Analytical Chemistry Insights

ORIGINAL RESEARCH

Evaluation of Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry for Second-Generation Lignin Analysis

Aurore Richel, Caroline Vanderghem, Mathilde Simon, Bernard Wathelet and Michel Paquot

Unit of Biological and Industrial Chemistry, Gembloux Agro-Bio Tech, University of Liege, Gembloux, Belgium.
Corresponding author email: a.richel@ulg.ac.be

Abstract: Matrix-Assisted Laser Desorption/Ionization time-of-flight (MALDI-TOF) mass spectrometry is evaluated as an elucidation tool for structural features and molecular weights estimation of some extracted herbaceous lignins. Optimization of analysis conditions, using a typical organic matrix, namely α-cyano-4-hydroxycinnamic acid (CHCA), in combination with α-cyclodextrin, allows efficient ionization of poorly soluble lignin materials and suppression of matrix-related ions background. Analysis of low-mass fragments ions (*m/z* 100–600) in the positive ion mode offers a "fingerprint" of starting lignins that could be a fine strategy to qualitatively identify principal inter-unit linkages between phenylpropanoid units. The molecular weights of lignins are estimated using size exclusion chromatography and compared to MALDI-TOF-MS profiles. Miscanthus (*Miscanthus x giganteus*) and Switchgrass (*Panicum Virgatum* L.) lignins, recovered after a formic acid/acetic acid/water process or aqueous ammonia soaking, are selected as benchmarks for this study.

Keywords: mass spectrometry, biomass, lignin

Analytical Chemistry Insights 2012:7 79–89

doi: 10.4137/ACI.S10799

This article is available from http://www.la-press.com.

Introduction

Lignin is the cell wall-resistant component of vascular plants and represents 15% to 30%, by weight, of the lignocellulosic feedstock.[1] Lignin is therefore the third most abundant biopolymer on earth after cellulose and hemicellulose. In the context of an integrated biorefinery concept, it is important to upgrade the entire lignocellulosic material.[2] Whilst the production of fermentable sugars, platforms and building blocks from both cellulose and hemicelluloses is extensively developed, the non-energetic valorization of lignin remains a major challenging task.[3] The final application of lignin can indeed significantly vary as a function of both extraction methods and type of lignocellulosic sources. The development of fast, non-invasive, and reliable methods for the physicochemical characterization of extracted lignin is thus a prerequisite to optimizing processing conditions such that they are adequate.

Native lignin is an amorphous, three-dimensional cross-linked heteropolymer consisting of phenylpropanoid units joined together by several specific ether or carbon-carbon bonds.[4] Lignin is composed of three phenylpropane derivatives in varying ratios, referred to as p-hydroxyphenyl (H), guaiacyl (G) and syringyl (S), linked by β-O-4, β-β, β-5 (phenylcoumaran), 4-O-5 or β-1 bonds in various bonding patterns (Fig. 1). Practically, lignin can be extracted from lignocellulosic feedstocks by physical and/or chemical processes that modify its native structure. Ether linkage of the β-O-4 type is the most frequent in herbaceous plants and its solvolysis under formic acid/acetic acid (Delmas) conditions is used to depolymerize lignin, as carbon-carbon bonds are more resistant to degradation. Conversely, under alkali conditions, ie, ammonia soaking, phenylcoumaran substructures are easily broken whilst β-O-4 linkages remain almost unaltered.[5]

Because value-added applications of extracted (depolymerized) lignins are envisioned, complete structural characterization of lignins is required to ensure a better comprehension and prediction of structure-functions. Due to lignin structural heterogeneity, traditional spectroscopic methods

