16563	373.16457	9.5	2.8	C_3H_6
	$C_{20}H_{23}O_6^+$	359.14997	359.14892	9.5	2.9	C_4H_8
<i>Schisantherin A</i>	$C_{21}H_{23}O_6^+$	371.14998	371.14892	10.5	2.9	C_2H_4O
	$C_{23}H_{25}O_6^+$	397.16573	397.16457	11.5	2.9	H_2O
	$C_{22}H_{25}O_6^+$	385.16568	385.16457	10.5	2.9	CH_2O
	$C_{20}H_{21}O_7^+$	373.12931	373.12818	10.5	3.0	C_3H_6
<i>Deoxyschisandrin</i>	$C_{19}H_{23}O_6^+$	347.14945	347.14892	8.5	1.5	C_5H_{10}
	$C_{18}H_{20}O_5^+$	316.13102	316.13053	9.0	1.5	$C_6H_{13}O$
	$C_{23}H_{30}O_6^+$	402.20431	402.20369	9.0	1.5	CH_3
	$C_{22}H_{26}O_5^+$	370.17805	370.17748	9.5	1.5	C_2H_7O
<i>Schisandrin B</i>	$C_{23}H_{30}O_5^+$	386.20938	386.20878	9.0	1.6	CH_3O
	$C_{18}H_{19}O_6^+$	331.11809	331.11762	9.5	1.4	C_5H_{10}
	$C_{17}H_{16}O_5^+$	300.09968	300.09923	10.0	1.5	$C_6H_{13}O$
	$C_{22}H_{26}O_6^+$	386.17295	386.17239	10.0	1.5	CH_3
	$C_{22}H_{27}O_5^+$	371.18579	371.18530	9.5	1.3	CH_2O
	$C_{22}H_{26}O_5^+$	370.17798	370.17748	10.0	1.4	CH_3O

Chemical structure of native and modified lignins



Scheme 1. Proposed MS^n fragmentation pathway of schisantherin A. Reproduced with permission from X. Huang *et al.* (2007). Copyright 2007 by John Wiley & Sons, Inc. All rights reserved.

polymer is constructed. Limited structure elucidation of oligomers have been obtained following dehydrogenative polymerization studies.^[135] However, such experiments fall short of being able to deduce the structure of lignin because it can vary among cell types and individual CW layers.^[136]

On the basis of the limited information from traditional methods and the complexity of lignin, new sequencing methods are likely to reveal novel oligolignol structures that do not concur with the current knowledge of lignin structures.^[52,53] Clearly, there is need for developing new and reliable sequencing methods for tackling the significant challenge of lignin structure determination. For this reason, the powerful technique of tandem mass spectrometric analysis has been explored for both VRL and PML. Indeed, MS/MS has been efficiently used to sequence native VRL, PML and the lignome.^[9,88,94,108,123,124]

Contrary to lignomics, proteomics has had the advantage of the availability of gene and genome sequence databases. In addition MS-based proteomics was initiated as a result of the technical and conceptual advances of protein ionization methods such as ESI-MS and MALDI-MS, which led to the 2002 Nobel Prize in chemistry.^[137–139] In comparison, the use of MS/MS in the lignomics field is a relatively new field, which will very likely surpass the conventional and routine use of Py-GC-MS. Therefore, until recent times, the study of lignomics has lacked versatile MS/MS sequencing approaches.

In this review, by analogy with proteomics, a new concept of top-down lignomic strategy employing MS/MS will be introduced. Herein, the top-down lignomic strategy is described as the identification by MS of the native extracted lignins (VRL) obtained directly from the destructuring of the vegetable matter. Therefore, the structural characterization, identification and the sequence of the lignin will be performed by MS/MS experiments. This will hopefully permit the identification of all the components of the intact non-modified virgin lignins (VRL) being analysed. Moreover, the top-down approach enables direct measurement of the intact mass of

the heterogeneous LOs, as well as fragment ions information relating to the original oligomer lignin sequences.

Regrettably and in contrast, the bottom up lignomic strategy concept is not possible because the resulting lignins are already degraded, without any further chemical manipulations because of their excessive chemical handlings in an attempt to purify them. For example, if an extracted lignin was treated with alkali, heated for a long time and then modified by chemical manipulations, it is obvious that some crucial structural information will be lost.

Accordingly, using the bottom-up lignomic strategy will result in the identification and sequencing of degraded products alone, which do not represent native VRL. Nevertheless, there is a large series of synthetic standards that exist to enable MS/MS comparison with most of the common types of PML (modified degraded) lignin components.^[91,94,96]

Lastly, it is important to note that contrary to Py-GC-EI-MS, the MS analyses of VRL by ESI-MS, APCI-MS and APPI-MS afford a large number of ions that usually lead to complex spectra that are challenging to interpret. This will be discussed in the following sections.

De novo top-down sequencing of VRL wheat straw lignin by APPI-QqTOF-MS/MS

Atmospheric pressure photoionization MS has proven to be a powerful analytical tool capable of ionizing small to large LOs. The APPI mass spectra of the wheat straw lignin dissolved in a mixture of chloroform-methanol (3:1) extracted by the Compagnie Industrielle de la Matière Végétale (CIMV) process were recorded in the positive and negative ion modes with a QqTOF-MS/MS hybrid instrument, using toluene as a dopant.^[88] It is important to mention that very significantly, ESI-MS and APPI-QqTOF-MS measurements of the wheat straw afforded identical spectra.

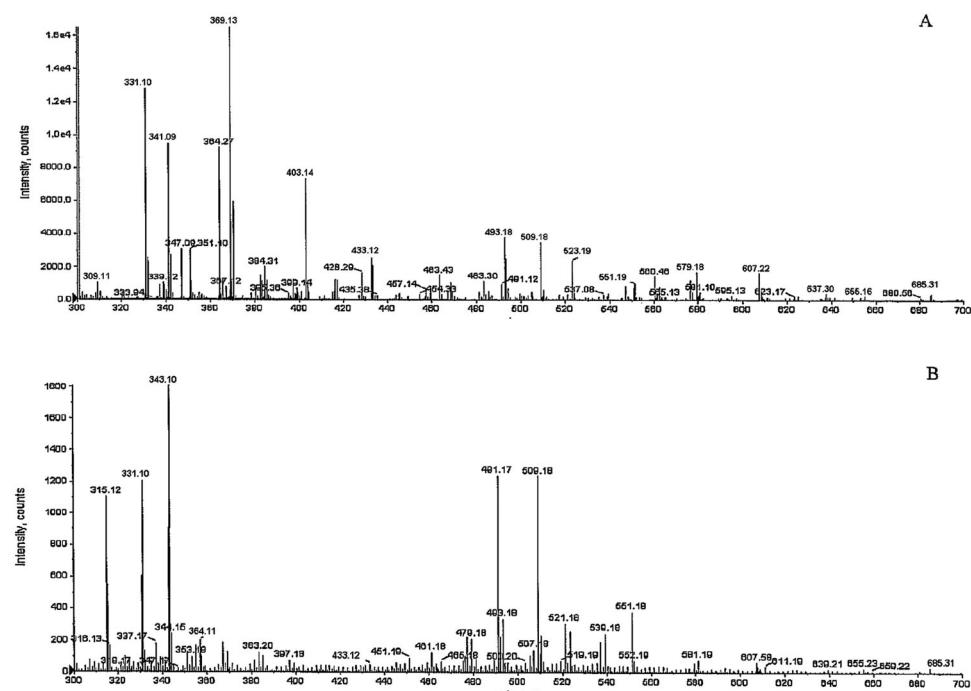
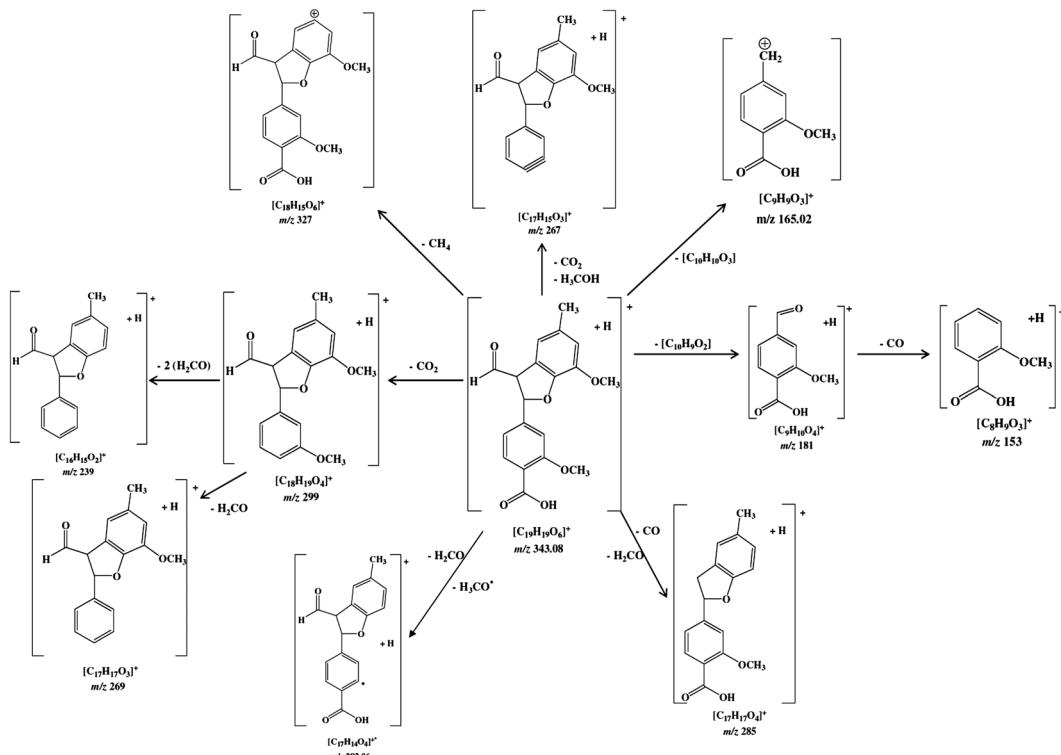


Figure 24. APPI-QqTOF-MS of the wheat straw lignins were recorded in the positive ion mode with scan ranges of m/z 300–700 with a DP = 60 V (A) and a DP of 100 V (B).

Banoub and Delmas have investigated the *de novo* sequencing of the VRL wheat straw lignin using the top-down approach. APPI-MS of technical wheat straw lignin was recorded in the positive ion mode with scan ranges of m/z 300–700 and m/z 500–900, with a DP = 60 V (A) and a declustering potential (DP) of 100 V (B) (Fig. 24).

Interestingly, the positive ion mode APPI-MS allowed us to tentatively propose structures for the 39 specific oligomeric ions, whereas negative ion APPI-MS allowed the identification of an additional presence of at least 18 specific oligomeric ions.^[88] Yet, it was noted that some oligomeric ion species photoionized in both



Scheme 2. Tentative breakdown processes in the MS/MS of the protonated phenylcoumaran derivative 4 $[C_{19}H_{19}O_6]^+$ at m/z 343.08. Reproduced with the permission of J. H. Banoub et al. (2007) Copyright 2007 by John Wiley & Sons, Ltd. All rights reserved.

Chemical structure of native and modified lignins

the positive and negative ion modes thus forming the respective protonated and deprotonated molecules. These observations are significant in revealing a high number of chemically distinct ions, indicating that the VRL oligomer was a very complex heterogeneous mixture. Of particular note, the use of APPI-QqTOF-MS has permitted the discovery of novel oligolignol structures that do not concur with the current knowledge of lignin structures. Most notably this series of oligomers was related and consistent with containing a repeating phenylcoumaran unit.

Low-energy CID-MS/MS analyses have been performed with a QqTOF-MS/MS hybrid instrument and provided unique dissociation patterns of the complete series of novel precursor ions. They also provided diagnostic product ions that enabled the tentative determination of structures and arrangement of the selected 57 different related ionic species. Similar MS/MS characterizations were carried out by Morreel and coworkers during their elucidation of the structures of large series oligomers possessing unknown linkages types present in *Populus* spp. and in the PML degraded products of *Nicotiana tabacum* (APCI-MS/MS-based Sequencing of Oligolignols).

In the APPI-CID-MS/MS study, the structure of the protonated molecule of the phenylcoumaran derivative $C_{18}H_{18}O_6$ $[M + H]^+$ at m/z 331.10 was unambiguously assigned as the protonated 4-carboxyl-coniferyl-(β -5')-(3-methoxylbenzene) unit, which forms the cyclic ether dimer constituent unit of all the phenylcoumaran derivatives. Unfortunately, in this study, it was not possible to make a comparison for the MS/MS dissociation behaviour of this series of the unknown ions with either the MS/MS of synthesized standards and/or any reported structures of biologically isolated lignin compounds.^[88]

One example of CID-MS/MS data is discussed in the following. The product ion scan of $[C_{19}H_{19}O_6]^+$ at m/z 343.09 afforded the product ion at m/z 327.05 by elimination of a molecule of methane. The precursor ion also lost a molecule of carbon dioxide to afford the product ion at m/z 299.06. The latter can lose a molecule of formaldehyde to produce the secondary product ion at m/z 269.06. Also, the precursor ion can lose simultaneously carbon monoxide and formaldehyde to form the product ion at m/z 285.05. The radical product ion at m/z 282.06 is created from the precursor ion by the consecutive losses of formaldehyde and a methoxyl radical. The secondary product ion at m/z 269.01 is formed from the ion at m/z 299.01 by loss of formaldehyde. The product ion at m/z 165.02 is produced from the precursor ion following the consecutive cleavages of the C-7-C-8 covalent bond of the coniferyl aldehyde, and the ether link formed between C-7 and the oxygen atom located on C-4' of the second coniferyl unit. (Scheme 2)

Recently, Del Río *et al.* in a repeated attempt to partially elucidate the structure of PML straw lignins reported that the flavone tricin was incorporated into lignin.^[66] According to del Rio *et al.*, tricin can form flavonolignan derivatives composed of a tricin skeleton linked to a phenylpropanoid (ρ -hydroxyphenyl or guaiacyl) moiety through a β -O-4' bond. For clarification purpose, tricin is 5,7-dihydroxy-2-(4-hydroxy-3,5-dimethoxyphenyl)-4H-chromen-4-one ($C_{17}H_{14}O_7$, MWt 330 u), and it is the aglycone moiety in the series of O-methylated flavone glycosides of the β -D-glucopyranoside-(1 → 4)-O-tricin type.^[66,140–144] It is important to prompt that del Rio *et al.* used a PML described by the authors as purified lignin. This PML was obtained by the classical Björkman method,^[145] which is known to result in 50% loss of the total lignin content.^[146] Likewise, it would be reasonable to expect that this PML does not bear any resemblance to the CIMV VRL wheat straw lignin.^[66]

It is needless to say that the protonated molecule of tricin $[C_{17}H_{14}O_7 + H]^+$, which appeared at m/z 331.10, has a different

molecular formulae and is isobaric to the ion reported by Banoub and Delmas for the phenylcoumaran derivative at $[C_{18}H_{18}O_6 + H]^+$ at m/z 331.10, obtained by APCI-QqTOF-MS, ESI-QqTOF-MS and APPI-QqTOF-MS analysis (+ ion mode) in study of the CIMV wheat straw lignin.^[88] Therefore, to clarify this position and avoid any ambiguity, we have analysed and compared the low-energy CID of the protonated precursor ion obtained from pure tricin and an extracted phenylcoumaran ion by ESI, APPI-QqTOF-MS in both ionization modes.

A quick perusal of the low-energy ESI-CID-MS/MS of both precursor protonated molecules at m/z 331 shown in Fig. 25 (inset m/z 290–307) reveals different intensities of product ions between the two isobaric precursor protonated molecules, which gas-phase fragment by different mechanisms.

Finally, the most essential element in this comparative study is that there is an unmistakable difference between the relative abundances of the isotopic profiles of the two isobaric protonated molecules obtained from tricin $[C_{17}H_{14}O_7 + H]^+$ and from the wheat straw $[C_{18}H_{18}O_6 + H]^+$ at m/z 331.10.

APPI-QqTOF-MS and ESI-QqTOF-MS of the extracted oak and pine (VRL) lignins

Recently, Delmas Jr., Delmas and Banoub measured the ESI-MS and the APPI-MS with a QqTOF-MS hybrid instrument of the VRL oak lignin extracted by the CIMV process (Fig. 26).^[147–149] It can be observed that there are differences between the two spectra, which

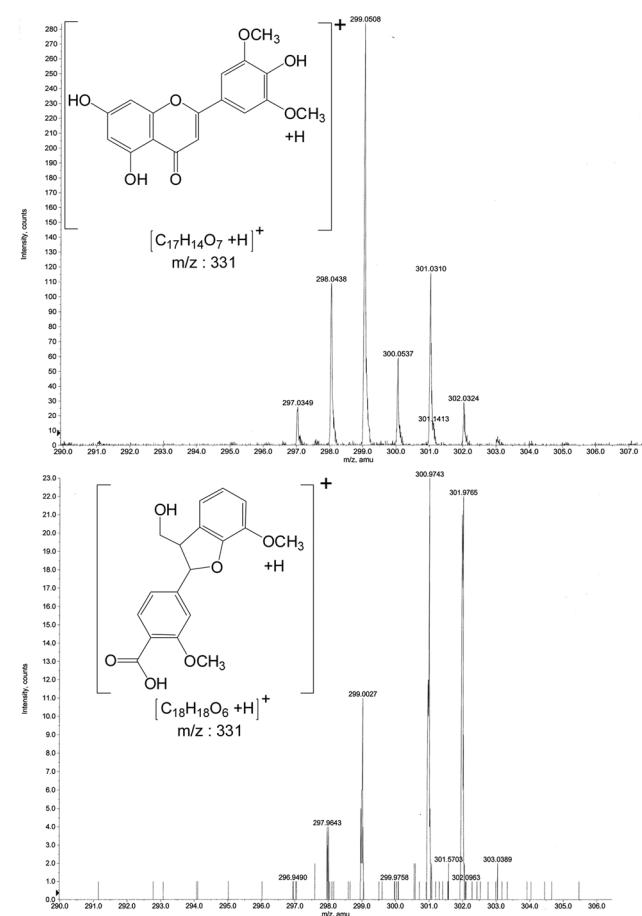


Figure 25. ESI-CID-MS/MS of the isobaric protonated molecules at m/z 331 recorded with a QqTOF-MS/MS hybrid instrument.

were recorded in the same mass range at m/z 100–700. Inessential to point out that, both the ESI-MS and APPI-MS contain a horde of ions that may look like background noise.^[150]

The ESI-MS and APPI-MS spectra VRL oak lignins show very clearly the complexity created by the huge number of monolignols, dilignols, trilignols and oligolignols present in the VRL. Additionally, it also shows clearly that for the same VRLs, there is a preferential specific ionization trend for the different oligomeric lignin analytes. However, this ionization trend was not noticed in the case of the VRL wheat straw, which gave almost identical spectra when recorded with both ESI and APPI ionizations.

Similar results were obtained by ESI-QqTOF-MS and APPI-QqTOF-MS of pine VRLs. The spectra obtained showed a multitude of overlapping ions in the mass range m/z 200–800, which showed approximately 600 signal ions (MS not shown). It is important to remember that ESI depends upon the electrostatic spraying of a solution containing ions, which is passed through a needle held at high voltage (kV) relative to a counter electrode and the formation of the spray is assisted by nebulization. Consequently, ESI is not an ion formation technique; rather, it is a technique for extracting ions

from the solution-phase into the gas-phase free of solvent for mass spectral analysis. On the other hand, APPI create ions that result from the light energy excitation, which favours the ion molecule reaction between the dopant and the analyte to pass on the charge.

