CrossMark

Characterisation of Authentic Lignin Biorefinery Samples by Fourier Transform Infrared Spectroscopy and Determination of the Chemical Formula for Lignin

Duy Michael Le[1,2] · Anders Damgaard Nielsen[2] · Hanne R. Sørensen[1] · Anne S. Meyer[2]

© Springer Science+Business Media, LLC 2017

Abstract Efficient methods for lignin characterisation are increasingly important as the field of lignin valorisation is growing with the increasing use of lignocellulosic feedstocks, such as wheat straw and corn stover, in biorefineries. In this study, we characterised a set of authentic lignin biorefinery samples in situ with no prior purification and minimal sample preparation. Lignin chemical formulas and lignin Fourier transform infrared (FTIR) spectra were extracted from mixed spectra by filtering out signals from residual carbohydrates and minerals. From estimations of C, H and O and adjustment for cellulose and hemicelluloses contents, the average chemical formula of lignin was found to be $C_9H_{10.2}O_{3.4}$ with slight variations depending on the biomass feedstock and processing conditions (between $C_9H_{9.5}O_{2.8}$ and $C_9H_{11.1}O_{3.6}$). Extracted FTIR lignin spectra showed many of the same characteristic peaks as organosolv and kraft lignin used as benchmark samples. Some variations in the lignin spectra of biorefinery lignin residue samples were found depending on biomass feedstock (wheat straw, corn stover or poplar) and on pretreatment severity, especially in the absorbance of bands at 1267 and 1032 cm^{-1} relative to the strong band at ~1120 cm^{-1}. The suggested method of FTIR spectral analysis with adjustment for cellulose and hemicellulose is proposed to provide a fast and efficient way of analysing lignin in genuine lignin samples resulting from biorefineries.

Introduction

Lignin characterisation is a growing field of interest, since current processes, involving lignocellulosic biomass feedstocks, such as wheat straw, corn stover and various wood species, produce a lignin residue as an unavoidable by-product. Increasing interest has also risen regarding valorisation of lignin [1, 2], which constitutes between 15 and 30% of many lignocellulosic biomasses [3]. Robust tools for characterising lignin are a crucially important prerequisite for rational, knowledge-based lignin valorisation. Fortunately, a wide range of techniques exist and are utilised for characterising lignin, including Fourier transform infrared spectroscopy (FTIR), size exclusion chromatography (SEC) and nuclear magnetic resonance spectroscopy (NMR) [4].

However, characterisation of authentic lignin residue samples from biorefinery processes remains a challenge. Firstly, lignin residues from biorefineries are usually crude, water-insoluble products, where lignin is associated with various other biomass components, such as cellulose, hemicellulose and mineral elements [5]. Characterisation by infrared spectroscopy of such lignin residues is challenged by the many strong bands of residual non-lignin compounds (especially cellulose and amorphous silica) overlapping the characteristic lignin bands in the FTIR spectra [6–8].

Another challenge is that most analytical procedures such as NMR, LC–MS and size-exclusion are liquid state analytical methods requiring lignin to stay stable in solution. Prior to NMR analysis, lignin is often ball mill ground followed by solubilisation and isolation in aqueous dioxane (96%) [9].

✉ Anne S. Meyer
am@kt.dtu.dk

[1] DONG Energy, Kraftværksvej 53, 7000 Fredericia, Denmark

[2] Center for Bioprocess Engineering, Department of Chemical and Biochemical Engineering, Technical University of Denmark, 2800 Lyngby, Denmark

However, the structure of isolated lignin might not resemble the native lignin structure since mainly low molecular weight lignin is solubilised [10]. Therefore, in situ NMR methods have been developed for solubilising the whole biomass sample, e.g. in a bisolvent system of *N*-methylimidazole (NMI) and dimethyl sulfoxide (DMSO) [11], but such treatment also causes modification of the lignin structure by acetylation.

For size exclusion chromatography, lignin is often either derivatised by acetylation or acetobromination and solubilised in tetrahydrofuran (THF) [10, 12, 13] or solubilised by DMSO/water or dimethylacetamide (DMAc) using a catalyst [14, 15]. However, the methods are developed for technical lignin from the pulping industry (e.g. kraft lignin or organosolv lignin) or isolated/purified lignin and are not necessarily directly applicable on crude biorefinery hydrolysis lignin, which is much more heterogeneous in composition [5]. Kraft lignin originates from an alkaline pulping process which conventionally involves alkaline heat treatment at around 170 °C in sodium hydroxide and sodium sulphide followed by precipitation of lignin from the solution [16]. Kraft lignin can therefore easily be solubilised by resuspension in alkaline solutions. Organosolv lignin is a relatively pure lignin product obtained by extraction in solutions of organic solvents and water (often ethanol or acetic acid) [17]. It is therefore also much easier to solubilise than biorefinery hydrolysis lignin residues, which are merely a crude solid by-product from the biorefinery process.

In this study, we aim to characterise authentic lignin residue samples from biorefineries. This is done by FTIR spectroscopy and elemental analysis for calculation of the chemical formula for lignin. To avoid modification of the lignin structure prior to analysis, crude lignin residue samples were analysed directly without sample pretreatment and, subsequently, signals from contaminants (cellulose, hemicellulose, and silica) were mathematically filtered out.