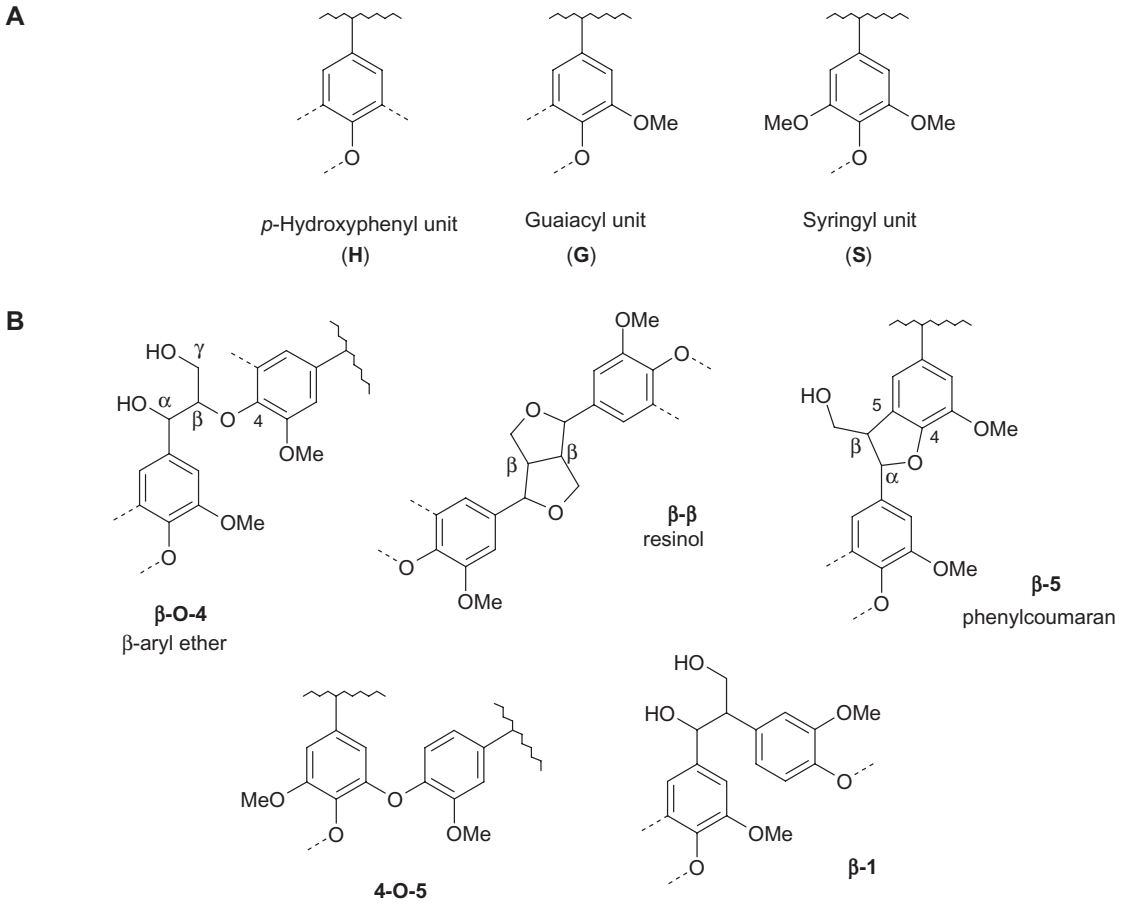

Figure 1. (**A**) Generic lignin units, and (**B**) main interunit linkages.

still reveal a fascinating complexity. Liquid nuclear magnetic resonance (NMR) spectroscopy, including two-dimensional techniques and quantitative ^{13}CNMR, provides extensive information of the salient structural features, mainly monolignols (G/S/H) and/or end-groups distribution, and quantification of inter-unit linkages.[6–8] Although strikingly useful, the obtainable signal-to-noise ratio for these analyses is only improved for high (derivatized) lignin concentrations (up to 50 mg/mL) in analysis solvents combined with long acquisition times (up to several days). The propensity of lignins to associate and aggregate in organic solutions, identifying apparent molecular weight changes over time and possible sedimentation of the samples during analysis, is another drawback of wet analysis methods. This aggregation behavior has been evidenced using static light scattering photometric (MALLS) measurements, and is notably dependent on the nature of the analysis solvent.

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) is commonly used for the characterization of synthetic or natural polymers.[9] Information such as average molecular weights (weight and number; M_w and M_n), nature and number of repeating units, and end-group determination can be deduced form MALDI-TOF-MS data. Even if widely accepted for synthetic polymers, the application of MALDI-TOF-MS for lignin analysis remains poorly explored. Most studies have been reported for high-mass "model" lignin samples, usually commercial Kraft lignins or synthetic lignins prepared by enzymatic polymerization of coniferyl alcohol, or natural milled wood lignins (from spruce, pine or birch).[10,11] Published spectra usually reveal low mass resolution and molecular weights accuracy, mainly due to a low efficiency of ionization of lignin samples and technical limitations (older instrumentations used).[12,13] Partial identification of the nature of repetitive units and regulation of the type of linkages is possible, but only for lignin samples with a weight superior to 1000 Da, as interferences with matrix-related peaks occur.

Due to their structural composition, their formal charge and/or their molecular weights, lignins are sometimes poorly soluble in classical analysis solvents (for example, water, chloroform, and toluene). The quest for a fast technical method for their investigation, without prior chemical derivatization that might deteriorate lignin structure, has prompted us to investigate MALDI-TOF-MS as a potential solution. This mass spectrometry methodology is non-destructive and, conversely to traditional wet chemical analytical tools, can possibly reflect the entire lignin sample composition. The scope of this work is to describe a versatile and straightforward analytical protocol for lignin mass fingerprint analysis and average molecular weight estimation. Miscanthus (*Miscanthus x giganteus*) and Switchgrass (*Panicum Virgatum* L.) lignins were chosen for this study because these two warm-season grasses have attracted considerable attention as possible dedicated energy crops.

Experiment

Lignin materials

Several herbaceous isolated lignins were selected for this study. These technical lignins were characterized by a S/G ratio of about 1/1 as estimated by quantitative ^{13}C NMR.[5] Three selected lignin fractions (FAL107, FAL80 and AL) were extracted from *Miscanthus x giganteus* under alkali or acid conditions according to a previously published protocol.[5] Lignin isolated after a formic acid/acetic acid pretreatment at 107 °C (FAL107) contained mainly β-5 linkages as evidenced by nuclear magnetic resonance. AL designates herein the lignin fraction obtained after aqueous ammonia soaking and presented predominantly 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. Lignin fractions SG-107 and SG-AL were extracted from Switchgrass feedstock under aforementioned formic acid/acetic acid pretreatment at 107 °C for 3 hours and aqueous ammonia conditions (60 °C for 6 hours), respectively. SG-107 was composed mainly of β-5 linkages; SG-AL presented classical β-O-4 bonds.[5]

The solubility of lignin samples was determined gravimetrically by filtration (0.45 µm filters) of the suspension (10 mg/mL lignins in organic solvent or water) after 1 hour of stirring at room temperature. Solubility of unmodified FAL107 was inferior to 0.1 g/L in water (18 °C, pH 6). In non-polar solvents such as toluene or chloroform, its solubility was measured at about 3.2 and 1.5 g/L respectively. In polar aprotic solvents such as acetone, its solubility was slightly improved with a value of 5.6 g/L. For FAL80 and AL, this solubility in acetone was of 6.4 and 1.2 g/L respectively and of 3.9 and 0.8 g/L for SG-107 and SG-AL.