Henceforth, a new challenge has emerged, as explained by Hockaday *et al.* in their elegant work on the ESI-MS and APPI-FT-ICR-MS for the characterization of organic matter in natural waters: 'If the composition of a complex mixture is unknown, how does one evaluate the performance of the spectroscopy used to characterize the unknown?'^[71]

Desorption/ionization MS of extracted and synthetic intact lignins

When dealing with VRLs and PMLs, which are non-volatile and practically insoluble, other softer ionization methods beside ESI, APCI and APPI can be used. This is where, desorption/ionization (DI) and particle-induced desorption MS shines, and these involve

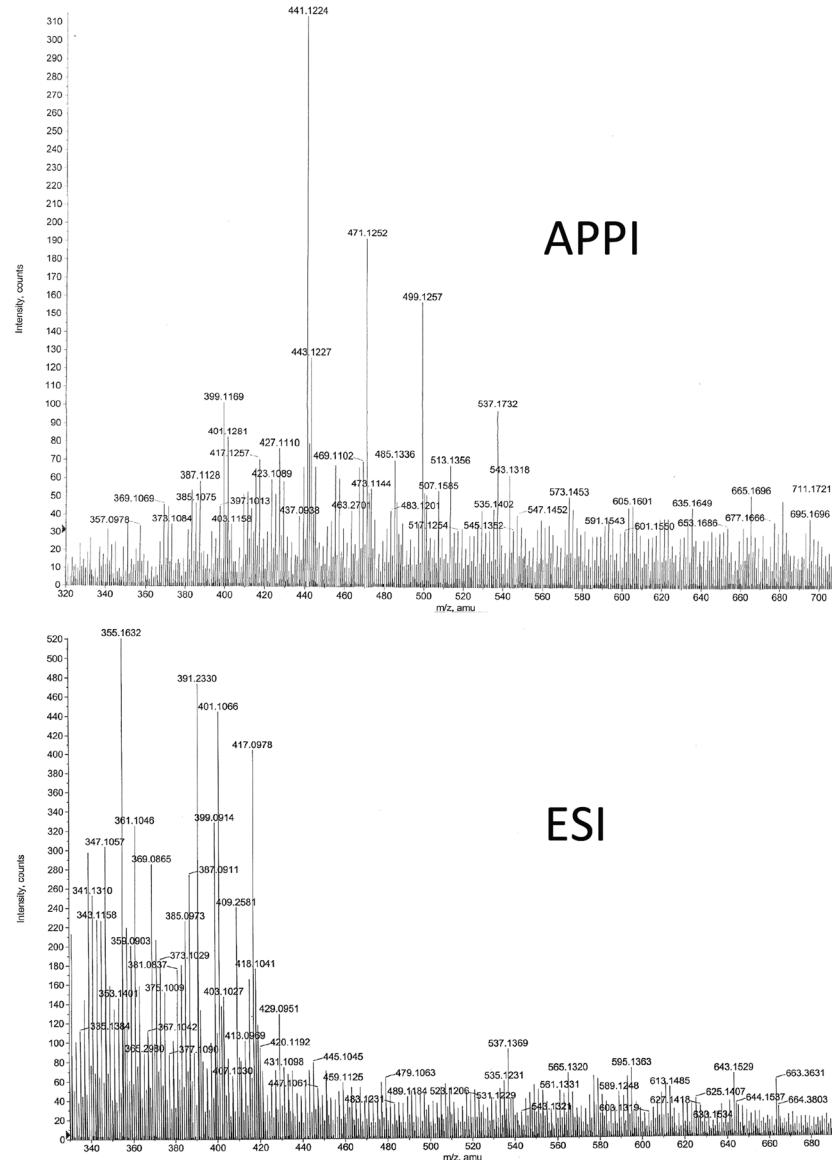


Figure 26. APPI-MS and ESI-MS of the CIMV extracted oak lignin recorded with a QqTOF-MS hybrid instrument.

Chemical structure of native and modified lignins

fast-atom bombardment, LSI-MS, plasma desorption and photon-induced desorption such as MALDI.^[151]

In the DI methods, the unifying aspect is the quick addition of energy into a condensed-phase sample, which results in the generation and release of gas-phase ions into the mass analyser.^[151] Instrumental details of the implementation of these ionization methods are not reviewed in this critique. However, some of the most recent and interesting applications reported on the structural analysis of lignin are included in this review.

MALDI-MS versus NALDI-MS analysis of isolated and synthetic lignins

Yoshioka *et al.* prepared synthetic lignin (DHP) by polymerization of coniferyl alcohol with horseradish peroxidase, according to the Zutropfverfahren method (drop-wise addition method), which favours end-wise polymerization. This polymerization system is known to produce larger numbers of β -O-4' bonds and more significantly represses condensation of the β - β' coupling reaction. If the β - β' coupling reaction was desired, it could be prepared by introducing coniferyl alcohol immediately into the dehydro-polymerizing enzyme solution (Zulaufverfahren DHP).^[152] The authors compared nano-assisted laser desorption/ionization time-of-flight MS (NALDI-TOF-MS) and MALDI-TOF-MS, to obtain the detailed structure of lignin. NALDI-MS is a novel innovative matrix-free approach, which eliminates the appearance of matrix peaks. NALDI-MS uses disposable nano-structured target plates on which the analyte is spotted. The nano-material on the target absorbs the laser energy and conveys it to the analyte for the DI process. As a result of the absence of the matrix, mass spectra present very low chemical background and increased sensitivity. The method is fast and reproducible and may be applicable to small molecular weight molecules.^[152]

When using MALDI-MS, the authors examined the matrices 2,5-dihydroxybenzoic acid (DHB), sinapinic acid and α -cyano-4-hydroxycinnamic acid, and they selected DHB matrix as the best. Because the matrix was immiscible with the sample solution, the sample was sandwiched between matrix layers.^[152]

The NALDI-TOF-MS data obtained for the synthetic lignin produced by the dehydrogenative polymerization of coniferyl alcohol

are shown in Fig. 27 in both the linear-positive and reflector-positive ion modes.

Henceforth, the authors observed characteristic ion masses of the Zutropfverfahren DHP when measured in the linear and reflector modes of NALDI-TOF-MS, mainly at intervals of 178 and 196 u, which correspond to the mass number of the monomeric unit of DHP and its Hydroxyl (HO) adduct. This means that the DHP from coniferyl alcohol polymerized with peroxidases, which proceed by oxidative radical couplings and a hydrogen atom, which was removed from the lignin monomer, to yield a mass increment of 178 u.^[152]

Accordingly, a dimer produced by oxidative radical coupling of coniferyl alcohols with a molecular weight of 358.4 u. This mass increment of 178.2 u corresponds to the formation of either β - β' and/or β -5' covalent bonds. Furthermore, a mass increment of 196 u indicated the existence of a β -O-4' bond, which is produced via the quinone methide intermediate followed by the addition of H₂O (Fig. 28).^[152]

When NALDI-TOF-MS was compared with the MALDI-TOF-MS of the DHP, it became apparent that the peak intensities of the ions in the high molecular mass range were higher than those of MALDI-TOF MS, but the spectral patterns were similar.

Similarly, Yoshioka *et al.* observed close similarity between the spectra of PML wood lignin from *P. densiflora* and DHP from coniferyl alcohol. Thus, a comparison between the NALDI-TOF-MS and MALDI-TOF-MS indicated sharp distinct differences in the low molecular mass regions between *m/z* 500 and 800 (Fig. 29).^[152]

Yoshioka *et al.* concluded that NALDI-TOF-MS of synthetic and PMLs, measured between the *m/z* 500 and 800 range, gave more intense specific ions than those obtained by MALDI-TOF-MS. Finally, they indicated that the detailed analysis of interunit linkages of lignin by the NALDI-TOF-MS method was an excellent means to clarify the structure of lignin.^[152]

MALDI-TOF-MS of PML and synthetic lignins

Bocchini and coworkers reported the MALDI-TOF-MS experimental conditions for the characterization of different types of PMLs and one type of synthetic lignin. Bocchini *et al.* measured the MALDI-TOF-MS of PML alkali extracted lignins from wheat straw, which is shown in Fig. 30. The authors described the data obtained for the

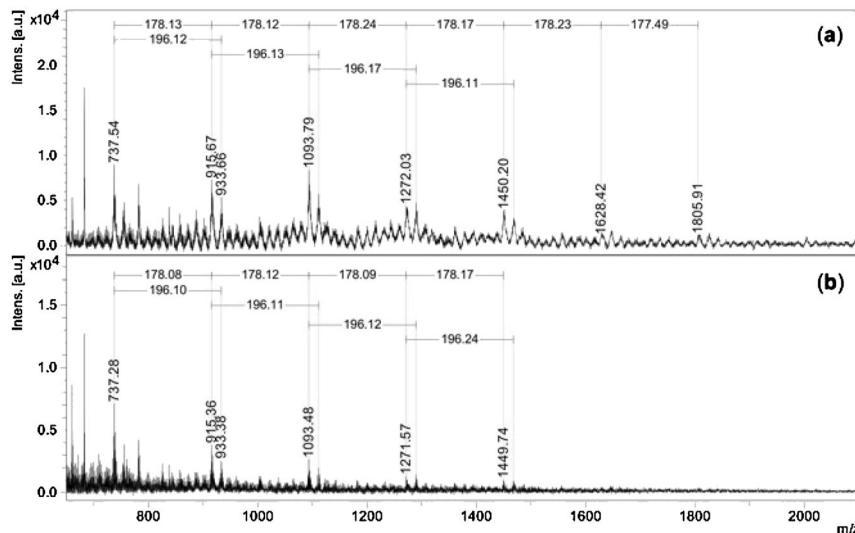


Figure 27. NALDI-TOF-MS of synthetic lignin obtained by dehydrogenative polymerization of coniferyl alcohol: (a) linear-positive ion mode and (b) reflector-positive ion mode. Reproduced with permission of K. Yoshiyoka *et al.* (2011). Copyright 2011 John Wiley & Sons, Ltd. All rights reserved.

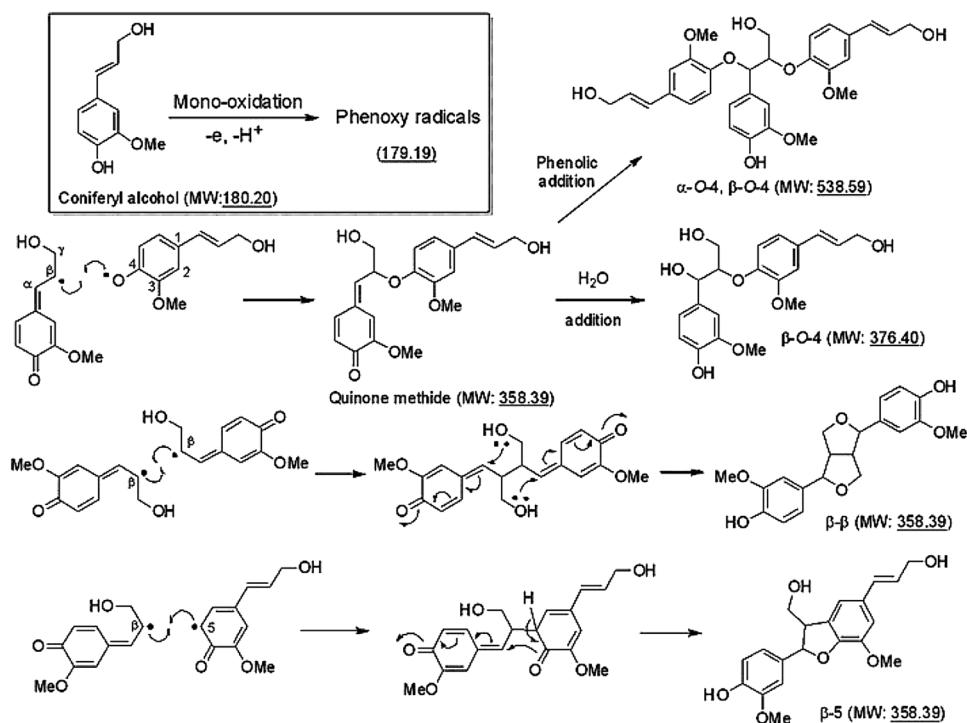


Figure 28. Formation of lignin substructures by dehydrogenation of coniferyl alcohol via radical intermediates. Reproduced with permission of K. Yoshiyoka et al. (2011). Copyright 2011 John Wiley & Sons, Ltd. All rights reserved.

molecular weight distribution of oligomers in the PMLs and also presented information on the monomeric constituents of different types of lignin.^[153]

The MALDI-TOF-MS showed a series of ions in the range of m/z 200 to m/z 3000. The authors assigned the structures of the ions obtained containing various oligomers, from monomers with m/z 200 to heptamers with m/z 1700: these are presented in Scheme 3. They also indicated that when the average molecular weight of the major fragment ion signals was divided by the degree of

oligomerization, the average molecular weight for the monomeric building units was found to be in the range of m/z 232–252. The authors noticed that the average monomeric molecular weight was higher than expected and attributed it to the presence of sodiated molecules.

Furthermore, Bocchini et al. measured the MALDI-TOF-MS of a synthetic G-type lignin obtained through oxidative polymerization of coniferyl alcohol.^[153] The MALDI-MS showed a series of ions, with better resolved fine structure than that of the PML alkali extracted

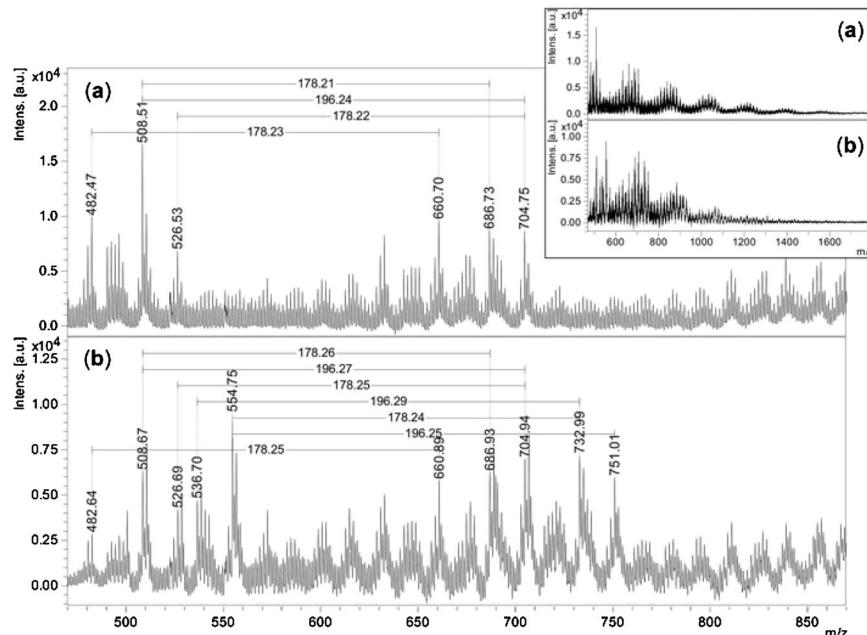


Figure 29. Mass spectra of isolated lignin obtained from *Pinus densiflora* (linear-positive ion mode). Measurement by (a) MALDI-TOF MS and (b) NALDI-TOF MS. Reproduced with permission of K. Yoshiyoka et al. (2011). Copyright 2011 John Wiley & Sons, Ltd. All rights reserved.

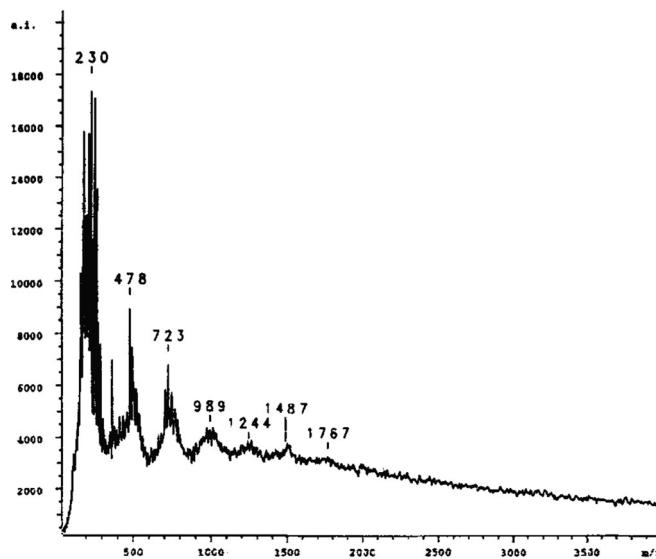


Figure 30. MALDI-TOF-MS of alkali lignin from wheat straw. Reproduced with permission from P. Bocchini *et al.* (1996). Copyright 1996 by John Wiley & Sons, Inc. All rights reserved.

lignin from wheat straw. The authors concluded that the synthetic lignin had a structure very similar to that of G-type natural lignin. Thus, it was found that MALDI-TOF-MS provided constructive information on the oligomeric distribution of different types of lignin.^[153]

Structural determination of β -O-4' synthetic lignin model polymer using MALDI-TOF-MS

Liu and co-workers synthesized hydroxyphenyl (H-type) and guaiacyl (G-type) lignin model polymers composed of the β -O-4'

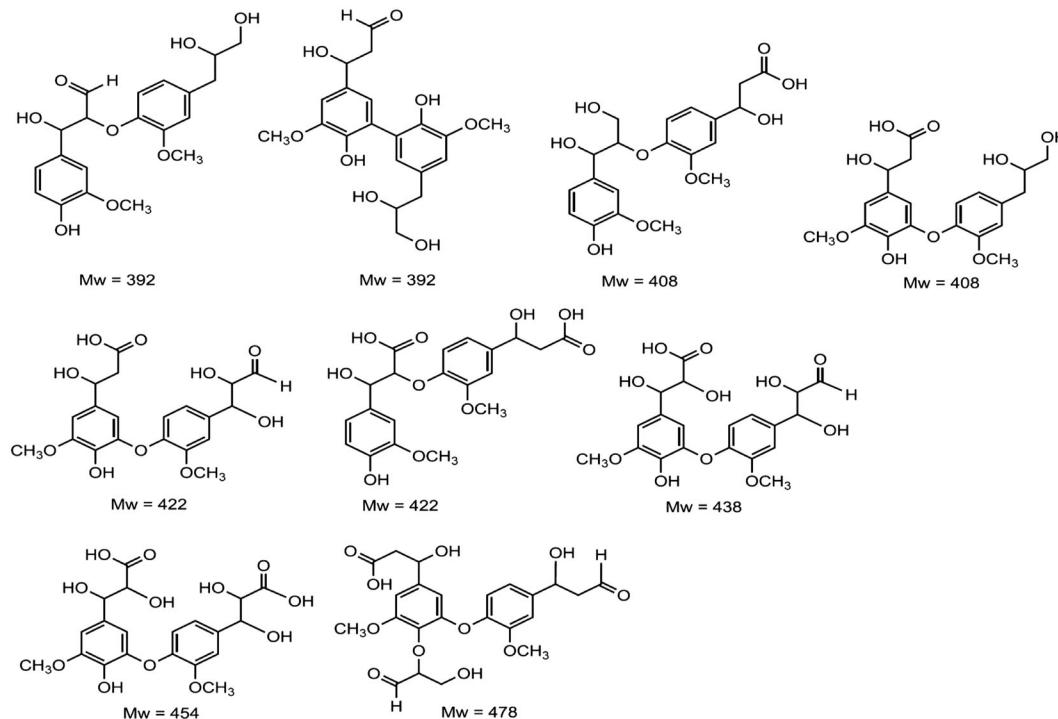
structure, which do not contain the γ -hydroxymethyl groups.^[154] The chemical structures of the H-type and G-type lignin models were characterized by $^1\text{H-NMR}$ and $^{13}\text{C-NMR}$, as well as MALDI-TOF-MS. The synthetic oligomers included exclusively the β -O-4' linkage, as assumed to be present in natural lignin. Accordingly, Liu *et al.* used the procedure described by Kishimoto *et al.*, which consisted of reacting the acetophenone derivative with CuBr_2 and a small amount of ethanol to prepare the corresponding bromides.^[155] Following addition of K_2CO_3 as the catalyst, the bromides were polymerized to yield the synthetic oligomeric lignin. Finally, the oligomer obtained was reduced with NaBH_4 in DMSO to obtain the H-type and G-type lignin model polymers containing only the β -O-4' linkage.^[155] The synthetic routes are depicted in Fig. 31.

The authors measured the molecular weight distributions by GPC in tetrahydrofuran. The GPC chromatogram indicated that for oligomers having an averaged molecular weight of 8285 u, 4832 u should be assigned to H-type and G-type lignin models.

Liu *et al.* analysed the H-type LO model by MALDI-TOF-MS and showed it to have a linear structure. Consequently, the molecular masses were found to be characterized by the factor $[136.16n + 2]$, in which n is the degree of polymerization. This indicated that the oligomers had regular structures and that reduction of carbonyl groups at the α -position was successfully conducted under the conditions used (Fig. 32).^[153]

MALDI-TOF-MS structural determination of laccase polymerized lignin

To study the efficiency of laccase-catalysed oxidation of lignin in the absence of mediators, Mattinen *et al.* characterized three different types of isolated lignin model compounds possessing high molecular weights by MALDI-TOF-MS.