Molecular weight determination using HPSEC-RI

Lignin molecular weights were estimated using size-exclusion chromatography (SEC) and were expressed as polystyrene-equivalent molecular weights (M_p).[14] Three Styragel columns (HR1, HR2 and HR3; Waters) were connected in series to an HPLC system (Agilent Technologies, 1200 series) equipped with a differential refractometer (BI-DNDC/GPC, 620 nm) using THF as the eluent (flow rate, 1 mL/min). Analyses were performed in duplicate. Calibration was ensured using commercial polystyrene standards. Measurements were performed on acetylated or non-acetylated lignins dissolved in THF (1 mg/mL). Acetylation was achieved according a published protocol.[5]

MALDI-TOF-MS instrumentation and conditions

High resolution mass spectra were recorded on an Ultraflex MALDI time-of-flight mass spectrometer (Bruker Daltonic) operating in the reflector and linear positive or negative mode using a pulsed nitrogen laser (λ 337 nm, pulse rate of 10 Hz) as the ionization source. The analyzer was used at an acceleration voltage of ±20 kV. Laser light was focused on the sample using a 5.2 kV lens. A pulsed ion extraction was optimized to 200 ns. A peptide calibration set was used as the external standard. For each lignin sample, MALDI-TOF-MS was ensured on 10 distinct sample deposit zones. A total of 300 shots were provided. Each sample preparation and analysis were duplicated.

α-CD/CHCA matrix preparation

α-CD/CHCA matrix was prepared by mixing 240 mL of a solution of 10 mM α-CD in water with 40 mL of CHCA 75 mM in acetonitrile/H_2O (7/3 v/v) containing 0.1 vol.% trifluoroacetic acid. The complex α-CD/CHCA was obtained after sonication at 50 °C for 1 hour. Interaction between α-CD and CHCA was confirmed by UV-Visible spectroscopy.

Results and Discussion

Optimization of MALDI-TOF-MS for extracted lignins

FAL107 lignin was arbitrarily selected for the optimization process. FAL107 absorbed conveniently at λ 337 nm and could be analyzed without a matrix, as proposed by Yoshioka et al,[15] using a nanostructured silicon-based target. However, in our study, preparation of the lignin sample with a traditional matrix was preferred, as partial oxidation of unsaturation sites was reported under matrix-free conditions and might cause confusions with the presence of some structural features characteristic of lignins.[16]

After a screening of conventional commercial MALDI matrices,[17] only aromatic acidic compounds, such as α-cyano-4-hydroxycinnamic acid (CHCA) or 2,5-dihydroxybenzoic acid (2,5-DHB), were found convenient for sufficient lignin-related ions production and adequate mass resolution with our incident λ 337 nm laser equipment ionization source. Best results were also obtained when MALDI-TOF-MS sample preparation was achieved using the "sandwich method", where the lignin analyte was deposited between two fast evaporated matrix layers. Practically, the lignin is not premixed with the matrix. A lignin sample droplet is applied on top of a fast-evaporated matrix layer (acetone), followed by the deposition of a second layer of matrix in non-volatile solvent (acetonitrile). The performance of CHCA and 2,5-DHB was then examined using four "classical" solvent systems: 30% aq. acetonitrile with 0.1% trifluoroacetic acid (TFA); 10% aq. acetone; 30% aq. acetone; and water-methanol (2:1). Saturated solutions of both matrices were used. Adjunction of sodium chloride was required to ensure suitable sample ionization for FAL107/DHB and addition of TFA was necessary for FAL107/CHCA. MALDI-TOF-MS spectra were acquired in the reflector positive ion mode by randomly irradiating 300 times the sample spot. The pulse ion extraction delay was optimized at 200 ns. Because of its solubility in polar aprotic solvents, FAL107 was prepared in acetone. A concentration of 0.5 to 5.6 mg/mL provided convenient MALDI-TO-MS sensitivity.

Figure 2 shows the best quality spectra obtained for FAL107 using either 2,5-DHB in acetone-water (7:3) in combination with NaCl (Fig. 2A) or CHCA in 30% aq. acetonitrile with 0.1% TFA (Fig. 2B). The MALDI-TOF-MS spectrum for FAL107 sandwiched between two layers of CHCA revealed an oligomeric distribution from about 100 to 600 Da, with a gradual decrease in intensity. Matrix clusters signals appeared in this low mass range and probably interfered with lignin fingerprints. This drawback was also noticed using DHB, and was amplified with the addition

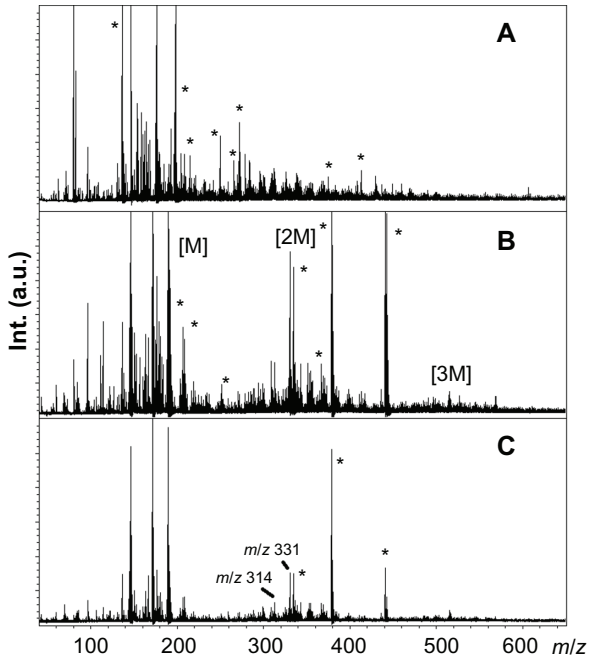

Figure 2. Reflector positive ion mode MALDI-TOF-MS spectrum of FAL107 with DHB (**A**), CHCA (**B**), and α-CD/CHCA (**C**).

of sodium chloride. Taking into account that constitutive monolignols in FAL107 were syringyl and coniferyl alcohols in a 1/1 ratio (average molar mass: 195.2 $g \cdot mol^{-1}$), the three main oligomeric distributions denoted in Figure 2B were attributed to monomers (denoted [M]), dimers ([2M]) and trimers ([3M]). Assignment of mass signals of these low-mass distributions was partially hampered by inherent matrix-related ions below *m/z* 500.