Scheme 3. Some proposed dimeric structures consistent with the MALDI-TOF-MS of sample 2. Reproduced with permission from P. Bocchini *et al.* (1996). Copyright 1996 by John Wiley & Sons, Inc. All rights reserved.

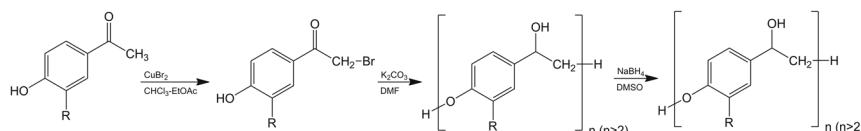


Figure 31. Synthetic route for H-type and G-type lignin models ($\text{R} = \text{H}$ or OCH_3). Reproduced with permission from J. Y. Liu *et al.* (2011). Copyright 2011 by North Carolina State University Department of Wood and Paper Sciences. All rights reserved.

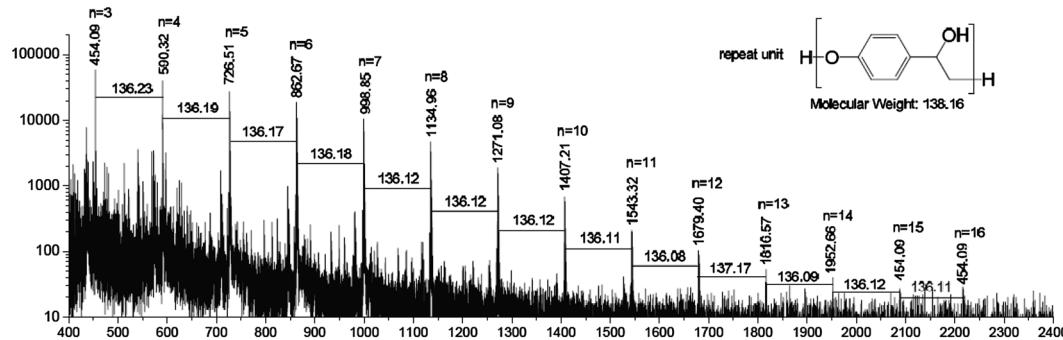


Figure 32. MALDI-TOF/MS of H-type lignin model polymer. Reproduced with permission from J. Y. Liu *et al.* (2011). Copyright 2011 by North Carolina State University Department of Wood and Paper Sciences. All rights reserved.

The tested model compounds of lignin originated from different botanical origins, that is, flax, softwood and hardwood.^[156] Accordingly, the authors used the method of Guerra *et al.* to isolate flax lignin from soda pulping black liquor and wood lignin from *Eucalyptus globulus* and Norway spruce pulp.^[146,157] Mattinen *et al.* postulated that the isolated lignin had the same morphology as the native lignin present in the CWS.

The authors used *Trametes hirsuta* laccase (ThL) for lignin activation and studying molecular weight distribution.^[156] After ThL treatments of all the lignins obtained from sources described in the aforementioned paragraph, the reaction mixtures were split into two parts. In the first part, the ThL activity was terminated with sodium azide, and the fraction was designed as liquid lignin. In the second part, the pH of the solution was acidified, and the precipitated lignin was centrifuged and washed with water. The products of each part were freeze dried and designated as solid lignin. The series of the lignin products obtained were studied by two different alkaline GPC methods.^[156]

Yet, it is important to note that when the GPC chromatograms were compared with the corresponding reference samples, they were surprisingly similar to each other. Clearly, the sensitivity of GPC was not high enough to be able to detect changes in the structure of ThL treated LOs.^[156]

Mattinen *et al.* also measured the MALDI-TOF-MS of the series of liquid and solid types of lignin obtained: the MS are shown in Fig. 33. In general, the MALDI-MS were louche, broad, had poor abundances and were noisy. They noted broad bands detected between m/z 30 000 and m/z 50 000 in all samples measured. The rather poorly louche resolved MS were presumably caused by the lignin samples not being fully ionized.^[156]

Molecular weight determination of PMLs by GPC and MALDI-TOF-MS

As shown throughout this review, the determination of the absolute molecular weight of lignin by GPC is an extremely tedious process. This is mainly due to the problems linked to the limited solubility, the presence of impurities and the interaction between lignin molecules. To further complicate matters, it is well known

that lignin can interact with the stationary phases in the chromatographic systems. In general, GPC can only be applied for relative estimation of the molecular mass of an identical class of lignin.

Jacobs and Dahlman showed that the PML Kraft lignin (Indulin AT) required to be better purified to obtain well defined and narrow dispersed fractions.^[158] For this reason, Indulin AT was fractionated by GPC for the accurate molar mass MALDI-TOF-MS to be determined. The MS obtained on the PML Kraft lignin sample (Indulin AT) showed a broad signal with the maximum intensity of around 700 u.^[158]

Gosselink fractionated the VRL by purification on an organic GPC column in THF.^[159] He obtained 12 GPC fractions (F1–F12), evenly distributed over the GPC chromatogram. The 12 fractions were analysed by MALDI-TOF-MS to give the broad signals (reflector and +ion mode) as shown in Fig. 34. Gosselink also indicated that the highest molar mass fraction F12, not shown in Fig. 34, could not be analysed by MALDI-TOF-MS.^[159]

Furthermore, Gosselink showed that there was a correlation between the molar mass of the highest peak in the MALDI-MS and the retention time in the organic GPC analysis.

In conclusion, Gosselink showed that the MALDI-TOF-MS of a series of unfractionated PMLs in alkaline solvent (DHB used as matrix) could be measured. Furthermore, he showed that the use of alkaline solvent did not interfere with the rather broad polydispersity of the MALDI-MS measurements.^[159] He also mentioned that addition of Ag-salts did not improve the MALDI results.

MALDI-TOF-MS of PML pyrolytic lignins

Bayerbach *et al.* used MALDI-TOF-MS, laser desorption/ionization time-of-flight MS (LDI-TOF-MS) and temperature-resolved analytical pyrolysis field ionization MS (Py-FIMS) for measuring the molecular masses of two extracts of pyrolytic lignin (PL).^[160] The PL was obtained from beech pyrolysis oil. The oil was produced in a fluidized bed reactor at a temperature of 470 °C. The oil was treated with ice cool demineralized water to give two PL fractions, which precipitated as powder-like brown product. The first one was precipitated

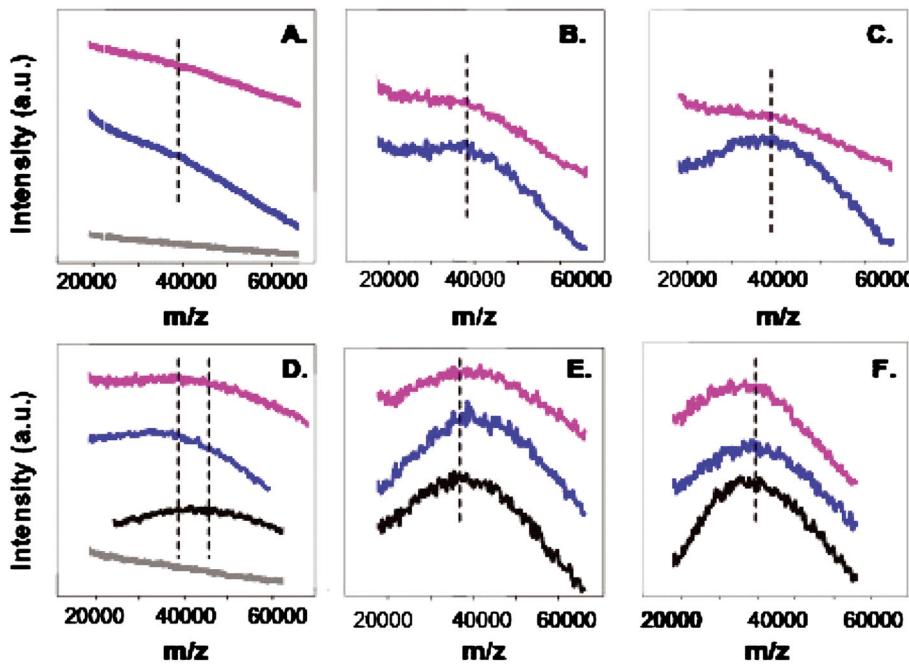


Figure 33. MALDI-TOF-MS of ThL-treated and untreated lignins. (A) 'liquid' and (D) 'solid' flax soda lignin; (B) 'liquid' and (E) 'solid' spruce EMAL lignin; (C) 'liquid' and (F) 'solid' *Eucalyptus* dioxane lignin. Pink = ThL-treated lignin, blue = the corresponding reference lignin and black = untreated lignin. The grey spectra in (A) and (D) were measured from the pure matrix. Reproduced with permission from M. L. Mattinen *et al.* (2008). Copyright 2008 by North Carolina State University Department of Wood and Paper Sciences. All rights reserved.

2 h after the oil was produced (PL-2h), and the second one (PL-2m) was precipitated 2 months later. The MALDI-TOF-MS of the PL-2h and PL-2m fractions showed ions that had a wide molar mass distribution.^[160]

Additionally, for comparison sake, the MALDI-TOF-MS and the GPC chromatograms were measured and are superimposed in Fig. 35. Hence, a smooth trend line can be seen on the MALDI-TOF-MS for the GPC data. The molecular masses of the PL fractions, measured by both methods, showed similar shapes, but the MALDI-TOF-MS showed more detail than the GPC chromatogram. In Fig. 35, the MALDI-TOF-MS of PL-2m shows high abundance ions at m/z 570, 750, 950, 1130, 1320 and 1500. With the GPC technique, only two maxima are located around 550 and 800 u. The MALDI-TOF-MS of PL-2h fraction shows less abundant ions at m/z 510 and 680. The GPC chromatogram shows only one maximum at 520 u

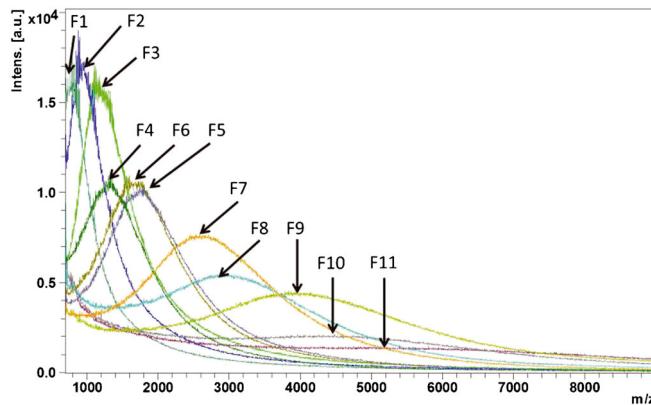


Figure 34. Positive reflective MALDI-TOF-MS spectra of fractionated Alcell lignin in THF (DHB used as matrix). Reproduced with permission from R. J. A. Gosselink (2011). Copyright 2011 by Wageningen Universiteit, The Netherlands. All rights reserved.

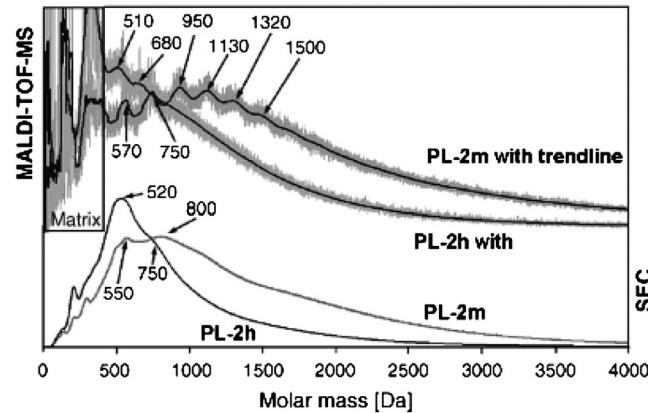


Figure 35. MMD of PL-2m and PL-2h, determined by MALDI-TOF-MS and SEC. For better comparison a trend line was added to the MALDI-TOF-MS signals. Reproduced with permission from R. Bayerbach *et al.* (2006). Copyright 2006 Elsevier. All rights reserved.

and a shoulder at 750 u. It was assumed that these masses could be assigned to PL-oligomers with different degrees of polymerization.

The intervals between the maxima determined by MALDI-TOF-MS (170–200 u) differ from the GPC values (230–250 u), and these intervals could represent the average absolute molar mass of PL-monomers.

Accordingly, the authors showed that the GPC mass distribution values of the PL-fractions differ from the visually interpolated MALDI-TOF-MS measurements by 20%.

Bayerbach *et al.* indicated that when they measured LDI-TOF-MS for the PL fractions, they obtained clearer spectra than when they were recorded with MALDI-TOF-MS. The authors noted that the intervals between main signals were mostly 14–16 u apart. In the spectra of different PL-fractions, the corresponding main signals

can vary between 2 and 4 u. These mass differences indicate variations in the aliphatic region of the PL molecules.^[160] Accordingly, they reported that the average molecular weight of the examined PL was between 560 and 840 u, with a maximal deviation of 20%, depending on the age of the oil. In addition, the authors showed that the masses of PL-dimers were between 270 and 400 u whereas, the PL monomers were between 130 and 200 u.

As expected, the Py-FIMS spectra contained ions of fragmented monomers and dimers in all samples investigated. The monomers detected corresponded to known lignin derived compounds obtained from bio-oil, and the dimers detected had similarities to phenylcoumaran structures. PL from aged bio-oil showed an increased content of higher oligomers and had a higher average molecular weight.^[160]

Bayerbach *et al.* concluded that the combination of various different methods for the molecular weight characterization, contributed to an understanding of the lignin structures, which would be extremely difficult to achieve by a single method.^[160]

SIMS-TOF-MS analysis of pine PML and synthetic dehydrogenated lignin oligomers

Secondary Ion Mass Spectrometry (SIMS-TOF-MS) has the advantages of rapid spectral acquisition, high spatial and mass resolutions, reasonable count rates and low damage offered by pulsed

liquid metal cluster primary ion sources.^[161,162] Saito and coworkers have shown that SIMS-TOF-MS with Ga primary ion bombardment was an excellent method for the structural analysis of different PML types of lignin because it was efficient in detecting the characteristic ions having a substituted guaiacyl aromatic ring.^[161,162]

The authors revealed that the SIMS-TOF-MS of the pine PML, containing the guaiacyl residues showed an ion at *m/z* 137, which was assigned as the $[C_8H_9O_2]^+$ ion composed of the C-7 benzylic ion of the guaiacyl unit G. In addition, the authors showed that the ions at *m/z* 151 were attributed to a pair of unresolved isobaric ions, namely, the C-7 benzoyl ion assigned as $[C_8H_7O_3]^+$ at *m/z* 151.0394 and the C-7-C β H₂-C γ H₂⁺ ion assigned as $[C_9H_{11}O_2]^+$ at *m/z* 151.0758, the latter ion being formed by the cleavage of the C-8-C-9 chain. The authors proposed that the fragment ions at *m/z* 137 and 151 were formed by rupture of common interunit linkages, such as the β -O-4', β -1', β - β' and β -5'. Neither of these ions could have been formed from the 5-5' interunit linkages (Fig. 36).^[161,162]

Also, both G and S types of lignin from beech wood were studied by SIMS-TOF-MS of PML (Fig. 36B). Analysis of this SIMS-TOF-MS showed that the ion at *m/z* 151 was derived from the G type and that the ions at *m/z* 167 and 181 were derived from the S type. As in the earlier studies, the ion at *m/z* 181 was an isobaric mixture of $[C_8H_9O_4]^+$ (the C-7 benzoyl ion) and the $[R-C\beta H_2-C\gamma H_2]^+$; both ions are derived from the S unit (Fig. 36).^[161,162]

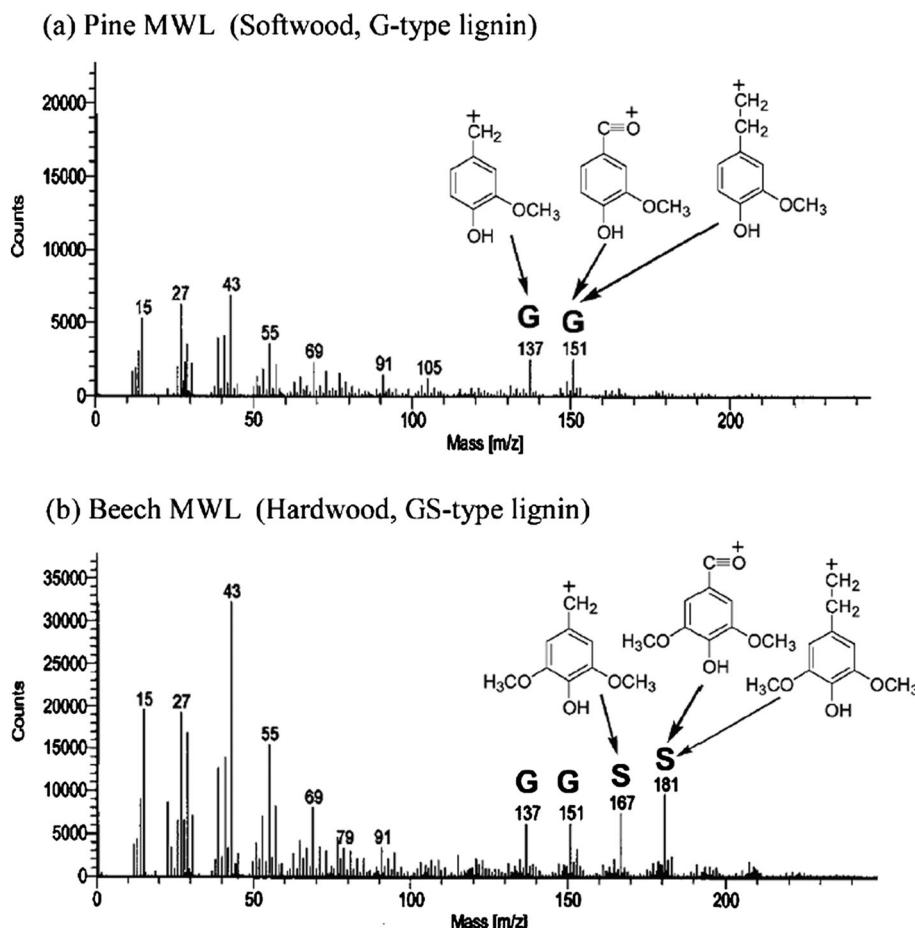


Figure 36. Common interunit linkages in lignin and positive TOF-SIMS spectra of (a) pine, and (b) beech milled wood lignin (MWL). The figure shows the tentative structures of the main secondary ions, each of which has a guaiacyl (G) or syringyl (S) ring. MWL is lignin isolated by extracting ball milled wood with dioxane-containing water (96 : 4). Pine (softwood) and beech (hardwood) MWL are composed of guaiacyl (G) and guaiacyl-syringyl (GS) lignins, respectively. Reproduced with permission from K. Saito *et al.* (2008). Copyright 2008 by Elsevier B.V. All rights reserved.

Chemical structure of native and modified lignins

Saito *et al.* suggested that the best lignin models were obtained from the dehydrogenation polymers DHPs (G type), which are usually synthesized from *in vitro* enzyme-catalysed polymerization of phenylpropanoid monomers.

The authors synthesized DHP from coniferyl alcohol aided by a peroxidase-H₂O system. The authors then used high-resolution SIMS-TOF-MS with Au⁺ primary ion for the ionization of DHP. The SIMS-TOF-MS of DHP showed two major ions: one at *m/z* 137 derived from G and the other at *m/z* 151. Once more, the isobaric ion at *m/z* 151 was presumed to be formed by the C-7 benzoyl ion assigned as [C₈H₇O₃]⁺ at *m/z* 151.0394 and as the ion [R-CβH₂-C₇H₂]⁺ at *m/z* 151.0758. Both of the ions were derived from the guaiacyl unit G and neither could be resolved (Fig. 36).^[161,162]

As depicted in Fig. 37A, the authors measured the SIMS-TOF-MS (+ ion mode) of the G(8-O-4')/C-7(CO) dimer, which gave the major ion at *m/z* 151 assigned as the C-7 benzoyl ion of G. They also showed in Fig. 37B that the SIMS-TOF-MS of the synthetic dimer composed of the 4-OMe-G(8-O-4')G produced ions at *m/z* 165, which was also assigned as the C-7 benzoyl ions.