To extend the range of applicability of MALDI-TOF-MS for lignin analysis, we checked a universal lignin/matrix sample preparation strategy based on the encapsulation of CHCA in a cyclodextrin cavity.[18] Then, only the protonated matrix-related ions were detected (mainly at *m/z* 189, and to a lesser extend at 293 and 334 and 379 Da), while their intensities and unwanted fragmentations were significantly suppressed (Fig. 3). Interference with low-molecular masses fragments of lignins could thus be minimized. The "cyclodextrin-supported matrix" (α-CD/CHCA) was prepared by mixing α-cyclodextrin (α-CD) with CHCA in a 1/1500 molar ratio. Exemplification of this new matrix for FAL107 analysis revealed a "cleaner" spectrum with less interference with matrix fragments (Fig. 2C). A decrease of sensitivity was, however, recorded (for instance, signal-to-noise ratio for the peak at *m/z* 314 falling from 18.8 for FAL107/CHCA to 12.6 for FAL107/α-CD/CHCA) and represented an identifiable limitation of this method. Good quality spectra were recorded for concentrations of FAL107 in acetone ranging from 1.5 to 3.5 mg/mL. Preparation

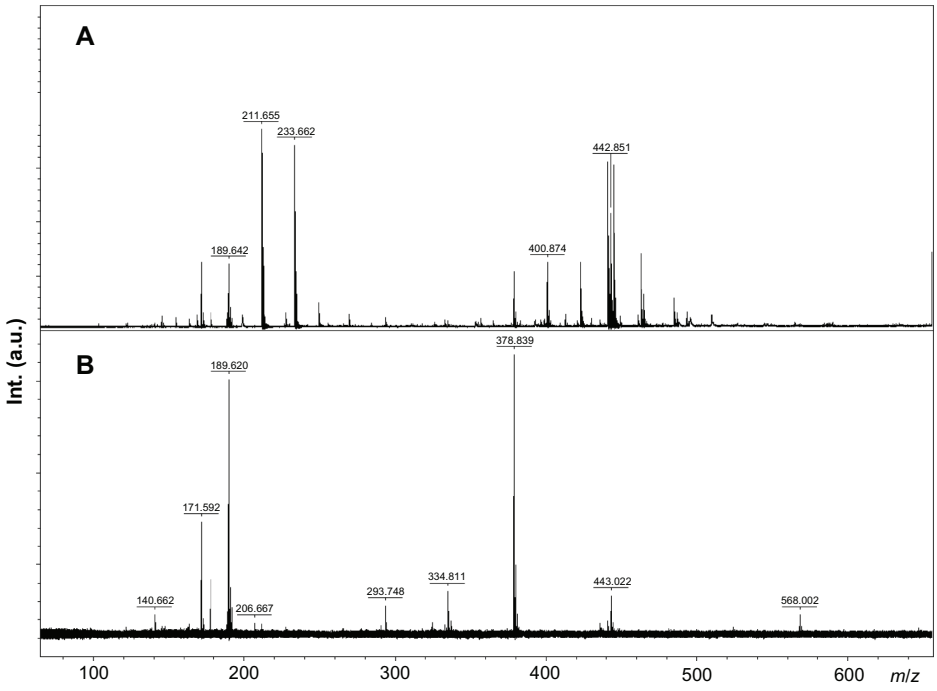

Figure 3. Reflector positive ion mode MALDI-TOF-MS spectrum of CHCA alone (without lignin) (**A**) and CHCA/α-CD alone (without lignin) (**B**) evidencing their corresponding protonated matrix-related ions.

of the matrix spot starting from an aqueous suspension of FAL107 between two deposits of α-CD/CHCA also provided a convenient spectrum with a signal-to-noise ratio of 9.2 for *m/z* 314. Ionization in the positive ion mode provided higher quality spectra than the corresponding negative ion spectra.

MALDI and HPSEC analysis of technical lignins

HPSEC analysis of FAL107 in THF revealed a mono-modal distribution curve with an apparent molecular weight (M_p) of about 2000 for non-acetylated FAL107 and of 2800 for its acetylated counterpart. MALDI-TOF-MS analysis for FAL107/α-CD/CHCA revealed, however, lower molecular weight compounds ranging from 100 to 600 Da, clearly inferior to the molecular weight measured by HPSEC (Fig. 2C). A depolymerization of the sample under laser beam could be suggested. We decided to explore if MALDI-TOF spectral area between *m/z* 100–600 could supply a characteristic fingerprint of the lignin, and to correlate this approach to previous NMR investigations.[5]

Constitutive monolignols ions were observed below 200 Da and were detected in the positive ion mode spectrum as their proton and alkali adduct ions. Specific MS peaks were recorded, with variable intensity, at *m/z* 137, 151 and *m/z* 167 and 181. For FAL107, which was a guaiacyl-syringyl (GS) lignin, these fragments ions were assigned respectively to benzyl cations of guaiacyl (G) and syringyl (S) units with free phenolic groups.[19] A poor Δm/m resolution hampered the possibility of a more accurate assignment. Secondary peak sequences at *m/z* 146 and 176 suggested the presence of sodium adduct of completely demethylated G and S units. In the negative ion mode, peaks at *m/z* 151, 181 and 163, 193 were only detected and tentatively assigned to deprotonated adducts of $[C_8H_7O_3]^-$, $[C_9H_9O_4]^-$, and $[C_{10}H_{11}O_2]^-$, $[C_{11}H_{13}O_3]^-$.[19]

With *m/z* ranging from about 250 to 400, the hyperfine structure of this dimeric distribution offered special indication regarding the occurrence of specific linkages between two phenylpropane units. Elucidation of the dimeric structure of FAL107 corroborated the seminal work of Banoub and Delmas using APCI-MS (atmospheric pressure chemical ionization) on acid-extracted wheat straw lignins.[20] The main experimental peak with a well-resolved isotopic distribution located at about *m/z* 331, associated to particular fragment at *m/z* 315, was assigned to protonated $[C_{18}H_{19}O_6]^+$ adduct, where two G units were linked through a β-5 moiety (Fig. 4). This fit well with our previous NMR investigations. This observation allowed to propose that main linkages in dilignols distribution of acid pretreated Miscanthus was of the β-5 type. In the trimers area, where the S/N ratio was quite low, prevalence of a MS signal at *m/z* 509 was the only valuable data. This peak was assigned to $[C_{28}H_{29}O_9]^+$ protonated structure. A mass increment of 178 Da between main peaks in the dimers and trimers distributions was, according to Yoshioka et al,[15] relevant of a repetitive β-5 bonding pattern (see Supplementary Information S1).[15] In the negative ionization mode, prevalence of signals at *m/z* 314 and 329 was detected with secondary *m/z* at 341 assigned respectively to deprotonated $[C_{18}H_{17}O_6]^-$, $[C_{17}H_{14}O_6]^-$, and $[C_{19}H_{17}O_6]^-$ fragments.[21] These MALDI-TOF-MS analyses thus provided "diagnostic" product ions, which enabled us to determine the main type of linkages between phenylpropane units at first sight.

Transfer of this MALDI-TOF-MS extrapolative approach was envisaged for Miscanthus FAL80 and AL lignins (Fig. 5). HPSEC-RI analyses revealed mono-modal mass distributions with M_p values of 1700 for non-acetylated FAL80 and of 2300 for its acetylated analogue. Acetylation of AL was mandatory for HPSEC analysis. An average molecular mass M_p of 3100 was measured. For FAL80 using the MALDI-TOF-MS positive ionization mode, peaks distribution in the *m/z* 100–400 region fit well with that of FAL107 underlying the same monolignols composition and the same β-5 linkage participation between

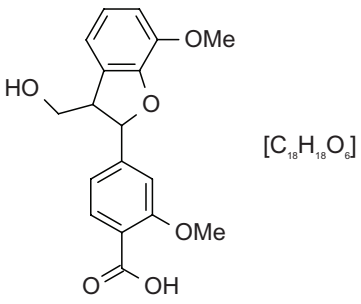

Figure 4. Chemical structure of lignin dimeric fragment as proposed by Delmas and Banoub.

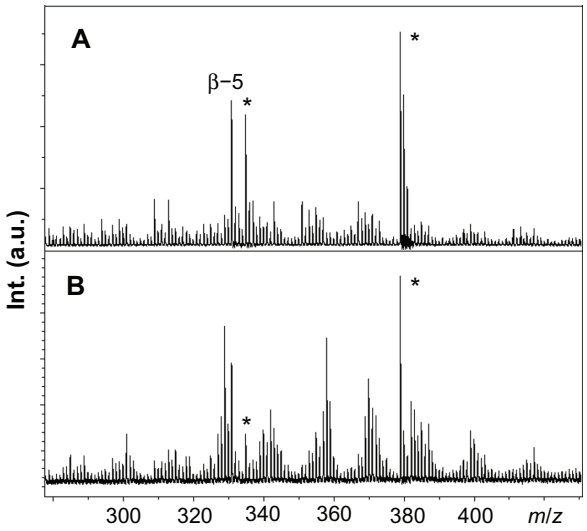

Figure 5. Reflector-positive ion mode MALDI-TOF spectra in the *m/z* 280–420 range of depolymerized Miscanthus lignin FAL80 (**A**) and ammonia lignin AL (**B**) with α-CD/CHCA.

phenylpropane units. For FAL107 and FAL80, the spectral patterns were similar with efficient ionization up to about 550 Da, smaller than molecular weights measured using HPSEC.

Conversely, MS signal for ammonia lignin AL was exploitable up to about *m/z* 800 whatever the initial concentration of AL in acetone (from 1.2 g/L to 0.05 g/L) used for sample preparation. Comparison of molecular weight estimated by MALDI-TOF-MS with M_p measured using HPSEC was quite hazardous due to required acetylation of lignin for size-exclusion analysis. Below 200 Da, the peaks distribution was sensibly identical to FAL107. The oligomeric distribution between *m/z* 280–420 was, however, quite dissimilar, with a peak located at *m/z* 331 combined with another major signal at *m/z* 359 (Fig. 5B). While the first signal was associated with residual β-5 linkages with a carboxylic acid end-group, the second was related to the occurrence of the β-O-4 dilignol substructure with an unsaturated alcohol end-group and fitted our 2D NMR investigations. The experimental signal at *m/z* 329 was assigned to the $[C_{19}H_{21}O_5]^+$ adduct, most probably formed by neutral loss of "CH_2O" from the cinnamyl-alcohol substructure. Identification of specific sequences of peaks with Δm = 12–14 mass units was obvious for the β-O-4 dimers and was related to the different number of carbon atoms or CH_2 groups present in the molecule. The same sequence for the trimers was also recorded, for instance, at *m/z* 521, 535, 548. A gap of Δm = 16–18 mass units was underlined between two "adjacent" peaks sequences, and underlined different numbers of oxygen atoms or hydroxyl groups in the molecule. This observation translated the existence of different linkages between the constitutive phenylpropane monomers. Such regularity in the sequence of peaks was not observed for acid extracted fractions FAL107 and FAL80. Characteristic fragments were also detected at *m/z* 669, 685 and 705 and were correlated with tetramers. As proposed by Yoshioka et al,[15] coherent mass intervals of 178 and 196 Da for β-5 and β-O-4 bonds between two guaiacyl moieties were notably evidenced between respectively *m/z* 521–343 and *m/z* 685–489. However, due to heterogeneity in repetitive monolignol units and linkages, this approach was not an accurate diagnostic. When switching the polarity of the acceleration potential, the spectrum of AL was less resolved, with available *m/z* ranging from 100 to only 350. Only $[M-H]^-$ lignin fragments at *m/z* 341, 329 and 314 characteristic of β-5 linkage were evidenced. No improvement was achieved using a different AL concentration for sample preparation, nor using DHB or CHCA only as the matrix.