When the authors measured the SIMS-TOF-MS of G(8-1'/C8-CH₂OH), they were able to resolve the isobaric ions at *m/z* 151 into the C-7 benzoyl ion assigned as [C₈H₇O₃]⁺ at *m/z* 151.0394 and the ion [R-CβH₂-C₇H₂]⁺ at *m/z* 151.0758 (Fig. 37C). The SIMS-TOF-MS of the G and S dimers showed that the [R-CβH₂-C₇H₂]⁺ ion was more intense than the C-7 benzoyl ion.

Finally, the authors measured the SIMS-TOF-MS of G(8-1'/C8-CH₃) having a C-8 side chain, and it appeared that ion at *m/z* 151 was mainly the major ion assigned as [R-CβH₂-C₇H₂]⁺ ion at *m/z* 151.0758, which existed with a small proportion of the C-6-C-1 benzoyl ion. Such an observation is indicative of a favoured cleavage between C-8 and C-9 being enhanced when the primary -CH₂OH group is replaced by a methyl group (Fig. 37D).^[161,162]

In conclusion, Saito *et al.* demonstrated that SIMS-TOF-MS using Au⁺ as the primary ion bombardment a new tool can be offered for the direct analysis of the fragments of lignin polymers.^[161,162]

Distinguishing between intact lignin and polysaccharides in SIMS-TOF-MS analysis of wood

As discussed earlier, Saito *et al.* were the first to use SIMS-TOF-MS as an analytical method for the evaluation of wood samples and for evaluating the distribution of the different molecular components of individual wood samples and different species.^[163] These analytical studies facilitated the characterization of engineered wood fibre and the impact of pre-treatment processes.

Recently, SIMS-TOF-MS imaging of wood surfaces has been used to show the location of the various types of inorganic ions.^[163] Tokareva and coworkers were able to measure SIMS-TOF-MS images of the different localization of the H-lignin subunits using the characteristic ions at *m/z* 107 and 121. They also measured the characteristic ions at *m/z* 137 and *m/z* 151 of the G-lignin subunits present in the transverse and longitudinal sections of aspen and spruce.^[164] As well, they also used SIMS-TOF-MS for imaging carbohydrates, using the characteristic ions at *m/z* 115 for hemicelluloses and the ions at *m/z* 127 and 145 for cellulose. Unfortunately, the lateral resolution of the ion images was insufficient to resolve both the carbohydrate and lignin constituents in the particular CW layers.^[164]

Similarly, Koljonen *et al.* described the structural changes in lignin and measured the lignin content on wood fibre surfaces by SIMS-TOF-MS imaging of the extractives and lignins on the pulp surfaces, using the ions at *m/z* 137 and *m/z* 151, which are diagnostic for the guaiacyl lignin residues, and the ion at *m/z* 77, which is characteristic of the phenyl ring (C₆H₅).^[165] Fardim and Duran

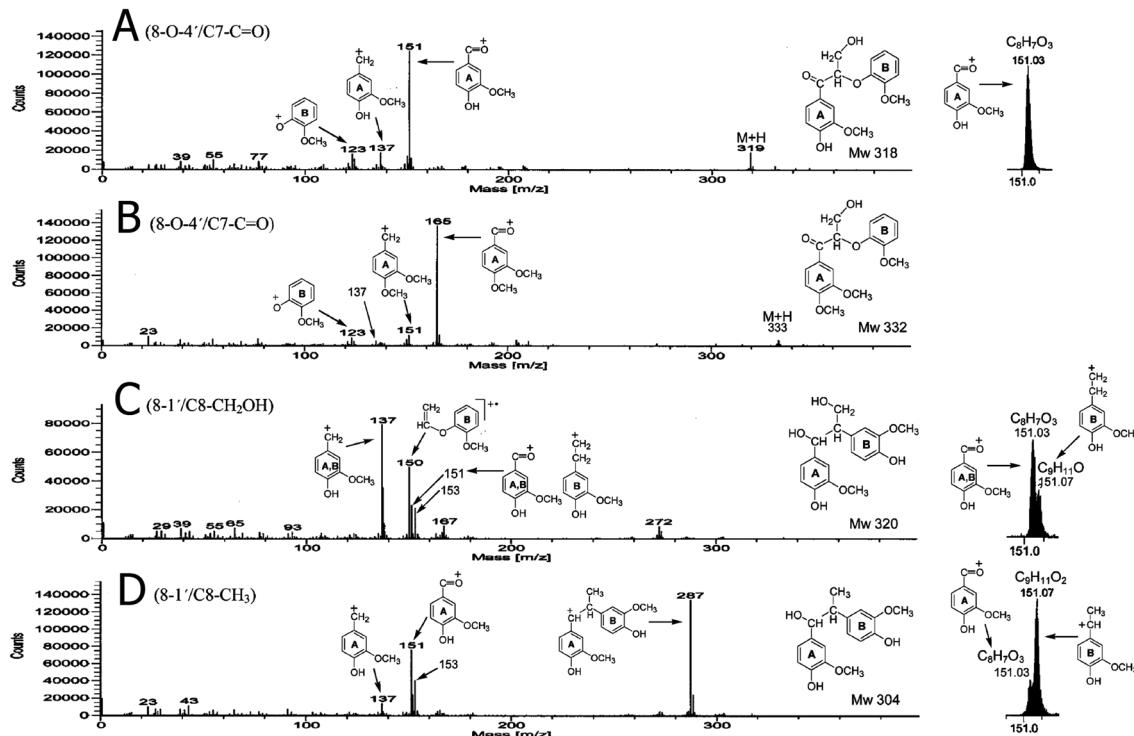


Figure 37. Positive TOF-SIMS spectra of lignin model dimers representing 8-O-4'/C7-C=O (compounds A and B), (C) 8-1'/C8-CH₂OH, (D) 8-1'/C8-CH₃. Reproduced with permission from K. Saito *et al.* (2005). Copyright 2005 by American Chemical Society. All rights reserved.

supplied lists of ions for extractives and ion assignments for the following: xylan (hemicellulose) at m/z 115 and 133; cellulose at m/z 127 and 145; and lignin at m/z 137, 151, 167 and 181.^[166] Berman *et al.* used SIMS-TOF-MS of the same series of ions to identify cellulose and xylan.^[167]

Goacher *et al.* used two different approaches to distinguish the ions in the SIMS-TOF-MS arising from the lignin and polysaccharide biopolymers in softwood.^[168] The first approach was to analyse samples with different relative proportions of lignin, cellulose and hemicellulose. The cellulose fraction samples were prepared by delignifying the extracted pine sawdust to holocellulose and removing both hemicellulose and the remaining lignin.^[168] The second approach was to generate high-resolution SIMS-TOF-MS images of solid pine cross-sections to determine the concentration variations between lignin and polysaccharide across CWs. This work formed the basis upon which Goacher *et al.* elucidated the main sources of variation among wood constituents using principal component analysis to highlight the differences between the spectra.^[168] These identifications were performed through combination of high mass resolution analyses of the pine fractions and the high lateral resolution image analyses to distinguish between the lignin-rich middle lamella and the secondary CW layers in solid wood cross-sections. Consequently, Goacher *et al.* was able to collect and compile ions from both high and nominal mass resolution spectra to extend the library of characteristic lignin and polysaccharide secondary ions in wood. It is important to appreciate that this library did not contain any ions that would create uncertainty between wood components.^[168]

Goacher *et al.* found that he could generate improved images using low-mass ions for the SIMS-TOF-MS imaging. Both the clarity and S/N improvements were attributed to the selective use of the low-mass diagnostic ions for polysaccharides and lignin, which are at higher intensities.

Finally, the lack of mass interferences from the ions identified by the present analysis was clearly demonstrated by the ability to obtain informative images of polysaccharide and lignin distributions using simple sum and division methods.^[168]

It is important to mention that in all the reported SIMS-TOF-MS imaging presented herein, no accounts were found of MS/MS analyses and resulting gas-phase fragmentations. It is needless to say that MS/MS analyses would have provided unequivocal and better identification of the lignin analytes.

Study of the removal of the lignin from cellulose and hemicellulose by imaging with direct matrix-assisted laser desorption/ionization tandem mass spectrometry

Lunsford, Peter and Yost reported that there is a variety of techniques used to image lignocellulosic materials.^[169] These methods include optical and fluorescence microscopies,^[170] magnetic resonance imaging,^[171] micro X-ray computed tomography,^[172] confocal Raman microscopy,^[173,174] and scanning electron and transmission electron microscopy. Although useful, these methods are disadvantaged by providing limited molecular specificity and structural information.

In contrast, Lunsford *et al.* showed that SIMS-TOF-MS and MALDI-TOF-MS were capable to provide much of the molecular specificity information that was unavailable by the other techniques.^[175–179] Nonetheless, even though SIMS-TOF-MS imaging can provide direct surface analysis and high spatial resolution of a lignocellulosic biomass, it is prone to extensive non-

specific fragmentations of the analyte, which limits its chemical specificity.^[180,181] This is attributed to the nature of the plant CW composition, which is made of complex building block carbohydrates such as glucose ($C_6H_{12}O_6$), xylose ($C_5H_{10}O_5$), glucuronic acid ($C_6H_{10}O_7$) and methyl glucuronic acid ($C_7H_{12}O_7$). As a result, SIMS ionization of these complex units yields complex fragmentations that are difficult to interpret.^[170]

To further complicate matters, the molecular mass similarities between the various sugars and the monolignols, which compose the lignocellulosic tissues, can produce isobaric ions. However, MS/MS analyses can overcome this complexity, by dissociating precursor ions and forming diagnostic product ions, which allow the isobaric ions to be distinguished and thereby allow an increase in the degree of confidence of ion identification.^[179,182]

For these reasons, Lunsford, Peter and Yost studied the direct imaging of cellulose and hemicellulose in intact wood tissue using a thermo MALDI-LTQ-MSⁿ XL tandem mass spectrometer equipped with a novel linear ion trap.^[169]

It is well known that microcrystalline cellulose contains semi-crystalline polymers of β -D-(1 → 4)-D-Glc linked monomers (Glc, $C_6H_{10}O_5$). Understandably, the molecular weight of cellulose can be determined by multiplying the mass of the glucose repeating unit, 162 u, by the number of glucose monomers and adding 18 u to account for the $^{\text{-}}\text{OH}$ and $^{\text{-}}\text{H}$ terminal groups.^[169]

Lunsford, Peter and Yost showed that the MALDI-LTQ-MS of microcrystalline cellulose of *Populus* wood tissue recorded between m/z 500–2000 formed ions that were separated by 162 u apart. They identified these signals as sodiated, dehydrated (singly charged) ions, differing by the number of glucose monomers, expressed as $[\text{Glc}_n\text{-H}_2\text{O} + \text{Na}]^+$ where $n = 3–12$. It is important to point out that the MALDI-LTQ-MS analysis of microcrystalline cellulose from *Populus* wood tissue, when coated with DHB matrix, showed very strong intensity ions throughout the range m/z 500–2000, which were not attributed to background noise. These ions were interpreted as arising from the ionization of various oligomers of lignin.^[169]

As mentioned earlier lignin is a complex polymer of polyphenol compounds found primarily in secondary CWs of gymnosperms, woody angiosperms and grasses and is abundant in the secondary xylem (wood) of the *Populus* tissue section.^[183] Therefore, to eliminate the complex series of the oligomeric lignin ion signals from intact tissue, the lignin was removed from the *Populus* wood sections using sodium hypochlorite extraction, by a procedure that leaves behind the delignified wood or holocellulose (composed of cellulose and hemicelluloses).^[146]

The delignified sections of the *Populus* tissue section were analysed by MALDI-LTQ-MS.^[169] Lunsford, Peter and Yost obtained spectra by averaging data from 50 MS scans recorded between m/z 500 and 2000 for untreated and delignified wood from serial sections, which were coated with DHB (Fig. 38). The authors noticed that the Total Ion Chromatogram (TIC) was reduced by more than 50% (from 3.08×10^8 to 1.26×10^8) and that the intensity of the cellulose analyte ion signal remained relatively unchanged (2.84×10^5 compared with 2.55×10^5). This suggests that the ionization of the LOs results in abundant background ion signal in wood tissue; thus, the signal-to-background can be improved by removing lignin from wood tissue.^[169]

Regrettably, the authors showed that despite the removal of the LOs, however, which improved the signal-to-background ratio, the obtained wood tissue had an altered wood structure that compromised both spatial and chemical information. This was explained by the presence of the remaining trace of lignin oligomeric ions

Chemical structure of native and modified lignins

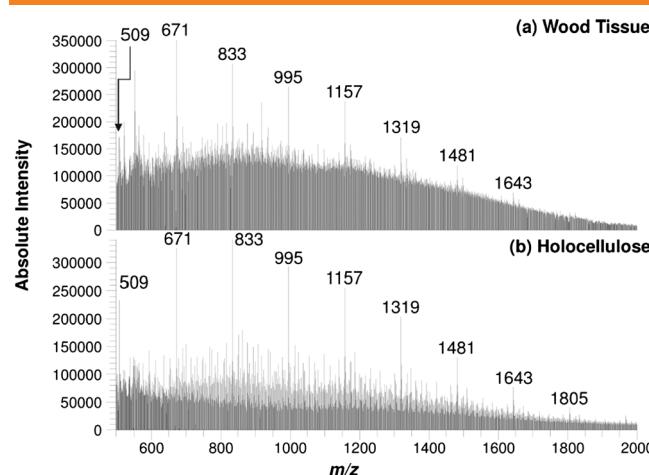


Figure 38. Comparison of MALDI-MS spectra of intact wood tissue (a) and holocellulose (b). Ions labelled were identified as singly charged, sodiated, dehydrated cellulose ions. Removing the lignin from the wood tissue removes much of the background ion signal (leaving behind the cellulose), illustrating that many of the ions observed in wood tissue are due to the ionization of lignin. Reproduced with the permission of K.A. Lunsford *et al.* Copyright 2011 by American Chemical Society All rights reserved.

complicating and downgrading the demand for more detail in the image dimension (spatial resolution) and in the spectral dimension (mass resolution and accuracy).

As an alternative means to solve this problem, the intrinsically high lignin ion signal can be effectively eradicated by using MS/MS.^[169] Hence, the MS^2 spectra for the precursor ion at m/z 1319 assigned as $[\text{Glc}_7\text{--H}_2\text{O} + \text{Na}]^+$ was recorded for the micro-crystalline cellulose standard (a) and the intact wood tissue (b). This MS^2 analysis of the precursor ion at m/z 1319 affords abundant product ions, arising from successive neutral losses (NLs) of 162 u resulting from glycosidic bond cleavages. The wood tissue spectrum shows additional product ions at m/z 1301 and 1275, (NLs of 18 and 44 u) resulting from loss of water and an acetyl

group, respectively. Additionally, this MS/MS result seems to indicate the existence of two isobaric ions obtained from wood tissue at m/z 1319.

Nevertheless when comparing the MS with the MS/MS results, a serious decrease in the lignins background ion signals compared with MS, and improves the signal-to-background ratio without removing the lignins (Fig. 39).^[169]

Similarly, the authors measured the MS^2 analysis of the precursor ion at m/z 1079 derived from Birch xylan extract and *Populus* wood tissue. In the xylan standard, the authors noticed the sequential NLs of 132 are observed in the MS^2 of m/z 1079 assigned as $[\text{Xyl}_8\text{--H}_2\text{O} + \text{Na}]^+$ which resulted from glycosidic bond cleavages between xylose monomers. Furthermore, the CID-MS/MS analysis of the precursor ion m/z 1079 derived from the *Populus* tissue showed product ions resulting from NLs at m/z 917 (NL of 162 u) at m/z 875 (NL of 204 u) and at m/z 847 (NL of 132 u). The authors indicated that as these product ions were not observed in the standard, at least two isobaric ions are likely to be present for the precursor ion at m/z 1079. Further stages of MS (MS^3 and MS^4) determined that, in addition to the linear xylan, the isobaric precursor ion at m/z 1079 was assigned as $[(162)_4 \times (204)_2 + \text{Na}]^+$ and was tentatively identified as an O-acetylgalactoglucomannan (Fig. 40).

In addition, the authors have shown a comparison between measured MALDI-LIT-MS with an average of 25 spectra averages in four different regions of wood tissue, inner pith, around the pith, secondary xylem and secondary phloem. The ions above the background were consistent with the m/z values of MCC and were observed in the spectra of the centre pith and secondary xylem. Nevertheless, it is important to realize that the intense lignin ions at nearly every m/z value are observed for the pith periphery and secondary phloem spectra. The differences in these spectra are consistent with the lignin composition of the regions: centre pith and secondary xylem are composed of mostly dead cells (more lignins and no polysaccharides), whereas the pith periphery and secondary phloem are composed of mostly living cells (Fig. 41).

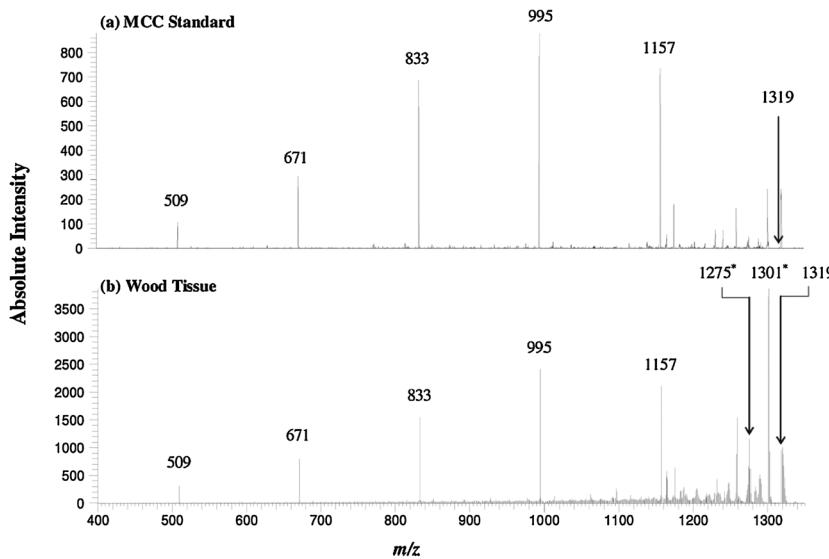


Figure 39. MS^2 spectra from MCC standard (a) and intact wood tissue (b) of m/z 1319, identified as $[\text{Glc}_7\text{--H}_2\text{O} + \text{Na}]^+$. Both spectra show abundant fragment ions, arising from successive NLs of 162 resulting from cleavages along the glycosidic bonds. The wood tissue spectrum shows additional fragment ions at m/z 1301 and 1275, (NLs of 18 and 44) resulting from loss of water and an acetyl group, respectively. This suggests at least two isobaric ions are observed from wood tissue at m/z 1319. Tandem MS decreases the background compound ion signal improving the signal-to-background ratio without removing lignin. Reproduced with the permission of K.A. Lunsford *et al.* Copyright 2011 by American Chemical Society All rights reserved.

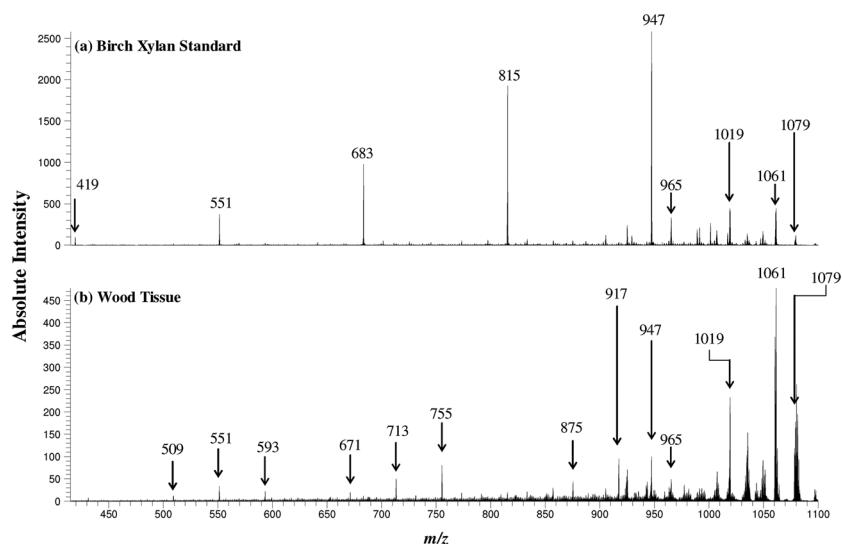


Figure 40. MS2 spectra from Birch Xylan standard (a) and intact wood tissue (b) of m/z 1079, identified as a linear xylan, $[Xyl_7 H_2O + Na]^+$. Both spectra show abundant fragment ions at m/z 1019 and 947, due to cross ring and glycosidic bond cleavages, respectively. The spectrum from wood tissue also shows abundant fragment ions at m/z 917 and 875 (NLs of 162 and 204), corresponding to a glucose and glucuronic acid residue, suggesting at least two isobaric species at the m/z value 1079. Reproduced with the permission of K.A. Lunsford et al. Copyright 2011 by American Chemical Society All rights reserved.