Our MALDI-TOF-MS protocol with α-CD/CHCA was then sampled for the study of other herbaceous lignins, namely Switchgrass lignins extracted under acidic (SG-107) or alkali (SG-AL) conditions (Fig. 6). SG-107 was characterized by an M_p value of 1600. MALDI MS spectrum in the positive ion mode, using α-CD/CHCA matrix system, revealed a peak distribution ranging from *m/z* 100 to about 900. No improvement was recorded when using either DHB or CHCA in combination with an ionizing agent (NaCl, KCl or TFA), nor was there any improvement while ionizing in the negative mode. Between 200 and 400 Da, spectral pattern was similar to acid-extracted Miscanthus lignin FAL107, evidencing the formation of specific phenylcoumaran β-5 linkages (*m/z* 315, 331 and 361) in dimeric species. Below 200 Da, signals detected at *m/z* 151 and 181 were classically related to G and S units.

For SG-AL, the best quality spectrum was obtained in the positive ion mode using CHCA alone, instead of α-CD/CHCA, and with an extraction delay of 20 ns. A suspension of SG-AL in acetone, dried between two layers of CHCA, offered convenient results. While the mass range covered by MALDI-TOF-MS

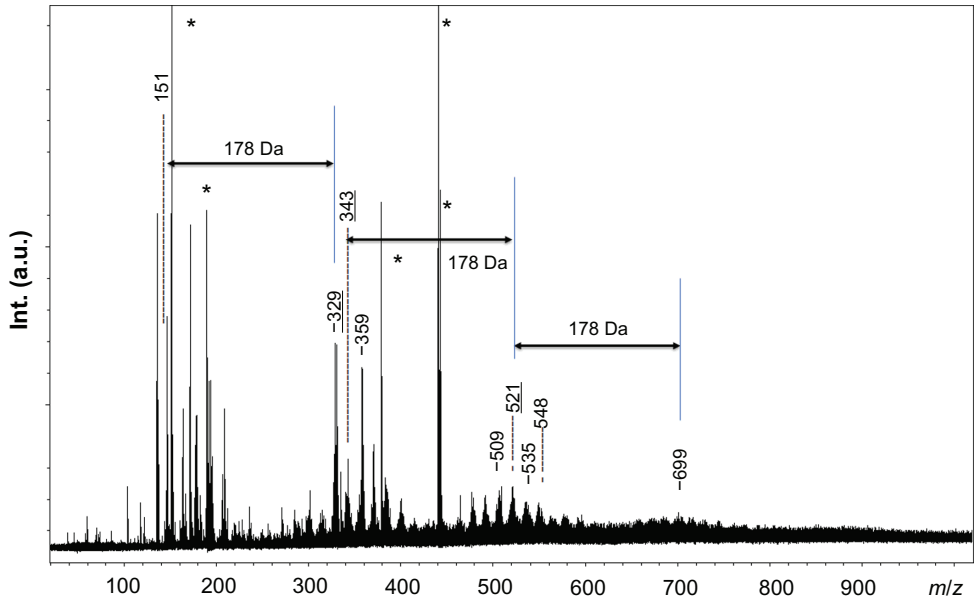

Figure 6. Positive-ion mode MALDI-TOF-MS spectrum of Switchgrass lignin recovered after ammonia soaking (SG-AL).

was detected between 100 and 1000 Da, the M_p value for THF-soluble SG-AL lignin was measured at 500 Da using HPSEC-RI, evidencing good agreement between both techniques. The mass spectrum of SG-AL, disturbed by the presence of matrix-related peaks, showed a fine oligomeric cluster distribution ranging from monomers (about 180 Da) to tetramers (about 750 Da). In this sample, the hyperfine structure of the spectrum highlighted peaks separated by Δm = 12–14 mass units (ie, *m/z* 509, 521, 535 and 548) characteristic of different numbers of carbon atoms or CH_2 in alkyl pending chains. Dual peaks at *m/z* 329 and 359 were recorded for the dimers and corroborated the occurrence of β-O-4 bonding patterns as mentioned for Miscanthus AL lignin. The peak at *m/z* 343 demonstrated the existence of β-5 linkages. Mass increments of 178 Da were measured, notably, between selected peaks and showed the incorporation of a G unit through β-5 bond sequences. These MALDI-TOF-MS results support our previous NMR investigations.[5]

Conclusions

Our results demonstrated that MALDI-TOF-MS was able to provide useful information on the oligomeric distribution of different types of herbaceous lignins containing mainly guaiacyl and syringyl moieties. Involvement of α-cyano-4-hydroxycinnamic acid/α-cyclodextrin as the matrix system allowed efficient sample ionization and suppression of matrix-related signals.

This technique could discriminate between samples extracted using ammonia-soaking and formic/acetic acid conditions with the occurrence of specific "fingerprints", being mass peaks in the 100–800 Da region. Ionization in the positive ion mode was found to be very efficient for low molecular weight samples, and conveniently highlighted β-5 bonds between phenylpropane units, notably at *m/z* 331 and 314. Conversely, ionization in the negative mode also seemed efficient for higher molecular weight ammonia lignins and allows visualization of β-O-4 aryl ether linkages. Discrepancy between molar masses estimated using HPSEC-RI and MALDI-TOF-MS was obvious for samples having a polystyrene-equivalent molar mass superior to 1000, as estimated by size-exclusion chromatography. Although it could provide fast and reliable information on the nature of linkages in a given lignin sample, without proceeding to time-consuming traditional NMR analysis, the MALDI technique remains explorative.

Author Contributions

Conceived and designed the experiments: AR, CV, MS, BW, MP. Analyzed the data: AR, CV, MS, BW, MP. Wrote the first draft of the manuscript: AR, CV, MS, BW, MP. Contributed to the writing of the manuscript: AR, CV, MS, BW, MP.

Agree with manuscript results and conclusions: AR, CV, MS, BW, MP. Jointly developed the structure and arguments for the paper: AR, CV, MS, BW, MP. Made critical revisions and approved final version: AR, CV, MS, BW, MP. All authors reviewed and approved of the final manuscript.

Funding

This work was financially supported by the Walloon Region (TECHNOSE Excellence Research Program, project number 716757). The authors also thank the "Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture" (FRIA) for a fellowship to M. Simon.

Competing Interests

Author(s) disclose no potential conflicts of interest.

Disclosures and Ethics

As a requirement of publication author(s) have provided to the publisher signed confirmation of compliance with legal and ethical obligations including but not limited to the following: authorship and contributorship, conflicts of interest, privacy and confidentiality and (where applicable) protection of human and animal research subjects. The authors have read and confirmed their agreement with the ICMJE authorship and conflict of interest criteria. The authors have also confirmed that this article is unique and not under consideration or published in any other publication, and that they have permission from rights holders to reproduce any copyrighted material. Any disclosures are made in this section. The external blind peer reviewers report no conflicts of interest.

References

1. Johnson DK, Chornet E, Zmierczak W, Shabtai J. Convers ion of lignin into a hydrocarbon product for blending with gasoline. *Fuel Chemistry Division Preprints*. 2002;47:380–1.
2. Octave S, Thomas D. Biorefinery: Toward an industrial metabolism. *Biochim*. 2009;91:659–64.
3. Zakzeski J, Bruijnincx PCA, Jongerius AL, Weckhuysen BM. The catalytic valorization of lignin for the production of renewable chemicals. *Chem Rev*. 2010;110:3552–99.
4. Ralph J, Lundquist K, Brunow G, et al. Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochem Rev*. 2004;3: 29–60.
5. Vanderghem C, Richel A, Jacquet N, Blecker C, Paquot M. 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;96:1761–70.
6. Capanema EA, Balakshin MY, Kadla JF. A comprehensive approach for quantitative lignin characterization by NMR spectroscopy. *J Agric Food Chem*. 2004;52:1850–60.
7. Boeriu CG, Bravo D, Gosselink RJA, van Dam JE. Characterisation of structure-dependent functional properties of lignin with infrared spectroscopy. *Ind Crops Prod*. 2004;20:205–18.
8. Alesiani M, Proietti F, Capuani S, Paci M, Fioravanti M, Maraviglia B. ^{13}C CPMAS NMR spectroscopic analysis applied to wood characterization. *Appl Magn Res*. 2005;29:177–84.
9. Nielen MW. MALDI time-of-flight mass spectrometry of synthetic polymers. *Mass Spectrom Rev*. 1999;18:309–44.
10. Metzger JO, Bicke C, Faix O, et al. Matrix-assisted laser desorption mass spectrometry of lignins. *Angew Chem Int Ed*. 1992;31:762–4.
11. De Angelis F, Fregonese P, Veri F. Structural investigation of synthetic lignins by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom*. 1996;10:1304–8.
12. Bocchini P, Galletti GC, Seraglia R, Traldi P, Camarero S, Martinez AT. Matrix-assisted laser desorption/ionization mass spectrometry of natural and synthetic lignin. *Rapid Commun Mass Spectrom*. 1996;10:1144–7.
13. De Angelis F, Nicoletti R, Spreti N, Veri F. A new in vitro model of lignin biosynthesis. *Angew Chem Int Ed*. 1999;38:1283–5.
14. Netopilık M, Kratochvıl P. Polystyrene-equivalent molecular weight versus true molecular weight in size-exclusion chromatography. *Polymer*. 2003;44:3431–6.
15. Yoshioka K, Ando D, Watanabe T. A comparative study of matrix- and nano-assisted laser desorption/ionisation time-of-flight mass spectrometry of isolated and synthetic lignin. *Phytochem Anal*. 2011;23:248–53.
16. Pavlásková K, Strnadová M, Strohalm M, Havlícek V, Sulc M, Volný M. Time-dependent oxidation during nano-assisted laser desorption ionization mass spectrometry: a useful tool for structure determination or a source of possible confusion? *Anal Chem*. 2011;83:5661–5.
17. Commercial matrices tested: sinapinic acid; 2,4,6-trihydroxyacétophenone; ferulic acid; 3-hydroxypicolinic acid; and a universal commercial MALDI matrix (combination of 2,5-DHB and sinapinic acid).
18. Yamaguchi S, Fujita T, Fujino T, Korenaga T. Suppression of matrix-related ions using cyclodextrin in MALDI mass spectrometry. *Anal Sci*. 2008;24(11):1497. DOI: 10.2116/analsci.24.1497, JOI: JST.JSTAGE/analsci/24.1497
19. Saito K, Kato T, Tsuji Y, Fukushima K. Identifying the characteristic secondary ions of lignin polymer using ToF-SIMS. *Biomacromolecules*. 2005;6:678–83.
20. Banoub JH, Delmas MJ. Elucidation of the complex molecular structure of wheat straw lignin polymer by atmospheric pressure photoionization quadrupole time-of-flight tandem mass spectrometry. *Mass Spectrom*. 2003;38:900–3.
21. Banoub JH, Benjelloun-Mlayah B, Ziarelli F, Joly N, Delmas M. 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;21:2867–88.

Supplementary Information S1

MALDI MS analysis of synthetic lignin sample DHP

Materials

Synthetic lignin, commonly referred to as dehydrogenative polymerizate (**DHP**), was prepared and purified by dehydrogenative polymerization of coniferyl alcohol in the presence of H_2O_2 with peroxidase according to a published protocol.[1] Solubility of DHP in acetone is estimated at 3.7 g/L at 25 °C. HPSEC-RI of non-acetylated DHP revealed a mono modal mass distribution with a M_p value of 1600.