Moreover, the authors measured the intact wood tissue on a quarter section of the *Populus* stem by MALDI-LTQ-MS imaging. This was achieved to ensure that all tissue types were examined in the quarter section. The MS image of the whole tissue shows the ion signal $[Glc_7-H_2O + Na]^+$ ion at m/z 1319.

In addition, the optical image of the transverse section of young *Populus* stem is also shown. Please note, that the white dotted line in the optical images illustrates the region of MS analysis. The pith is positioned in the centre of the stem and is composed of thin non-

lignified, primary-walled cells. The secondary xylem extends from the pith to the vascular cambium and is composed of living ray parenchyma, nonliving lignified fibre and vessel cells with thickened secondary walls. The vascular cambium is a thin layer ($\sim 50 \mu\text{m}$) of living tissue between the secondary xylem and secondary phloem. The secondary phloem is composed of living primary walled parenchyma and companion cells as well as dead sieve and fibre cells with thickened secondary walls. In contrast to the sieve cells, the fibre walls are lignified (Fig. 42).^[169]

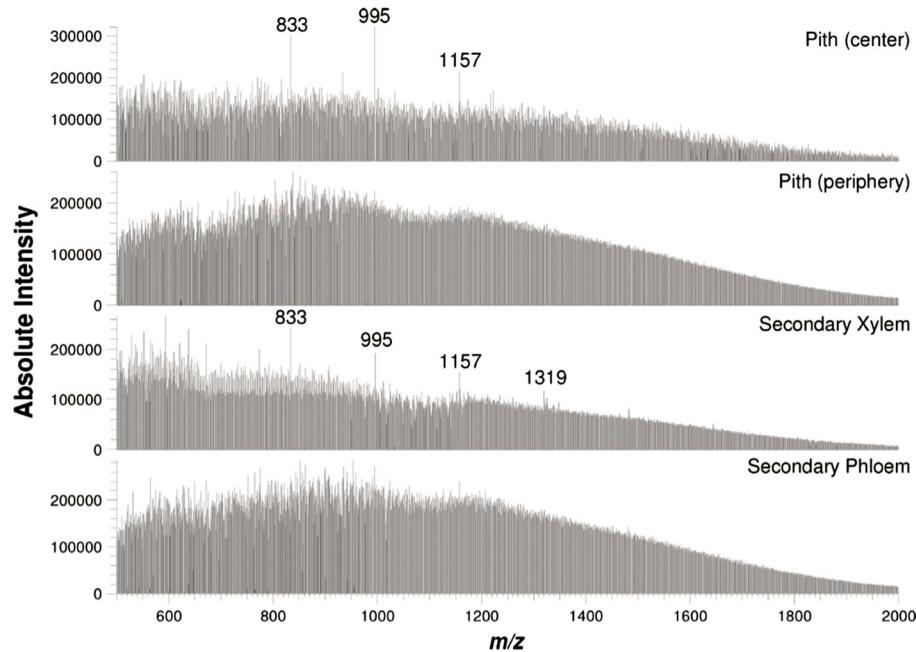


Figure 41. Comparison of an average of 25 spectra averages in four different regions of wood tissue, inner pith, around the pith, secondary xylem, and secondary phloem. Ions above the background, consistent with m/z values of MCC, are observed in the spectra of the centre pith and secondary xylem. Intense ions at nearly every m/z value are observed in the pith periphery and secondary phloem spectra. The differences in spectra are consistent with the composition of the regions: centre pith and secondary xylem are composed of mostly dead cells, and the pith periphery and secondary phloem are recomposed of mostly living cells. Reproduced with the permission of K. A. Lunsford et al. Copyright 2011 by American Chemical Society All rights reserved.

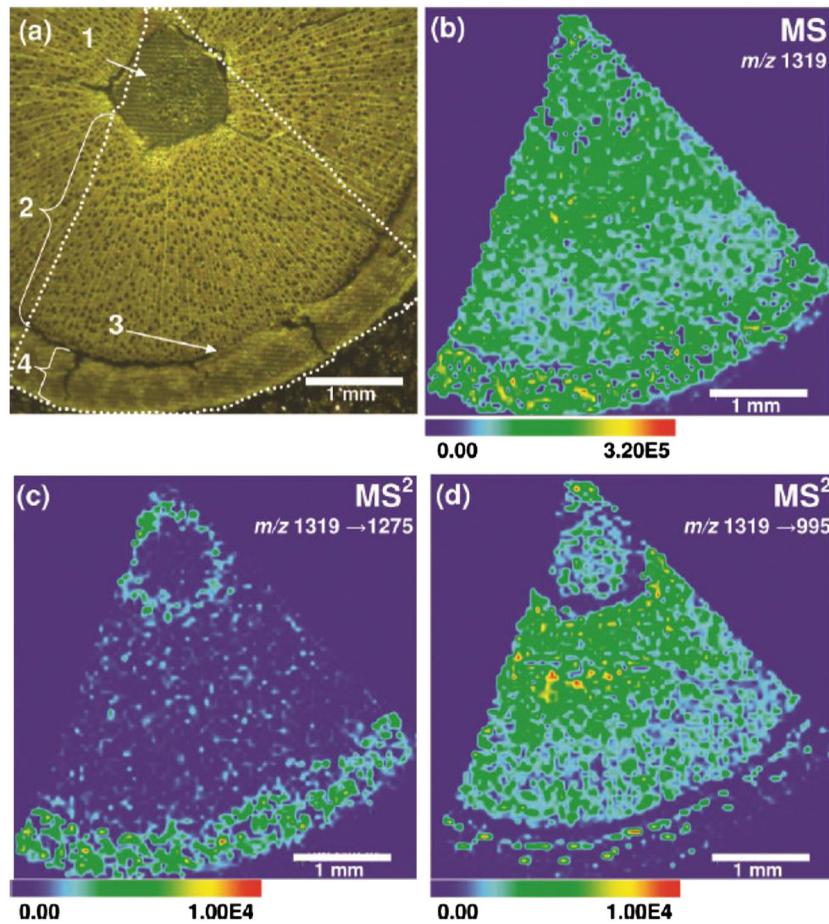


Figure 42. (a) Optical image of *Populus* wood tissue showing the pith (1), secondary xylem (2), vascular cambium (3), and secondary phloem (4). The white outline shows the area of MS and MS² analyses. (b) MS image of m/z 1319 ($[Glc_7-H_2O + Na]^+$) shows ion signal over the whole tissue. (c) The MS² image of m/z 1319 → 1275 (NL of 44) displays more abundant ion signal in the region around the pith and the secondary phloem, which is consistent with tissue compositions. (d) MS² image of m/z 1319 → 995, resulting from two NLs of 162, shows localization in the secondary xylem, closest to the pith, and reduced ion signal in the centre of the pith, vascular cambium and secondary phloem. Reproduced with the permission of K. A. Lunsford *et al.* Copyright 2011 by American Chemical Society All rights reserved.

The MS² image of m/z 1319 → 1275 (−44 u) exhibits a more abundant ion signal in the region around the pith and the secondary phloem, which is consistent with tissue compositions. The MS² image of m/z 1319 → 995, resulting from two NLs of 162 u, exhibit localization in the secondary xylem, closest to the pith, and reduced ion signal in the centre of the pith, vascular cambium and secondary phloem.^[169]

Although a complete chemical composition analysis of all the regions of wood tissue was not attempted by these authors, their experiments showed the viability of MALDI-MSⁿ for the imaging of lignocellulosic tissue. They concluded that MS alone was incapable of providing accurate spatial distributions of different ions at a single m/z . Instead, the necessity and advantages of MS/MS analyses of wood tissue were evident after comparing MS² images with the MS images.^[169]

MALDI-LTQ-MSI imaging a rapid and facile method for quick determination of the composition of intact lignin in *Eucalyptus*

Recently, Araújo *et al.* have used MALDI imaging MS (IMS) to study the native lignin composition of *Eucalyptus*.^[184] The authors used hand-cut sections of stems of two *Eucalyptus* species that were covered with silica and analysed by MALDI-IMS. In these experiments, the authors used imaging data and the MALDI spectra were

acquired with a linear trap quadrupole-XL (MALDI-LTQ-XL) linked to a solid imaging camera.^[184]

The MADLI-LTQ-IMS results reported by Araújo *et al.* and those obtained by UPLC-ESI-MS and MS/MS as described by Kiyota *et al.* were in agreement with each other in characterizing the main soluble lignin monomers and oligomer structures that were present in the sugar cane.^[108]

Accordingly, the authors were able to trace lignin monomer, dimers and trimers distribution in the *Eucalyptus* sections by means of CID-MS/MS through the measurements of multiple-reaction monitoring scans mode of the transitions of the m/z precursor ions → m/z product ions and comparing the results with their MS/MS library data.^[91,95,108] In addition, by availing of modern software imaging technique, the authors completed a relative quantification of the lignin components (Fig. 43).^[184]

With such advances, Araújo *et al.* used MALDI-LTQ-MSⁿ to image 22 out of the 24 compounds described by Kiyota *et al.*, which embodied the main soluble lignin monomers and dimers and trimers in cane sugar.^[108] In addition, these analyses allowed the determination of their localization in the cross sections. (Table 5)

It is of note that the authors obtained their images for each of the compounds identified in each stem cuts, using ImageJ software (National Institutes of Health, USA – Open Source) on grey scale images.^[185] The quantifying software that assigned the different

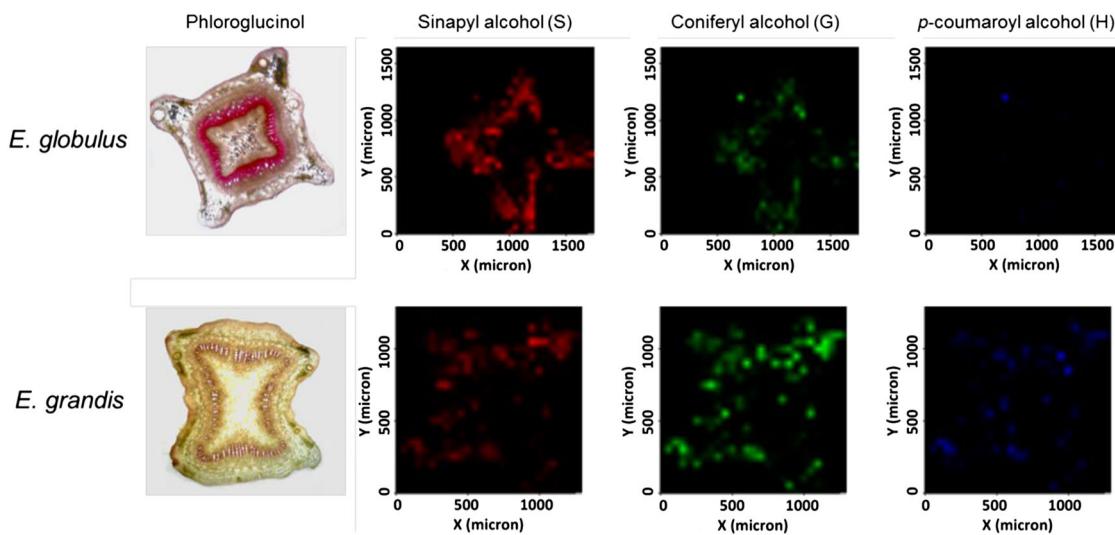


Figure 43. Phloroglucinol staining and mapping of sinapyl (m/z 209), coniferyl (m/z 179), and p-coumaroyl (m/z 149) alcohols in freshly hand-cut sections of Eucalyptus stems. Different colours were selected for each unit for illustrative reasons. For relative quantification, the figures were converted to grey images. Reproduced with permission from P. Araújo et al. Copyright by 2014 American Chemical Society.

multiple-reaction monitoring values for each compound image resulted from the sum of the intensity of each pixel. This sum was then divided by the sum of all the values, which corresponded to all the other compounds obtained in the same cut. Therefore, this elegant analysis allowed a relative novel quantification method for each compound in each cut.^[185] This once more points to the important fact that we can find small LOs in the intact stem of eucalyptus.

A striking point made by the authors was that neither coniferyl nor sinapyl aldehydes could be detected. However, Araújo et al. indicated that the presence of the 8-O-4' (or β -O-4') linkages were predominant in the lignin structure (Table 5). The authors also indicated the presence of a higher amount of S at m/z 209 residues over those samples attributed to G at m/z 179 and H at m/z 149.^[184] Although the authors indicated that the H unit image was less intense

Table 5. Relative pixel quantification of oligomers in *E. globulus* and *E. grandis* using the ImageJ Software. Reproduced with permission from P. Araújo et al. Copyright by 2014 American Chemical Society

Parent ion (m/z)	Product ions (m/z)	Structure	<i>E. globulus</i>	<i>E. grandis</i>
149	131, 103	<i>p</i> -coumaryl alcohol (H)	0.024 ± 0.010	0.010 ± 0.000
179	164, 161	coniferyl alcohol (G)	0.068 ± 0.010	0.079 ± 0.010
209	191, 176, 161	sinapyl alcohol (S)	0.088 ± 0.020	0.064 ± 0.010
297	279, 191, 121 267, 149, 121	H(8-5)H H(8-8)H	0.066 ± 0.020	0.029 ± 0.000
327	313, 309, 191	G(8-5)H	0.056 ± 0.020	0.042 ± 0.000
357	339, 221, 191 341, 151	G(8-5)G G(8-8)G	0.048 ± 0.010	0.075 ± 0.010
375	327, 195, 179	G(8-O-4)G	0.042 ± 0.010	0.050 ± 0.010
387	369, 221, 191 357, 203, 181	S(8-5)G S(8-8)G	0.065 ± 0.010	0.098 ± 0.010
405	375, 225, 179	S(8-O-4)G	0.048 ± 0.020	0.046 ± 0.000
417	181, 166	S(8-8)S	0.064 ± 0.000	0.100 ± 0.010
435	225, 209, 195	S(8-O-4)S	0.060 ± 0.000	0.037 ± 0.010
493	313, 297, 191	H(8-5)H(8-O-4)G	0.090 ± 0.020	0.106 ± 0.010
535	517, 505, 221	G(8-5)G(8-5)G	0.034 ± 0.000	0.054 ± 0.000
553	357, 339, 195	G(8-O-4)G(8-5)G	0.046 ± 0.010	0.060 ± 0.000
571	391, 375, 179	G(8-O-4)G(8-O-4)G	0.025 ± 0.010	0.032 ± 0.010
583	387, 369, 195 357, 339, 225	G(8-O-4)S(8-5)G S(8-O-4)G(8-5)G	0.044 ± 0.030	0.054 ± 0.010
601	583, 209, 195	G(8-O-4)G(8-O-4)S	0.028 ± 0.000	0.033 ± 0.000
613	387, 369, 225 417, 195, 181	S(8-O-4)S(8-5)G G(8-O-4)S(8-8)S	0.025 ± 0.000	0.035 ± 0.000
631	435, 421, 209	S(8-O-4)G(8-O-4)S	0.025 ± 0.000	0.033 ± 0.010
643	417, 225, 181	S(8-O-4)S(8-8)S	0.015 ± 0.000	0.030 ± 0.000
661	435, 225, 209	S(8-O-4)S(8-O-4)S	0.026 ± 0.010	0.042 ± 0.000
851	627, 433, 417	S(8-O-4)S(8-8)S(8-O-4)S	0.014 ± 0.020	0.045 ± 0.000

than what were observed for the S and G unit images, they also showed that the trimer H(8-5')H(8-O-4')G gave one of the highest relative intensity images (Table 5).

In addition, the authors were able to show that the normalized data of the S/G ratio obtained for *E. globulus* (1.37 ± 0.05) was higher than that observed for *E. grandis* (1.16 ± 0.13).^[184]

Nevertheless, a very important point worthy of mention is that when a comparison is made between the quantitative results to determine the S/G ratio obtained by the Py-GC/MS and nitrobenzene oxidation methods of similar *Eucalyptus spp.*, data from hybrids indicated a huge overestimation of the S/G ratio when compared with the MALDI-LTQ-MSⁿ imaging method.^[184]

It is significant that the authors indicated that normal matrixes routinely used in MALDI-imaging analysis were found to be unsatisfactory for the analysis of plant materials; however, they indicated that these could be efficiently substituted by thin-layer chromatography grade silica.

In conclusion, the authors indicated that the convenience of sample preparation required for and efficiency of MALDI-LTQ-MSⁿ imaging offer advantages over the complex and laborious MS methods that have been used in studies of lignin composition.^[184]

Second-generation types of VRLs analysed by MALDI-TOF-MS

Richel and coworkers evaluated MALDI-TOF-MS (+ ion mode) for the structural elucidation and molecular weight assessment of some extracted herbaceous types of lignin derived from *Miscanthus (Miscanthus x giganteus)* and Switchgrass (*Panicum Virgatum L.*).^[186] Three selected lignin fractions (FAL107, FAL80 and AL) were extracted from *Miscanthus x giganteus* under alkaline or acid conditions according to a previously published protocol.^[187] Lignin isolated after a formic acid/acetic acid pretreatment at 107 °C (FAL107) contained mainly β-5' linkages as evidenced by NMR. The AL lignin fraction obtained after aqueous ammonia soaking contained mainly β-O-4' and residual phenylcoumaran (β-5') substructures. FAL80 was obtained using a milder formic/acetic acid protocol at 80 °C and contained both β-O-4' and β-5' moieties. The lignin fractions SG-107 and SG-AL were extracted from Switchgrass feedstock under formic acid/acetic acid pretreatment (107 °C for 3 h) and aqueous ammonia conditions (60 °C for 6 h), respectively.^[187]

The authors showed MALDI-TOF-MS (+ ion mode) to be best recorded using α-cyano-4-hydroxycinnamic acid (CHCA), in combination with α-cyclodextrin. This permitted the efficient ionization of the poorly soluble lignin materials and suppressed the matrix-related background ions. Effectively, Richel and coworkers were able to analyse low-mass fragments ions in the mass range at m/z 100–600. The MS data obtained were considered as a diagnostic fingerprint of the starting VRLs and allowed quantitative identification of the oligomers containing principal interunit linkages between phenylpropanoid units. The authors were also able to provide useful information on the oligomeric distribution of different types of herbaceous types of lignin containing mainly guaiacyl and syringyl moieties.

The MALDI-TOF-MS of FAL 107 lignin, which was isolated from *Miscanthus* after a formic acid/acetic acid pretreatment at 107 °C (FAL107), was obtained using either 2,5-DHB in combination with NaCl (Fig. 44A) or CHCA in 30% aq. acetonitrile with 0.1% TFA (Fig. 44B). (Matrix cluster signals appeared in this low mass range and probably interfered with the lignin fingerprints.)^[187]

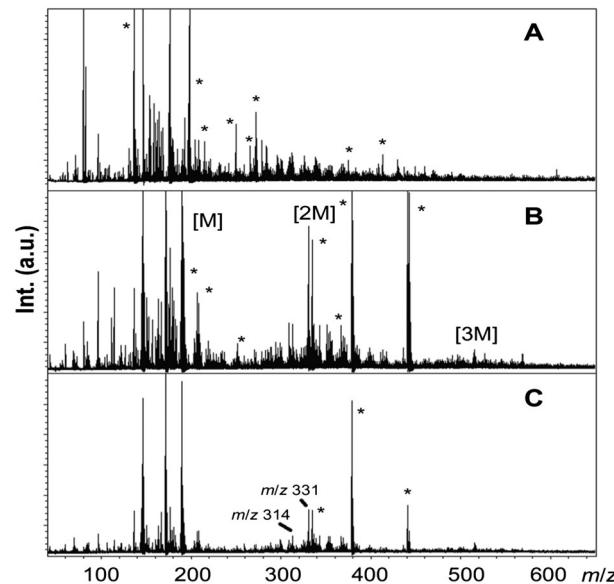


Figure 44. Reflector positive ion mode MALDI-TOF-MS spectrum of FAL107 with DHB (A), CHCA (B) and α -CD/CHCA (C) [monomers (denoted [M]), dimers ([2M]) and trimers ([3M])]. Reproduced with the permission of A. Richel et al. (2012). Copyright 2012 by Libertas Academica Ltd. All rights reserved.