MALDI-TOF-MS analysis

Our results demonstrated that best results were obtained using either CHCA in 30% aq. acetonitrile with 0.1% TFA or 2,5-DHB in acetone-water (7:3) (Fig. 1). A concentration of 1.0 mg/mL of **DHP** in acetone offered optimal mass resolution. The incident laser intensity was optimized for each matrix to obtain the best signal-to-noise ratios and the best resolution. Less noise and higher molecular weight search scores were obtained using 2,5-DHB instead of CHCA, evidencing an apparent molecular weight distribution for **DHP** ranging from several hundreds to about *m/z* 2200 for 2,5-DHB and *m/z* 1800 for CHCA. For both matrices, clear groups of **DHP**-related oligomers can be distinguished between 600 and 1500 Da, with more or less interference with matrices clusters signals. Regular mass increments of 178.1 ± 0.3 Da were denoted on both MS spectra and revealed the of β-5 bonds between coniferyl alcohol units during the radical polymerization process in batch conditions.[2] Below *m/z* 600 for CHCA and *m/z* 800 for DHB, matrix-related fragments were underlined, hampering accurate lignin peaks attribution for low weight oligomeric distributions. The average signal-to-noise ratio (S/N) of peak at *m/z* 1094 from duplicate runs were calculated and estimated at 76.7 and 22.3, respectively using DHB and CHCA-based systems. For the peak located at *m/z* 358 and attributed to the coniferyl dimer, the S/N values were of 52.0 and 41.0 for **DHP** prepared respectively with CHCA and DHB. These results underlined that CHCA provided a slight sensitivity enhancement for low mass range

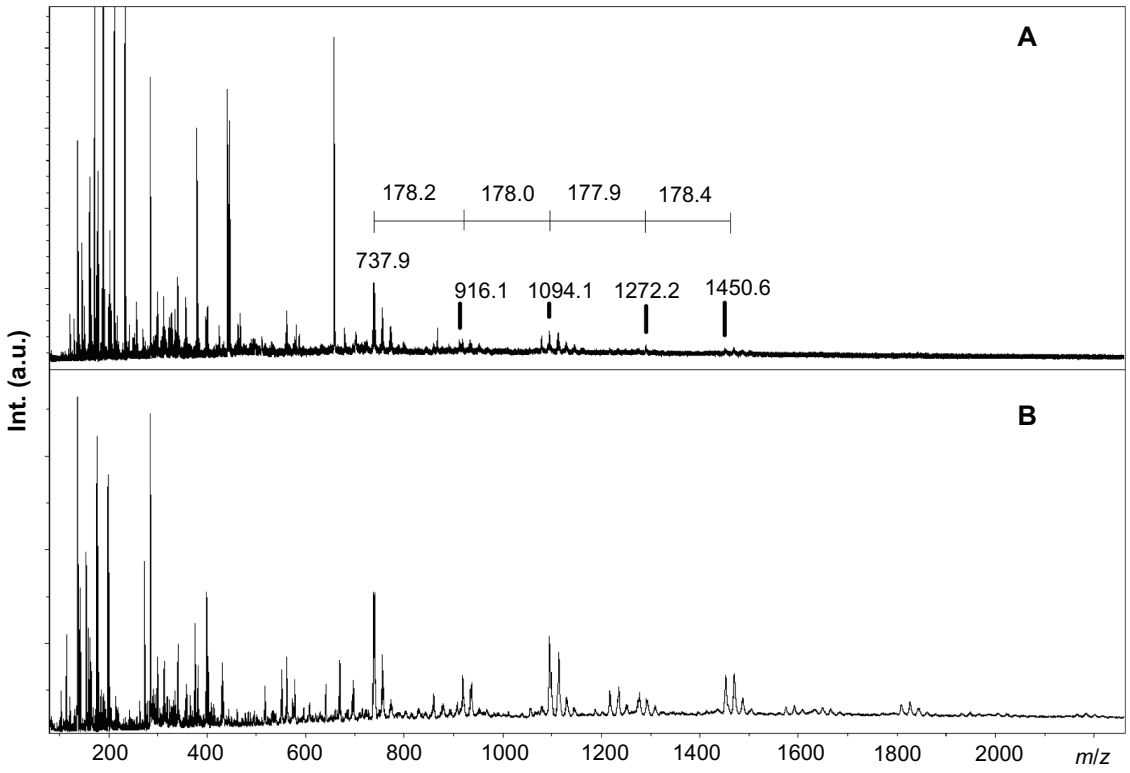

Figure 1. Positive ion mode MALDI-TOF MS spectra of **DHP** with CHCA/TFA (**A**) or 2,5-DHB/NaCl (**B**).
Notes: Samples preparation: 0.5 µl of **DHP** sample (1.0 mg/mL in acetone) between two deposits of 0.5 µL of CHCA (10.0 mg/mL in 30% aq. acetonitrile + TFA 0.1 vol.%) or 2,5-DHB (10.0 mg/mL in 30% aq. acetone). Pulsed extraction delay, 200 ns.

investigations, conversely to DHB which was preferred when exploring the higher mass region. This peak at 358 Da was undoubtedly assigned to the phenylcoumaran (β-5) dimer of coniferyl alcohol with a pendant unsaturated alcohol end-group by comparison with an authentic synthetic sample.[2] More accurate description of the fragmentation pathway of **DHP** was hampered by unwanted signals below 800 Da for 2,5-DHB and 600 Da for CHCA arising from the decomposition of these matrices under MALDI soft ionization conditions. No appreciable results were collected when applying the negative ionization mode.

References

1. Kirk TK. Brunow G. *Methods in Enzymology*. In: Wood WA, Kellogg ST, editors. Academic Press, San Diego; 1988;161:65–73.
2. Synthesis and physico-chemical characterization of model β-5 dilignol compound are in agreement with: Chioccara F, Poli S, Rindone B, et al. *Acta Chem Scand*. 1993;47:610–6.