In experiments designed to extend the range of applicability of MALDI-TOF-MS for FAL107, it was found that CHCA encapsulated in a cyclodextrin cavity enabled improvements of the aforementioned conditions by giving a cleaner spectrum with less interference from matrix fragments (Fig. 44C). In Fig. 44, [M] was assigned to the monomer, [2M] to dimers and [3M] to trimers.

The authors observed that the hyperfine structure of this dimeric distribution, within the mass range m/z 250 to 400, offered special diagnostic information concerning the occurrence of specific linkages between two phenylpropane units. It is noteworthy that the elucidation of the dimeric structure of FAL107 corroborated the work of Banoub and Delmas using APCI-MS and APPI-MS of acid-extracted lignin from wheat straw.^[9,88] The authors also found a main ion at m/z 331, which had a well-resolved isotopic distribution that related to the product ion at m/z 315. This was assigned to the protonated $[C_{18}H_{19}O_6]^+$ dimer, constructed of two G units linked through a β-5' moiety. This observation allowed the authors to propose that β-5' linkages predominate in dilignols of *Miscanthus* pretreated with acid. Richel and coworkers were able to use this technique to discriminate between the samples extracted using ammonia soaking and those extracted by formic/acetic acid conditions. In agreement with earlier studies, the authors showed that there were discrepancies between molar masses, above 1000 u, when they were estimated using on-line high-performance size-exclusion chromatography coupled to differential refractometry (HPSEC-RI) in conjunction with GPC.^[187]

MALDI-TOF-MS analysis of the PML alfa grass (*Stipa tenacissima L.*) lignins

This study was chosen to exemplify the type of chemical manipulation that the lignin can be submitted to, prior of initiation of structural elucidation.^[188] The authors treated the alfa grass lignin (VRL)

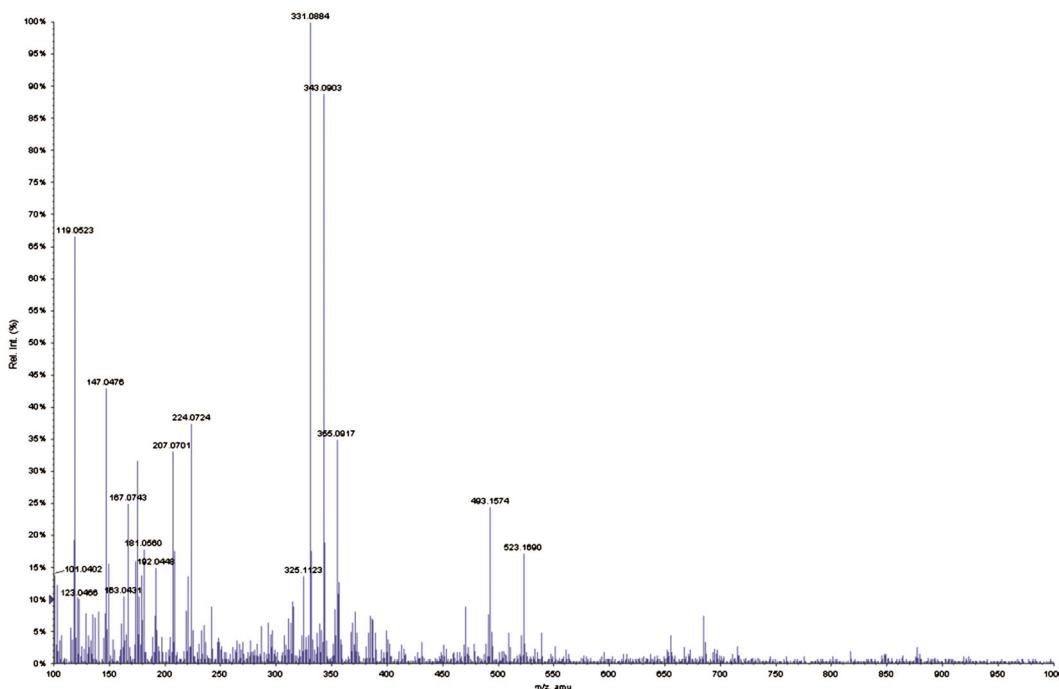


Figure 45. MALDI-QqTOF-MS of the CIMV extracted alfa grass lignin.

obtained by the CIMV process,^[147–149] by a complicated series of dissolution/precipitation and alkaline hydrolysis manipulations, which were accomplished as follows. The following is included to underscore the tedious process of converting VRL into an altered PML. The VRL was suspended in 1,4-dioxane; the insoluble fraction was removed by filtration and the solvent removed by evaporation. The residue was then dissolved in diethylether to separate insoluble lignin from soluble low molar mass compounds. Finally, the precipitate was filtered and dried to yield purified lignin L_p . In addition, the authors hydrolysed the VRL in aqueous 3N NaOH for 4 h at 80 °C followed by precipitation in acidic aqueous medium to afford, the hydrolysed lignin (L_h). The samples were preliminarily analysed by size exclusion chromatography to determine their molar masses. VRL and L_h could not be analysed without further treatment because of their low solubility in THF. For this reason, the corresponding acetylated samples (VRL-Ac, L_p -Ac and L_h -Ac) were prepared. Following conventional analytical analysis of these samples using FTIR, ^1H NMR, ^{13}C - ^1H 2D-HSQC NMR and ^{31}P -NMR spectroscopies, the authors concluded that their alfa purified lignin was of the SGH type with low H content (*ca.* 12%) and contained β -O-4' ether bonds as the major interunit linkage. These results are very similar to those of del Rio *et al.* for wheat straw lignin.^[66] According to these authors, the structures obtained for these different fractions were analogous to those of the native polymer in the plant.^[189]

Delmas Jr., Delmas and Banoub have analysed the same alfa grass lignin (VRL) described earlier (of which the only modification was to wash extensively in water) by ESI-QqTOF-MS, APPI-QqTOF-MS and MALDI-QqTOF-MS (+ ion mode) and CID-MS/MS. These series of spectra obtained were comparable with the MS and MS/MS spectra of the obtained from wheat straw (VRL).^[150] However, they were not identical because the plant origin was different. In addition, the MS/MS analyses of some of the major ions formed gave CID fragmentations that were identical to some fragment ions of the wheat straw lignin (VRL). This finding seems to indicate that the VRL alfa grass lignins may be composed preferentially by series

of repeating oligomeric phenylcoumaran units, as is the case with the wheat straw lignin.

Interestingly, MALDI-QqTOF-MS (+ ion mode) of the VRL alfa grass lignin (*Stipa tenacissima* L.) obtained by the CIMV process indicated the complete absence of protonated molecules of syringyl aldehyde at m/z 209 and syringyl alcohol at m/z 211 and any derived dimers or trimers and derived units (Fig. 45).

These results seem to confirm that additional chemical manipulations of the VRL alfa grass lignin give specious results. The results presented in this section seem in general to make a distinction between VRL and PML analytes obtained by series of different analytical methods.

Conclusion

In this review article, we have introduced two new terms to distinguish lignins in accordance with the processes used to extract them from crops. Accordingly, we have named VRLs as those that have been released from the lignocellulosic cellular material by a simple one-step hydrolysis process. They could be considered as unmodified lignins with structures similar to native lignins present in the glycolignin network. The term PMLs is used for lignins obtained by other chemical or mechanical means followed by other steps including purification. Such multiple steps of extraction and purification appear to substantially affect the structure and properties of the lignins obtained.

Forthrightly, it would be over simplistic at this juncture to speculate if the structure of the oligomers composing the VRL could be a true representation of the natural lignin.

Analytical data currently available obtained from samples extracted by established methods are likely to be insufficient to be confident in assuring that the structure elucidated for VRL represents the actual structure present in the glycolignin network than for PML. However, the structures determined for VRL samples

obtained directly from vegetable matter, after several simple steps such as washing with water and drying, should represent LOs with the closest resemblance to native lignin biomolecules.^[66,189]

To appreciate the enormous heterogeneity and complexity of the ionization of native lignin, the reader should consult the work of Lunsford, Peter and Yost^[170] (refer to section Study of the Removal of the Lignin from Cellulose and Hemicellulose by Imaging with Direct Matrix-assisted Laser Desorption/Ionization Tandem Mass Spectrometry). They used direct imaging by MALDI-TOF-MS of cellulose and hemicellulose in intact wood tissue, in the presence of assorted LOs to show that there is a multitude of real oligomeric signal ions, which cannot be assigned as background noise, but results from the presence of a pack of different LOs. Evidently, these do not represent one and/or various linear and/or branched lignin polymers. If that were to have been the case, then it would have been logical to expect fewer ions for such intact hypothetical lignin polymer chain to have been present between the cellulose polymers as viewed by MALDI-TOF-MS direct imaging. Even after base treatment of the analysed tissue, when the oligomeric signals had been removed from the wood tissue, their presence was still detected.

Similar presence of a multitude of oligomeric ions was also observed for APPI-QqTOF-MS, ESI-QqTOF-MS and MALDI-TOF/TOF-MS analyses of various VRL extracted from pine and oak lignins. Therefore, the presence of such complex series of oligomeric ions provide excellent evidence to support the similarity between VRL and the complex series of oligomeric ions observed by direct imaging using MALDI-TOF-MS, of cellulose and hemicellulose in intact wood tissue.

In addition, Araújo *et al.* have used MALDI-IMS to study the native lignin composition of Eucalyptus.^[184] The authors were capable to image 22 out of the 24 compounds described by Kiyota *et al.*, which embodied the main soluble lignin monomers and dimers and trimers in cane sugar.^[108,184] This indicated that the native lignins oligomers are indeed present as an enormous conglomerate of simple monomer, dimer, trimer and oligomers.^[108,184]

Accordingly, on the basis of the results of Lunsford *et al.*, Araújo *et al.*^[169,184] and our ESI-MS and APPI-MS of the CIMV straw, oak and pine lignins, we propose a new paradigm. Intact natural lignins present in the lignocellulosic biomass are not actually either one and/or series of similar biomolecules, like an individual cellulose fibre; instead, they are composed of a series of different length linear related biosynthesized oligomers. These oligomers may well be formed either from homoligomers repeating units and/or could be heteroligomers formed by mixed units. These have never been completely described in their natural unprocessed form. In addition, it is proposed that lignins present in the lignocellulosic biomass are attached covalently in a criss-cross manner to both cellulose and hemicellulose fibres, forming a glycolignin network.

Because of the complementary nature of APPI and ESI, when analysing VRL and PML by MS, it is recommended that both sources should be used to obtain best structural indices for lignin structures and compositions. In addition, the use of either ESI, APPI and/or MALDI ionizations with either a high-field FT-ICR-MS and/or an orbitrap analysers is advised as it should provide unprecedented resolution and should enable the allocation of more precise structural information on the lignin molecules.^[71]

We propose that obtaining purified lignin by any process that is outlined in the supplementary part of this review is unachievable, especially when using GPC as a tool for purification. It is the view of the authors that a purified lignin has never been isolated. On the other hand, what lignin researchers most

probably have meant, when they attribute a lignin as being pure, is that they have obtained a fractionated heterogeneous mixture of LOs. Such were obtained by chemical modifications of the native lignin material.

In the 21st century and in the age of lignomics, the authors feel that lignin scientists should not be afraid to seek solutions away from their *modus operandi* established protocols. Yet, the field should rely more on adapting the state-of-the-art new technologies for the better structural elucidation of VRL rather than PML.

It is particular significance in this field to note that no high molecular weight lignins between 10 000 to 100 000 Da have not yet been reported using various mass spectrometric analyses including APPI-MS, ESI-MS and MALDI-MS. This is not due to any limitations of the mass ranges of these modern MS instruments but to the probability that such large lignin molecules have either not been isolated and or do not exist.

In this review, the new concept of top-down lignomic strategy employing MS/MS is introduced. The *de novo* top-down lignomic strategy is described as the identification by MS of the VRL-type lignins. Moreover, the top-down approach enables direct measurement of the intact mass of the heterogeneous LOs, as well as fragment ion information relating to the original oligomer lignin sequences. The structural characterization and sequence of the various oligomers present in VRL are performed by MS/MS. This strategy allows the identification of all the components of the intact non-modified virgin lignins (VRL) being analysed, which can also be considered as crude lignin. In addition to our VRL sequencing of the CIMV wheat straw, it is noteworthy that the following authors, Moreel *et al.* (APCI-MS/MS-based Sequencing of Oligolignols), Kyota *et al.* (Ultrahigh-performance Liquid Chromatography ESI-MS and MS/MS Analysis of Sugar Cane Soluble Lignin), Hauper *et al.* (APCI-MS/MS and ESI-MS/MS Characterization of PML and Model Lignin Compounds) and Owen *et al.* (Characterization of VRL Degradation Products by HPLC-ESI-MSⁿ) have already studied *de novo* top-down lignomic approach but without naming it as such.

We sincerely hope that the reader can comprehend the intricacy of the different types of lignin and the convolution resulting from the chemistry directed towards their structural elucidation.

In summary, the complexity of the structural determination of lignins has been described, and the search for the native structure of a pure lignin will continue. We do hope that the reader of this review is able to capture the excitement of recent achievements in MS-based lignomics and how they might play a role towards a more profound understand of the intrigues of lignin structure elucidation.

Acknowledgements

The authors would like to acknowledge an anonymous referee for his courage editing our draft and who was extremely generous with his constructive comments and corrections that have contributed to this new edition of this review. The authors would like to thank the Editors of JMS Professors Richard Caprioli and Pierre Chaurand for their resilient support and encouragement.

References

- [1] J. Bidlack, M. Malone, R. Benson. Molecular structure and component integration of secondary cell walls in plants. *Proc. Oklahoma Acad. Sci.* **1992**, 72, 51.

- [2] R. D. Perlack, L. L. Wright, A. F. Turhollow, R. L. Graham, B. J. Stokes, D. C. Erbach. Biomass as a feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion ton annual supply. *The US Department of Agriculture (USDA) and Energy (DOE) report*, **2005**. http://www1.eere.energy.gov/biomass/pdfs/final_billionton_vision_report2.pdf (accessed on April 2005).
- [3] K. Esau. Cell Wall. In *Plant anatomy*, Wiley-Blackwell (ed). John Wiley and Sons: New York, **1977**, 43–60.
- [4] K. Keegstra, K. W. Talmadge, W. D. Bauer, P. Albersheim. The structure of plant cell walls. III. A model of the walls of suspension-cultured sycamore cells based on the interconnections of the macromolecular components. *Plant Physiol.* **1973**, *51*, 188.
- [5] (a) D. W. Jones. *Cellulose and Cellulose Derivatives*, vol. 5, part 4, *High Polymers*, M. N. Bikales, L. Segal (Eds). Wiley Interscience: New York, **1971**, 117–121; (b) O. Ellefsen, B. A. Tonnesen. *Cellulose and Cellulose Derivatives*, M. N. Bikales, L. Segal (Eds). Wiley Interscience: New York, **1971**, 151–180; (c) B. A. Tonnesen, O. Ellefsen. *Cellulose and Cellulose Derivatives*, M. N. Bikales, L. Segal (Eds). Wiley Interscience: New York, **1971**, 265–304; (d) R. H. Atalla. *Adv. Chem. Ser.* **1979**, *181*, 55.
- [6] R. D. Preston. Polysaccharides conformation and cell wall function. *Ann. Rev. Plant. Physiol.* **1979**, *30*, 55.
- [7] K. V. Sarkonen, C. H. Ludwig. *Lignin: Occurrence, Formation, Structure and Reactions*, K. V. Sarkonen, C. H. Ludwig (Eds). Wiley-Interscience: New York, **1971**, 916–931.
- [8] W. O. S. Doherty, P. Mousavioun, C. M. Fellows. Value-adding to cellulosic ethanol: lignin polymers. *Ind. Crop Prod.* **2011**, *33*(2), 259.
- [9] J. H. Banoub, M. Delmas. Structural elucidation of the wheat straw lignin polymer by atmospheric pressure chemical ionization tandem mass spectrometry and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J. Mass Spectrom.* **2003**, *38*, 900 and references therein.
- [10] W. D. Bauer, K. W. Talmadge, K. Keegstra, P. Albersheim. The structure of plant cell walls. II. The hemicellulose of the walls of suspension-cultured Sycamore cells. *Plant Physiol.* **1973**, *51*, 174.
- [11] G. Brunow. Methods to reveal the structure of lignin. In *Lignin, Humic Substances and Coal, Biopolymers*, Vol. 1, M. Hofrichter, A. Steinbüchel (Eds). Wiley Blackwell: Michigan, USA, **2001**, 89–116.
- [12] A. Caño-Delgado, S. Penfield, C. Smith, M. Catley, M. Bevan. Reduced cellulose synthesis invokes lignifications and defense responses in *Arabidopsis thaliana*. *Plant J.* **2003**, *34*(3), 351.
- [13] M. Tronchet, C. Balaguer, T. Kroi, L. Jouanin, D. Roby. Cinnamyl alcohol dehydrogenases-C and D, key enzymes in lignin biosynthesis, play an essential role in disease resistance in *Arabidopsis*. *Mol. Plant Pathol.* **2010**, *11*, 83.
- [14] (a) K. Freudenberg. Biosynthesis and constitution of lignin. *Nature* **1959**, *183*, 1152; (b) K. Freudenberg, H. Hess. *Liebigs Ann. Chem.* **1926**, *448*, 126; (c) K. Freudenberg, H. Richtzenhain. *Liebigs Ann. Chem.* **1942**, *552*, 126.
- [15] Y.-R. Chen, S. Sarkonen. Macromolecule replication during lignin biosynthesis. *Phytochemistry* **2010**, *71*, 453.
- [16] C. Lapierre. Chap 2. Determining lignin structure by chemical degradations. In *Lignin and Lignans: Advances in Chemistry*, C. Heitner, D. Dimmel, J. Schmidt (Eds). CRC Press: Boca Raton, Florida, **2010**, 13–48 and references cited therein.
- [17] H. Wang, P. de Vries Frits, Y. Jin. A win-win technique of stabilizing sand dune and purifying paper mill black-liquor. *J. Environ. Sci.* **2009**, *21*, 488.
- [18] W. Boerjan, J. Ralph, M. Baucher. Lignin Biosynthesis. *Annu. Rev. Plant Biol.* **2003**, *54*, 519.
- [19] D. Harris, S. DeBolt. Synthesis, regulation and utilization of lignocellulosic biomass. *Plant Biotechnol. J.* **2010**, *8*, 244.
- [20] M. M. Campbell, R. R. Sederoff. Variation in lignin content and composition. Mechanisms of control and implications for genetic improvement of plants. *Plant Physiol.* **1996**, *110*, 3.
- [21] W. G. Glasser. Lignin. In *Pulp and Paper. Chemistry and Chemical Technology*, Vol. 1, 3rd ed, J. P. Casey (Ed.). John Wiley & Sons: New York, **1980**, 39–111.
- [22] L. B. Davin, N. G. Lewis. Lignin primary structures and dirigent sites. *Curr. Opin. Biotechnol.* **2005**, *16*, 407 and references therein.
- [23] M. Dolk, F. Pla, J. F. Yan, J. L. McCarthy. Lignin. 22. Macromolecular characteristics of alkali lignin from western hemlock wood. *Macromolecules* **1986**, *19*, 1464.
- [24] J. Mlynar, S. Sarkonen. Renaissance in ultracentrifugal sedimentation equilibrium calibrations of size exclusion chromatographic elution profiles. In *Strategies in Size Exclusion Chromatography*, Vol. 635, M. Potschka, P. L. Dubin (Eds). ACS Symposium Series, American Chemical Society: Washington, **1996**, 379–400.
- [25] K. Ruel, F. Barnoud, D. A. I. Goring. Lamellation in the S2-layer of softwood tracheids as demonstrated by scanning transmission electron microscopy. *Wood Sci. Technol.* **1978**, *12*, 287.
- [26] T. Hatakeyama, T. Yoshida, H. Hatakeyama. Thermal and viscoelastic properties of *in situ* lignin. *Tenth International Symposium on Wood and Pulping Chemistry*, Vol. 1, **1999**, 478.
- [27] C. Crestini, F. Melone, M. Sette, R. Saladino. Milled Wood Lignin: A linear oligomer. *Biomacromolecules* **2011**, *12*, 3928.
- [28] L. B. Davin, M. Jourdes, A. M. Patten, K.-W. Kim, D. G. Vassão, N. G. Lewis. Dissection of lignin macromolecular configuration and assembly: comparison to related biochemical processes in allyl/propenyl phenol and lignan biosynthesis. *Nat. Prod. Rep.* **2008**, *25*, 1015.
- [29] L. B. Davin, A. M. Patten, M. Jourdes, N. G. Lewis. Lignins: a twenty-first century challenge. In *Biomass Recalcitrance – Deconstructing the Plant Cell Wall for Bioenergy*, M. E. Himmel (Ed.). Blackwell: Oxford, **2008**, 213–305.
- [30] N. G. Lewis, L. B. Davin, S. Sarkonen. Lignin and lignan biosynthesis: distinctions and reconciliations. *ACS Symposium Series*, Vol. 697, **1998**, 1.
- [31] G. P. Moss. Nomenclature of lignans and neolignans (IUPAC Recommendations 2000). *Pure Appl. Chem.* **2000**, *72*, 1493.
- [32] J. R. Weng, H. H. Ko, T. L. Yeh, H. C. Lin, C. N. Lin. Two new arylnaphthalide lignans and anti-platelet constituents from *Justicia procumbens*. *Arch. Pharm. Pharm. Med. Chem.* **2004**, *337*, 207.
- [33] H. Y. Hsu. *Treating Cancer with Chinese Herbs*. Oriental Healing Arts Institute: Los Angeles, **1982**, 238–250.
- [34] C. C. Chen, W. C. Hsin, F. N. Ko, Y. L. Huang, J. C. Ou, C. M. Teng. Antiplatelet arylnaphthalide lignans from *Justicia procumbens*. *J. Nat. Prod.* **1996**, *59*, 1149.
- [35] C. C. Chen, W. C. Hsin, Y. L. Huang. Six new diarylbutane lignans from *Justicia procumbens*. *J. Nat. Prod.* **1998**, *61*, 227.
- [36] N. Fukamiya, K. H. Lee. Antitumor agents. 81. Justicidin-A and Diphyllin, two cytotoxic principles from *Justicia procumbens*. *J. Nat. Prod.* **1986**, *49*, 348.
- [37] T. Umezawa. Phylogenetic distribution of lignan producing plants. *Wood Res.* **2003**, *90*, 27.
- [38] S. Suzuki, T. Umezawa. Biosynthesis of lignans and norlignans. *J. Wood Sci.* **2007**, *53*, 273.
- [39] I. E. J. Milder, I. C. W. Arts, B. van de Putte, D. P. Venema, P. C. H. Hollman. Lignan contents of Dutch plant foods: a database including lariciresinol, pinoresinol, secoisolariciresinol and matairesinol. *Br. J. Nutr.* **2005**, *93*, 393.
- [40] A. I. Smeds, P. C. Eklund, R. E. Sjöholm, S. M. Willfor, S. Nishibe, T. Deyama, B. R. Holmbom. Quantification of a broad spectrum of lignans in cereals, oilseeds, and nuts. *J. Agric. Food Chem.* **2007**, *55*, 1337.
- [41] C. Frankenstein, U. Schmitt, G. Koch. Topochemical studies on modified lignin distribution in the xylem of poplar (*Populus spp.*) after wounding. *Ann. Bot.* **2006**, *97*(2), 195.
- [42] A. George, K. Tran, T. J. Morgan, P. I. Benke, C. Berreco, E. Lorente, B. C. Wu, J. D. Keasling, S. Blake, B. M. Holmes. The effect of ionic liquid cation and anion combinations on the macromolecular structure of lignins. *Green Chem.* **2011**, *1*, 3375.
- [43] R. J. A. Gosselink. Lignin as a renewable aromatic resource for the chemical industry. PhD Thesis, Wageningen University, Wageningen, NL, **2011**.
- [44] R. Vanholme, B. Demedts, K. Morreel, J. Ralph, W. Boerjan. Update on lignin biosynthesis and structure lignin biosynthesis and structure. *Plant Physiol.* **2010**, *153*, 895.
- [45] Y. Barriere, C. Riboulet, V. Mechini, S. Maltese, M. Pichon, A. Cardinal, C. Lapierre, T. Lubberstedt, J. P. Martinat. Genetics and genomics of lignification in grass cell walls based on maize as a model species. *Genes, Genomes Genomics* **2007**, *1*(2), 133.
- [46] T. I. Eglinton, M. A. Goni, J. J. Boon, E. R. E. Van der Hage, N. Terashima, Y. Xie. Incorporation of ¹³C-labeled coniferyl alcohol into developing *Ginkgo biloba* L. lignin revealed by analytical pyrolysis and CuO oxidation in combination with isotope ratio monitoring – Gas chromatography-mass spectrometry. *Holzforschung* **2000**, *54*, 39.
- [47] S. Reale, A. Di Tullio, N. Spreti, F. De Angelis. Mass spectrometry in the biosynthetic and structural investigation of lignins. *Mass Spectrom. Rev.* **2004**, *23*, 87.

Chemical structure of native and modified lignins

- [48] K. V. Sarkanen, A. Islam, C. D. Anderson. Ozonation. In *Methods in Lignin Chemistry*, S. Y. Lin, C. W. Dence (Eds). Springer-Verlag: Berlin, **1992**, 387 and reference therein.
- [49] K. V. Sarkanen. Lignin precursors and their polymerization. In *Lignins: Occurrence, Formation and Reactions*, K. V. Sarkanen, C. W. Ludwig (Eds). Wiley Interscience: New York, **1971**, 95.
- [50] S. Reale, A. Di Tullio, N. Spreti, F. De Angelis. Mass spectrometry in the biosynthetic and structural investigation of lignins. *Mass Spectrom. Rev.* **2004**, 23, 87.
- [51] K. Syrjanen, G. Brunow. Oxidative cross coupling of *p*-hydroxycinnamic alcohols with dimeric arylglycerol β -aryl ether lignin model compounds. The effect of oxidation potentials. *J. Chem. Soc. Perkin Trans.* **1998**, 1(20), 3425.
- [52] L. B. Davin, N. G. Lewis. Dirigent proteins and dirigent sites explain the mystery of specificity of radical precursor coupling in lignan and lignin biosynthesis. *Plant Physiol.* **2000**, 123, 453.
- [53] D. R. Gang, M. A. Costa, M. Fujita, A. T. Dinkova-Kostova, H.-B. Wang, V. Burlat, W. Martin, S. Sarkanen, L. B. Davin, N. G. Lewis. Regiochemical control of monolignol radical coupling: a new paradigm for lignin and lignan biosynthesis. *Chem. Biol.* **1999**, 6, 143.
- [54] V. Burlat, M. Kwon, L. B. Davin, N. G. Lewis. Dirigent proteins and dirigent sites in lignifying tissues. *Phytochemistry* **2001**, 57, 883.
- [55] A. Maureen Rouhi. Only facts will end the lignin war. *Chem. Eng. News* **2001**, 79, 52.
- [56] J. Ralph, G. Brunow, W. Boerjan. *Lignins*. John Wiley & Sons, Ltd: Hoboken, New Jersey, **2007**.
- [57] K. Morreel, H. Kim, F. Lu, O. Dima, T. Akiyama, R. Vanholme, C. Niculaes, G. Goeminne, D. Inzé, E. Messens, J. Ralph, W. Boerjan. Mass spectrometry-based fragmentation as an identification tool in lignomics. *Anal. Chem.* **2010**, 82, 8095.
- [58] J. H. Lora, W. G. Glasser. Recent industrial applications of lignin: a sustainable alternative to non-renewable materials. *J. Polym. Environ.* **2002**, 10, 39.
- [59] D. Stewart. Lignin as a base material for materials applications: chemistry, application and economics. *Ind. Crops Prod.* **2008**, 27, 202.
- [60] R. T. Marthers. How well can renewable resources mimic commodity monomers and polymers? *J. Polym. Sci. Part A* **2012**, 50(1), 1.
- [61] J. Zakzeski, P. C. A. Bruijnincx, A. L. Jongerius, B. M. Weckhuysen. The catalytic valorization of lignin for the production of renewable chemicals. *Chem. Rev.* **2010**, 110, 3552.
- [62] J. MacKay, D. R. Dimmel, J. J. Boon. Pyrolysis mass spectral characterization of wood from CAD-deficient pine. *J. Wood Chem. Tech.* **2001**, 21, 19.
- [63] I. Fachuang, J. Ralph. Derivatization followed by reductive cleavage (DFRC Method), a new method for lignin analysis: protocol for analysis of DFRC monomers. *J. Agric. Food Chem.* **1997**, 45, 2590.
- [64] J. O. Metzger, C. Bicke, O. Faix, W. Tuszyński, R. Angermann, M. Karas, K. Strupat. Matrix-assisted laser desorption mass spectrometry of lignins. *Angew. Chem. Int. Ed. Engl.* **1992**, 31, 762.
- [65] F. de Angelis, R. Nicoletti, N. Spreti, F. Veri. A new *in vitro* model of lignin biosynthesis. *Angew. Chem. Int. Ed. Engl.* **1999**, 38, 1283.
- [66] J. C. del Rio, J. Rencoret, P. Prinsen, A. T. Martinez, J. Ralph, A. Gutierrez. Structure characterization of wheat straw lignin as revealed by analytical pyrolysis, 2D-NMR, and reductive cleavage methods. *J. Agric. Food Chem.* **2012**, 60, 5922.
- [67] R. Seraglia, L. Molin, I. Isak, P. Traldi. Investigation on the role of pneumatic aspects in electrospray, desorption electrospray surface ionization and surface activated chemical ionization. *Eur. J. Mass Spectrom.* (Chichester, Eng) **2012**, 18(2), 195.
- [68] T. J. Kauppila, T. Kotiaho, R. Kostainen, A. P. Bruins. Negative ion atmospheric pressure photoionization-mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2004**, 15(2), 203.
- [69] P. Kebarle, M. Pesche. On the mechanism by which the charged droplets produced by electrospray lead to gas phase ions. *Anal. Chim. Acta* **2000**, 406, 11 and references therein.
- [70] T. J. Kauppila, R. Kostainen, A. P. Bruins. Anisole, a new dopant for atmospheric pressure photoionization-mass spectrometry of low proton affinity, low ionization energy compounds. *Rapid Commun. Mass Spectrom.* **2004**, 18, 808.
- [71] W. C. Hockaday, J. M. Purcell, A. G. Marshall, J. A. Baldock, P. G. Hatcher. Electrospray and photoionization mass spectrometry for the characterization of organic matter in natural waters: a qualitative assessment. *Limnol. Oceanogr. Meth.* **2009**, 7, 81 and references therein.
- [72] S. S. Cai, J. A. Syage. Comparison of atmospheric pressure photoionization, atmospheric pressure chemical ionization, and electrospray ionization mass spectrometry for analysis of lipids. *Anal. Chem.* **2006**, 78(4), 1191.
- [73] J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P. Schatz, J. Marita, R. Hatfield, S. Ralph, J. Christensen, W. Boerjan. Lignins: natural polymers from oxidative coupling of 4-hydroxyphenyl propanoids. *Phytochemistry Rev.* **2004**, 3, 29.
- [74] D. B. Robb, M. W. Blades. State of the art in APPI for LC/MS. *Anal. Chim. Acta* **2008**, 627(1), 34.
- [75] M. Takino, K. Yamaguchi, T. Nakahara. Determination of carbamate pesticide residues in vegetables and fruits by liquid chromatography-atmospheric pressure photoionization-mass spectrometry and atmospheric pressure chemical ionization-mass spectrometry. *J. Agric. Food Chem.* **2004**, 52(4), 727.
- [76] N. Yoshioka, Y. Akiyama, T. Matsuoka, T. Mitsuhashi. Rapid determination of five post-harvest fungicides and metabolite in citrus fruits by liquid chromatography/time-of-flight mass spectrometry with atmospheric pressure photoionization. *Food Control* **2010**, 21(2), 212.
- [77] S. S. Cai, J. A. Syage, K. A. Hanold, M. P. Balogh. Ultra performance liquid chromatography-atmospheric pressure photoionization-tandem mass spectrometry for high-sensitivity and high-throughput analysis of US Environmental Protection Agency 16 Priority Pollutants Polynuclear Aromatic Hydrocarbons. *Anal. Chem.* **2009**, 81(6), 2123.
- [78] S. Roy, A. Delobel, K. Gaudin, D. Touboul, D. P. Germain, A. Baillet, P. Progron, O. Laprévote, P. Chaminade. Liquid chromatography on porous graphitic carbon with atmospheric pressure photoionization and tandem mass spectrometry for the analysis of glycosphingolipids. *J. Chromatogr. A* **2006**, 1117(2), 154.
- [79] Y. X. Cai, D. Kingery, O. McConnell, A. C. Bach. Advantages of atmospheric pressure photoionization mass spectrometry in support of drug discovery. *Rapid Commun. Mass Spectrom.* **2005**, 19, 1717.
- [80] J. L. Gomez-Ariza, T. Garcia-Barrera, F. Lorenzo. Anthocyanins profile as fingerprint of wines using atmospheric pressure photoionisation coupled to quadrupole time-of-flight mass spectrometry. *Anal. Chim. Acta* **2006**, 570, 101.
- [81] M. Hiroshi, I. Miho, Y. Shisaku, M. Hidezaku, A. Jean-Francois. Determination of polycyclic aromatic hydrocarbons in sediment by liquid chromatography atmospheric pressure photoionization mass spectrometry. *Anal. Sci.* **2004**, 20, 375.
- [82] E. A. Straube, W. Dekant, W. Volkel. Comparison of electrospray ionization, atmospheric pressure chemical ionization, and atmospheric pressure photoionization for the analysis of dinitropyrene and aminonitropyrene LC-MS/MS. *J. Am. Soc. Mass Spectrom.* **2004**, 15, 1853.
- [83] J. M. Purcell, C. L. Hendrickson, R. P. Rodgers, A. G. Marshall. Atmospheric pressure photoionization Fourier transform ion cyclotron resonance mass spectrometry for complex mixture analysis. *Anal. Chem.* **2006**, 78, 5906.
- [84] J. Zheng, S. A. Shamsi. Capillary electrochromatography coupled to atmospheric pressure photoionization mass spectrometry for methylated benzo[α]pyrene isomers. *Anal. Chem.* **2006**, 78(19), 6921.
- [85] M. Takino, S. Daishima, T. Nakahara. Determination of chloramphenicol residues in fish meats by liquid chromatography-atmospheric pressure photoionization mass spectrometry. *J. Chromatogr. A* **2003**, 1011, 67.
- [86] N. Yoshioka, Y. Akiyama, K. Teranishi. Rapid simultaneous determination of o-phenylphenol, diphenyl, thiabendazole, imazalil and its major metabolite in citrus fruits by liquid chromatography-mass spectrometry using atmospheric pressure photoionization. *J. Chromatogr. A* **2004**, 1022, 145.
- [87] J. M. Purcell. Atmospheric pressure photoionization proton transfer for complex organic mixtures investigated by Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2007**, 18, 1682.
- [88] J. H. Banoub, B. Benjelloun-Mlayah, F. Ziarelli, N. Joly, M. Delmas. Elucidation of the complex molecular structure of wheat straw lignin polymer by atmospheric pressure photoionization quadrupole time-of-flight tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2007**, 21, 2867.
- [89] A. C. Stenson, A. G. Marshall, W. T. Cooper. Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization fourier transform ion cyclotron resonance mass spectra. *Anal. Chem.* **2003**, 75, 1275.

- [90] E. B. Kujawinski, P. G. Hatcher, M. A. Freitas. High-resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids: improvements and comparisons. *Anal. Chem.* **2002**, *74*, 413.
- [91] K. Morreel, J. Ralph, H. Kim, F. Lu, G. Goeminne, S. Ralph, E. Messens, W. Boerjan. Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. *Plant Physiol.* **2004**, *136*, 3537.
- [92] I. Damiani, K. Morreel, S. Danoun, G. Goeminne, N. Yahiaoui, C. Marque, J. Kopka, E. Messens, D. Goffner, W. Boerjan, A. M. Boudet, S. Rochange. Metabolite profiling reveals a role for atypical cinnamyl alcohol dehydrogenase CAD1 in the synthesis of coniferyl alcohol in tobacco xylem. *Plant Mol. Biol.* **2005**, *59*, 753.
- [93] J. C. Lepié, R. Dauwe, K. Morreel, V. Storme, C. Lapierre, B. Pollet, A. Naumann, K. Y. Kang, H. Kim, K. Ruel, J.-C. Lepié, R. Dauwe, K. Morreel, V. Storme, C. Lapierre, B. Pollet, A. Naumann, K.-Y. Kang, H. Kim, K. Ruel, A. Lefebvre, J.-P. Joseleau, J. Grima-Pettenati, R. De Rycke, S. Andersson-Gunnerås, A. Erban, I. Fehrle, M. Petit-Conil, J. Kopka, A. Polle, E. Messens, B. Sundberg, S. D. Mansfield, J. Ralph, G. Pilate, W. Boerjan. Downregulation of cinnamoylcoenzyme A reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. *Plant Cell* **2007**, *19*, 3669.
- [94] K. Morreel, O. Dima, H. Kim, F. Lu, C. Niculaes, R. Vanhome, R. Dauwe, G. Goeminne, D. Inzé, E. Messens, J. Ralph, W. Boerjan. Mass spectrometry-based sequencing of lignin oligomers. *Plant Physiol.* **2010**, *153*, 1464.
- [95] J. H. Bowie. The fragmentations of even-electron organic negative ions. *Mass Spectrom. Rev.* **1990**, *9*, 349.
- [96] K. Morreel, H. Kim, F. Lu, O. Dima, T. Akiyama, R. Vanholme, C. Niculaes, G. Goeminne, D. Inzé, E. Messens, J. Ralph, W. Boerjan. Mass spectrometry-based fragmentation as an identification tool in lignomics. *Anal. Chem.* **2010**, *82*, 8095.
- [97] R. Huis, K. Morreel, O. Fliniaux, A. Lucau-Danila, S. Fénart, S. Grec, G. Neutelings, B. Chabbert, F. Mesnard, W. Boerjan, S. Hawkins. Natural Hypolignification is associated with extensive oligolignol accumulation in flax stems. *Plant Physiol.* **2012**, *158*(4), 1893.
- [98] M. Thevis, W. Schanzer, H. Schmidkler. Effect of the location of hydrogen abstraction on the fragmentation of diuretics in negative electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2003**, *4*, 658.
- [99] C. Cheng, M. L. Gross. Applications and mechanisms of charge-remote fragmentation. *Mass Spectrom. Rev.* **2000**, *19*, 398.
- [100] C. E. Costello, J. M. Contado-Miller, J. F. Cipollo. A glycomics platform for the analysis of permethylated oligosaccharide alditoins. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1799.
- [101] B. Domon, C. E. Costello. Structure elucidation of glycosphingolipids and gangliosides using high-performance tandem mass spectrometry. *Biochemistry* **1988**, *27*, 1534.
- [102] B. Domon, C. E. Costello. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J.* **1988**, *5*, 397.
- [103] S. A. McLuckey, G. J. Van Berkel, G. L. Glish. Tandem mass-spectrometry of small multiply charged oligonucleotides. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60.
- [104] P. Roepstorff, J. Fohlman. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biol. Mass Spectrom.* **1984**, *11*, 601.
- [105] R. S. Johnson, S. A. Martin, K. Biemann, J. T. Stults, J. T. Watson. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal. Chem.* **1987**, *59*, 2621.
- [106] D. V. Evtuguin, C. P. Neto, F. M. L. Amado, A. M. S. Silva, D. Robert, P. M. Domingues, O. Faix. Comprehensive study on the chemical structure of dioxane lignin from plantation eucalyptus globulus wood. *J. Agric. Food Chem.* **2001**, *49*, 4252.
- [107] D. V. Evtuguin, F. M. L. Amado. Application of ESI-MS to the elucidation of the primary structure of lignin. *Macromol. Biosci.* **2003**, *3*, 339.
- [108] E. Kiyota, P. Mazzafera, A. C. Sawaya. Analysis of soluble lignin in sugarcane by ultrahigh performance liquid chromatography-tandem mass spectrometry with a do-it-yourself oligomer database. *Anal. Chem.* **2012**, *84*(16), 7015.
- [109] M. Dachtler, F. H. M. van de Put, F. von Stijn, C. M. Beindorff, J. Fritzsche. On-line LC-NMR-MS characterization of sesame oil extracts and assessment of their antioxidant activity. *Eur. J. Lipid Sci. Tech.* **2003**, *105*, 488.
- [110] A. Kamal-Eldin, L. A. Appelqvist, G. Yousif. Lignan analysis in seed oils from four *Sesamum* species: comparison of different chromatographic methods. *J. Am. Oil Chem. Soc.* **1994**, *71*, 141.
- [111] A. Kamal-Eldin, L. A. Appelqvist. Variations in the composition of sterols, tocopherols and lignans in seed oils from four *Sesamum* species. *J. Am. Oil Chem. Soc.* **1994**, *71*, 149.
- [112] Y. S. Shyu, L. S. Hwang. Antioxidative activity of the crude extract of lignin glycosides from unroasted Burma black sesame meal. *Food Res. Int.* **2002**, *35*, 357.
- [113] A. A. Moazzami, R. E. Andersson, A. Kamal-Eldin. Characterization and analysis of sesamolinol diglucoside in sesame seeds. *Biosci. Biotechnol. Biochem.* **2006**, *70*, 1478.
- [114] A. A. Moazzami, R. E. Andersson, A. Kamal-Eldin. HPLC analysis of sesamolin glucosides in sesame seeds. *J. Agric. Food Chem.* **2006**, *54*, 633.
- [115] K. P. Suja, A. Jayalekshmy, C. Arumughan. Free radical scavenging behavior of antioxidant compounds of sesame (*Sesamum indicum* L.) in DPPH center dot system. *J. Agric. Food Chem.* **2004**, *52*, 912.
- [116] Y. Jiao, L. B. Davin, N. G. Lewis. Furanofuran lignan metabolism as a function of seed maturation in *Sesamum indicum*: methylenedioxy bridge formation. *Phytochemistry* **1998**, *49*, 387.
- [117] E. Ono, M. Nakai, Y. Fukui, N. Tomimori, M. Fukuchi-Mizutani, M. Saito, H. Satake, T. Tanaka, M. Katsuta, T. Umezawa, Y. Tanaka. Formation of two methylenedioxy bridges by a *Sesamum* CYP81Q protein yielding a furofuran lignan, (+)-sesamin. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 10116.
- [118] K. Struijs, J.-P. Vincken, H. Gruppen. Comparison of atmospheric pressure chemical ionization and electrospray ionization mass spectrometry for the detection of lignans from sesame seeds. *Rapid Commun. Mass Spectrom.* **2008**, *22*(22), 3615.
- [119] M. Nagata, T. Osawa, M. Namiki, Y. Fukuda, T. Ozaki. Stereochemical structures of antioxidative bisepoxylignans, sesamolin and its isomers, transformed from sesamolin. *Agric. Biol. Chem.* **1987**, *51*, 1285.
- [120] P. C. Eklund, M. J. Backman, L. A. Kronberg, A. I. Smeds, R. E. Sjoholm. Identification of lignans by liquid chromatography-electrospray ionization ion-trap mass spectrometry. *J. Mass Spectrom.* **2008**, *43*, 97.
- [121] M. Ye, Y. N. Yan, D. A. Guo. Characterization of phenolic compounds in the Chinese herbal drug Tu-Si-Zi by liquid chromatography coupled to electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 1469.
- [122] G. Q. Yan, Q. R. Li, H. R. Tan, T. F. Ge. Electrospray ionization ion-trap time-of-flight tandem mass spectrometry of two furanofurans: sesamin and gmelinol. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 3613.
- [123] L. J. Haupert, B. C. Owen, C. L. Marcum, T. M. Jarrell, C. J. Pulliam, L. M. Amundson, P. Narra, M. S. Aqueel, T. H. Parsell, M. M. Abu-Omar, H. I. Kenttämäa. Characterization of model compounds of processed lignin and the lignome by using atmospheric pressure ionization tandem mass spectrometry. *Fuel* **2012**, *95*, 634.
- [124] B. C. Owen, L. J. Haupert, T. M. Jarrell, C. L. Marcum, T. H. Parsell, M. M. Abu-Omar, J. J. Bozell, K. Black, H. Kenttämäa. High-performance liquid chromatography/high-resolution multiple stage tandem mass spectrometry using negative-ion-mode hydroxide-doped electrospray ionization for the characterization of lignin degradation products. *Anal. Chem.* **2012**, *84*, 6000.
- [125] P. C. Pinto, D. V. Evtuguin, C. P. Neto, A. J. D. Silvestre, F. M. L. Amado. Behavior of *Eucalyptus globulus* lignin Kraft, NMR, ESI-MS, and GPC. *J. Wood Chem. Tech.* **2002**, *22*(2–3), 109.
- [126] A. A. Shatalov, D. V. Evtuguin, C. Pascoal Neto. (2-O- α -D-Galactopyranosyl-4-O-methyl- α -D-glucuronol-D-xylan from *Eucalyptus globulus* labill. *Carbohydr. Res.* **1999**, *320*, 93.
- [127] B. L. Browning. *Methods in Wood Chemistry*, Vol. II. Interscience Publishers: New York, **1967**, 732–734.
- [128] G. F. Zakis. Synthesis of Lignin Model Compounds, Zinatne, Riga, **1980**.
- [129] H. Onnerud, M. Palmlad, G. Gellerstedt. Investigation of lignin oligomers using electrospray ionisation mass spectrometry. *Holzforschung* **2003**, *57*, 37.
- [130] C. Rolando, B. Monties, C. Lapierre. *Thioacidolysis*. Springer Series in Wood Science, Springer Verlag: Berlin, **1992**, 334–349.
- [131] E. Kendrick. A mass scale based on $CH_4 = 14.0000$ for high resolution mass spectrometry of organic compounds. *Anal. Chem.* **1963**, *35*, 2146.
- [132] E. B. Kujawinski, M. A. Freitas, X. Zang, P. G. Hatcher, K. B. Green-Church, R. B. Jones. The application of electrospray ionization mass spectrometry (ESI-MS) to the structural characterization of natural organic matter. *Org. Geochem.* **2002**, *33*, 171.

- [133] K. D. Henry, F. W. McLafferty. Electrospray ionization with Fourier-transform mass spectrometry. Charge state assignment from resolved isotopic peaks. *Org. Mass Spectrom.* **1990**, 25, 490.
- [134] X. Huang, F. Song, Z. Liu, S. Liu. Studies on lignan constituents from *Schisandra chinensis* (Turcz.) Baill. Fruits using high-performance liquid chromatography/electrospray ionization multiple-stage tandem MS. *J. Mass Spectrom.* **2007**, 42(9), 1148.
- [135] R. Vanholme, B. Demedts, K. Morreel, J. Ralph, W. Boerjan. Lignin biosynthesis and structure. *Plant Physiol.* **2010**, 153(3), 895.
- [136] K. Ruel, J. Berrio-Sierra, M. Mir Derikvand, B. Pollet, J. Thévenin, C. Lapierre, L. Jouanin, J.-P. Joseleau. Impact of CCR1 silencing on the assembly of lignified secondary walls in *Arabidopsis thaliana*. *New Phytol.* **2009**, 184(1), 99.
- [137] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, 246(4926), 64.
- [138] K. Tanaka. The origin of macromolecule ionization by laser irradiation (Nobel Lecture). *Angew. Chem. Int. Ed.* **2003**, 42(33), 3860.
- [139] K. Wüthrich. Nuclear magnetic resonance spectroscopy for determining the three-dimensional structure of biological macromolecules in solution. *Nobel Lecture*, **2002**, 235.
- [140] J. M. Zhou, R. K. Ibrahim. Tricin – a potential multifunctional nutraceutical. *Phytochem. Rev.* **2010**, 9, 413.
- [141] D. H. Kim, B. G. Kim, Y. S. Lee, Y. R. Ji, Y. H. Lim, H. G. Hur, J. H. Ahn. Regiospecific methylation of naringenin to ponciretin by soybean O-methyltransferase expressed in *Escherichia coli*. *J. Biotechnol.* **2005**, 119(2), 155.
- [142] G. Schröder, E. Wehinger, R. Luka, F. Wellmann, W. Seefelder, W. Schwab, J. Schröder. Flavonoid methylation: a novel 4'-O-methyltransferase from *Catharanthus roseus*, and evidence that partially methylated flavanones are substrates of four different flavonoid dioxygenases. *Phytochemistry* **2004**, 65(8), 1085.
- [143] B. G. Kim, Y. S. Lee, H. G. Hur, Y. H. Lima, J. H. Ahn. Flavonoid 3'-O-methyltransferase from rice: cDNA cloning, characterization and functional expression. *Phytochemistry* **2006**, 67(4), 387.
- [144] H. Cai, W. P. Steward, A. J. Gescher. Determination of the putative cancer chemopreventive flavone tricin in plasma and tissues of mice by HPLC with UV-visible detection. *Biomed. Chromatogr.* **2005**, 19, 518.
- [145] A. Björkman. Studies on finely divided wood. Part I. Extraction of lignin with neutral solvents. *Sven. Papperstidn* **1956**, 13, 477.
- [146] A. Guerra, I. Filpponen, L. A. Lucia, C. Saquing, S. Baumberger, D. S. Argyropoulos. Toward a better understanding of the lignin isolation process from wood. *J. Agric. Food Chem.* **2006**, 54, 5939.
- [147] B. Benjelloun-Mlayah, G. Avignon, M. Delmas. Installation for implementing a method for producing paper pulp, lignins and sugars and production method using such an installation. Patent n° WO 2006117295 A1, **2006**.
- [148] M. Delmas. Vegetal Refining and Agrochemistry. *Chem. Eng. Technol.* **2008**, 31(5), 792.
- [149] G.-H. Delmas, B. Benjelloun-Mlayah, Y. Le Bigot, M. Delmas. Functionality of wheat straw lignin extracted in organic acid media. *J. App. Sci.* **2011**, 127(1), 191.
- [150] J. H. Banoub, G.-H. Delmas, N. Cachet, B. Benjelloun-Mlayah, M. Delmas, unpublished results.
- [151] K. L. Busch. Special feature: *Tutorial*. Desorption ionization mass spectrometry. *J. Mass Spectrom.* **1995**, 30, 233.
- [152] K. Yoshioka, D. Ando, T. Watanabe. A comparative study of matrix and nano LDI-TOF-MS of isolated and synthetic lignin. *Phytochem. Anal.* **2011**, 23, 248.
- [153] P. Bocchini, G. C. Galletti, R. Seraglia, P. Traldi, S. Camarero, A. T. Martinez. Matrix-assisted laser desorption/ionization mass spectrometry of natural and synthetic lignin. *Rapid Commun. Mass Spectrom.* **1996**, 10, 1144.
- [154] J.-Y. Liu, S.-B. Wu, R. Lou. Chemical structure and pyrolysis response of β -O-4 lignin model polymer. *BioResources* **2011**, 6(2), 1079.
- [155] T. Kishimoto, Y. Uraki, M. Ubukata. Synthesis of bromoacetophenone derivatives as starting monomers for β -O-4 type artificial lignin polymers. *J. Wood Chem. Tech.* **2008**, 28, 97.
- [156] M.-L. Mattinen, T. Suortti, R. Gosselink, D. S. Argyropoulos, D. Evtuguin, A. Suurnakki, E. de Jong, T. Tamminen. Polymerization of different lignins by laccase. *BioResources* **2008**, 3(2), 549.
- [157] A. Guerra, L. A. Lucia, D. S. Argyropoulos. Isolation and characterization of lignins from *Eucalyptus grandis* Hill ex Maiden and *Eucalyptus globulus* Labill. by enzymatic mild acidolysis (EMAL). *Holzforschung* **2008**, 62, 24.
- [158] A. Jacobs, O. Dahlman. Absolute molar mass of lignins by size exclusion chromatography and MALDI-TOF mass spectrometry. *Nord. Pulp Pap. Res. J.* **2000**, 15(2), 120.
- [159] R. J. A. Gosselink. Lignin as a renewable aromatic resource for the chemical industry. Ph.D. Thesis, Wageningen University, Wageningen, NL, **2011**.
- [160] R. Bayerbach, V. D. Nguyen, U. Schurr, D. Meier. Characterization of the water-insoluble fraction from fast pyrolysis liquids (pyrolytic lignin). Part III. Molar mass characteristics by SEC, MALDI-TOF-MS, LDI-TOF-MS and Py-FIMS. *J. Anal. Appl. Pyrolysis* **2006**, 77, 95.
- [161] K. Saito, T. Kato, Y. Tsuji, K. Fukushima. Identifying the characteristic secondary ions of lignin polymer using SIMS-TOF. *Biomacromolecules* **2005**, 6, 678.
- [162] K. Saito, T. Kato, H. Takamori, T. Kishimoto, A. Yamamoto, K. Fukushima. A new analysis of the depolymerized fragments of lignin. *Appl. Surf. Sci.* **2006**, 252(9), 6734.
- [163] K. Saito, T. Mitsutani, T. Imai, Y. Matsushita, A. Yamamoto, K. Fukushima. Chemical differences between sapwood and heartwood of *Chamaecyparis obtusa* detected by SIMS-TOF. *Appl. Surf. Sci.* **2008**, 255, 1088.
- [164] E. N. Tokareva, A. V. Pranovich, P. Fardim, G. Daniel, B. Holmbom. Analysis of wood tissues by time-of-flight secondary ion mass spectrometry. *Holzforschung* **2007**, 61, 647.
- [165] K. Koljonen, M. Oesterberg, M. Kleen, A. Fuhrmann, P. Stenius. Precipitation of lignin and extractives on kraft pulp: effect on surface chemistry, surface morphology and paper strength. *Cellulose (Dordrecht, Neth.)* **2004**, 11, 209.
- [166] P. Fardim, N. Duran. Modification of fibre surfaces during pulping and refining as analysed by SEM, XPS and ToFSIMS. *Colloids Surf. A* **2003**, 223, 263.
- [167] E. S. F. Berman, K. S. Kulp, M. G. Knize, L. Wu, E. J. Nelson, D. O. Nelson, K. J. Wu. Distinguishing monosaccharide stereo- and structural isomers with SIMS-TOF and multivariate statistical analysis. *Anal. Chem.* **2006**, 78, 6497.
- [168] R. E. Goacher, D. Jeremic, E. R. Master. Expanding the library of secondary ions that distinguish lignin and polysaccharides in time-of-flight secondary ion mass spectrometry analysis of wood. *Anal. Chem.* **2011**, 83, 804.
- [169] K. A. Lunsford, G. F. Peter, R. A. Yost. Direct matrix-assisted laser desorption/ionization mass spectrometric imaging of cellulose and hemicellulose in *Populus* tissue. *Anal. Chem.* **2011**, 83(17), 6722.
- [170] S. Ding, Q. Xu, M. K. Ali, J. O. Baker, E. A. Bayer, Y. Barak, R. Lamed, J. Sugiyama, G. Rumbles, M. E. Himmel. Versatile derivatives of carbohydrate-binding modules for imaging of complex carbohydrates approaching the molecular level of resolution. *Biotechniques* **2006**, 41, 435.
- [171] N. Ishida, M. Koizumi, H. Kano. The NMR microscope: a unique and promising tool for plant science. *Ann. Bot.* **2000**, 86, 259.
- [172] J. V. D. Bulcke, B. Masschaele, M. Dierick, J. V. Acker, M. Stevens, L. V. Hoorebeke. Three-dimensional imaging and analysis of infested coated wood with X-ray submicron CT. *Int. Biodeterior. Biodegrad.* **2008**, 61, 278.
- [173] U. P. Agarwal. Raman imaging to investigate ultrastructure and composition of plant cell walls: distribution of lignin and cellulose in black spruce wood (*Picea mariana*). *Planta* **2006**, 224, 1141.
- [174] N. Gierlinger, M. Schwanninger. Chemical imaging of poplar wood cell walls by confocal Raman microscopy. *Plant Physiol.* **2006**, 140, 1246.
- [175] A. M. Belu, D. J. Graham, D. G. Castner. Time-of-Flight secondary ion mass spectrometry: techniques and applications for the characterization of biomaterial surfaces. *Biomaterials* **2003**, 24, 3635.
- [176] R. N. S. Sodhi. Time-of-flight secondary ion mass spectrometry (SIMS-TOF): versatility in chemical and imaging surface analysis. *Analyst* **2004**, 129, 483.
- [177] M. Karas, F. Hillenkamp. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **1988**, 60, 2299.
- [178] R. M. Caprioli, T. B. Farmer, J. Gil. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* **1997**, 69(23), 4751.
- [179] T. J. Garrett, M. Prieto-Conaway, V. Kovtoun, H. Bui, N. Izgarian, G. Stafford, R. A. Yost. Imaging of small molecules in tissue sections with a new intermediate-pressure MALDI linear ion trap mass spectrometer. *Int. J. Mass Spectrom.* **2007**, 260, 166.

- [180] L. A. McDonnell, R. M. A. Heeren. Imaging mass spectrometry. *Mass Spectrom. Rev.* **2007**, *26*, 606.
- [181] K. Kato, K. Endo, D. Matsumoto, T. Ida, K. Saito, K. Fukushima, N. Kato. Simulation of SIMS for monomer and dimer of lignin under the assumption of thermal decomposition using QMD method. *Appl. Surf. Sci.* **2008**, *255*(4), 1040.
- [182] R. R. Landgraf, M. C. P. Conaway, T. J. Garrett, P. W. Stacpoole, R. A. Yost. Imaging of lipids in spinal cord using intermediate pressure matrix-assisted laser desorption-linear ion trap/Orbitrap MS. *Anal. Chem.* **2009**, *81*, 8488.
- [183] R. F. Evert. *Esau's Plant Anatomy: Meristems, Cells, and Tissues of the Plant Body: Their Structure, Function and Development*, 3rd ed. John Wiley & Sons: Hoboken, NJ, **2006**.
- [184] P. Araújo, M. S. Ferreira, D. N. de Oliveira, L. Pereira, A. C. Sawaya, R. R. Catharino, P. Mazzafera. Mass spectrometry imaging: an expeditious and powerful technique for fast *in situ* lignin assessment in eucalyptus. *Anal. Chem.* **2014**, *86*(7), 3415.
- [185] D. N. de Oliveira, S. de Bona Sartor, M. S. Ferreira, R. R. Catharino. Cosmetic analysis using matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI). *Materials* **2013**, *6*, 1000.
- [186] A. Richel, C. Vanderghem, M. Simon, B. Wathelet, M. Paquot. Evaluation of MADLI-TOF-MS for second-generation lignin analysis. *Anal. Chem. Insights* **2012**, *7*, 79.
- [187] C. Vanderghem, A. Richel, N. Jacquet, C. Blecker, M. Paquot. Impact of formic/acetic acid and ammonia pre-treatments on chemical structure and physico-chemical properties of *Miscanthus x giganteus* lignins. *Polym. Degrad. Stab.* **2011**, *96*, 1761.
- [188] F. Abdelkafi, H. Ammar, B. Rousseau, M. Tessier, R. El Gharbi, A. Fradet. Structural analysis of Alfa grass (*Stipa tenacissima* L.) lignin obtained by acetic/formic acid delignification. *Biomacromolecules* **2011**, *12*, 3895.
- [189] L. B. Davin, D. L. Bedgar, T. Katayama, N. G. Lewis. On the stereoselective synthesis of (1)-pinoresinol in *Forsythia suspensa* from its achiral precursor, coniferyl alcohol. *Phytochemistry* **1992**, *31*, 3869.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.