Material and Methods

Lignin Sample Preparation

Five authentic lignin residue samples (WS, WS-hE, WS-hPHE, CS, POP; further details in Table 1) were obtained from the Inbicon pilot plant (Skærbæk, Denmark). Pretreatment conditions were adjusted for the different biomasses (WS was wheat straw lignin residue, CS was corn stover lignin residue and POP was poplar lignin residue) to obtain approximately the same glucan conversion after fermentation. This meant that POP had been subjected to higher pretreatment severity and enzyme dosage than the others. Furthermore, the effects of pretreatment severity and enzyme dosage on the lignin residue were studied by comparing WS

Table 1 Overview of lignin residue samples prepared at the Inbicon pilot plant

Sample name	Pretreatment severity	Enzyme loading	Biomass
WS	Low	Low	Wheat straw
WS-hE	Low	High	Wheat straw
WS-hPhE	High	High	Wheat straw
CS	Low	Low	Corn stover
POP	High	High	Poplar

Low pretreatment severity is $\log(R_o) < 4$, while high pretreatment severity is $\log(R_o) \geq 4$ (calculated according to [18]). Low-enzyme loading was <200 mg/g glucan and high-enzyme loading was ≥200 mg/g glucan. WS corresponds to wheat straw lignin residues; CS corresponds to corn stover lignin residue, while POP corresponds to poplar lignin residue. hE designates lignin residue from a high-enzyme dosage process, while hP designates lignin residue from a high pretreatment severity process

with WS-hE (high-enzyme dosage) and WS-hPhE (high pretreatment severity and high-enzyme dosage).

All Inbicon process samples had been dried at 50 °C for at least 24 h and milled to <1 mm in an IKA MF10 grinder. Indulin AT (KRAFT) was a mixed softwood kraft lignin supplied by MeadWestvaco (North Charleston, USA). OS was a mixed hardwood organosolv lignin from Sigma-Aldrich (St. Louis, USA).

Chemical Composition of Lignin Samples

All samples were analysed for chemical composition by methods based on the standard NREL analytical procedures [19]. The analysis of all samples included dry matter and ash content determination and strong sulphuric acid hydrolysis (72% H_2SO_4) for structural carbohydrates and lignin.

Elemental Analysis

As preparation for X-ray fluorescence (XRF) spectrometry, 3 g milled sample was pressed to pellets of 40 mm in diameter under 20 t in an Atlas 25T Manual Hydraulic Press (Specac, Orpington, UK). XRF spectrometry was performed using a WD-XRF Supermini 200 (50 kV, 200 W Pd-anode, Rigaku, Tokyo, Japan) in helium atmosphere. Initially, a long continuous scan using a LiF(200) crystal at a scan speed of 10°/min was performed to quantify the heavy elements (Ti to U). All detected elements were automatically annotated, and the net intensities were calculated by subtracting the background intensity (automatically measured in a valley adjacent to the peak) from the peak intensity. The long continuous scans were followed by separate scans for each of the lighter elements. Si, P, S, Cl, K and Ca were analysed using a PET crystal at a scan speed of 10°/min, Al using a PET crystal at a scan speed of 4°/min and Na and Mg using a RX25 crystal at a scan speed of 0.5°/min.

Table 2 Chemical composition of lignin residues

	Lignin (%)	Glucose (%)	Xylose (%)	Arabinose (%)	Acetate (%)	Ash (%)
OS	87.3 ± 3.3 a	0.11 ± 0.02 f	0.2 ± 0.1 e	0.04 ± 0.02 e	0.16 ± 0.05 d	0.00 ± 0.00 d
POP	72.6 ± 0.5 b	13.2 ± 0.1 c	2.2 ± 0.0 c	n.d.	0.75 ± 0.02 c	3.42 ± 0.03 c
KRAFT	72.6 ± 2.9 b	0.15 ± 0.01 f	1.1 ± 0.1 d	0.31 ± 0.02 c	0.16 ± 0.01 d	4.13 ± 0.06 c
WS-hPhE	71.2 ± 1.1 b	3.54 ± 0.07 e	2.2 ± 0.0 c	0.25 ± 0.00 d	0.55 ± 0.19 c	13.2 ± 0.1 a
WS-hE	59.7 ± 0.4 c	9.89 ± 0.01 d	6.5 ± 0.1 a	0.67 ± 0.00 a	1.10 ± 0.04 b	11.1 ± 0.3 b
WS	52.8 ± 0.3 d	21.7 ± 0.2 a	5.8 ± 0.0 b	0.60 ± 0.01 b	1.47 ± 0.00 a	10.2 ± 0.3 b
CS	48.9 ± 0.2 d	19.5 ± 1.0 b	6.4 ± 0.4 a	0.62 ± 0.01 b	1.25 ± 0.09 ab	13.2 ± 1.5 a

Glucose (Glc), xylose (Xyl) and arabinose (Ara) and acetate contents are presented as dehydrated weight percentages of dry biomass samples. Values are presented as mean ± standard deviation, and different lowercase letters in each column indicate groups of statistically significant different values according to Tukey's HSD test (p value <0.05). WS corresponds to wheat straw lignin residues; CS corresponds to corn stover lignin residue, while POP corresponds to poplar lignin residue. hE designates lignin residue from a high-enzyme dosage process, while hP designates lignin residue from a high pretreatment severity process. Sample details of lignin residues are found in Table 1

n.d. not detected

In the software ZSX (version 7.42, Rigaku, Tokyo, Japan), Scan Quantitative X (SQX) was utilised to calculate mass% from net intensities for each detected element using a built-in sensitivity library. All measured intensities were corrected for absorption by the helium atmosphere. Absorption/enhancement effects between all measured elements were corrected by the SQX method using built-in correction coefficients. All detected elements were included in these calculations even though only Na, Mg, Al, Si, P, S, Ca, K and Fe results are reported. C, H and N contents measured by the method described below were included in the calculation for quantitative estimates of element concentrations. O was used as a balance for the calculations.

C, H and N contents were measured using a standardised procedure on a Euro EA 3000 Elemental Analyser (EuroVector, Pavia, Italy). O contents in the organic part of samples were calculated by assuming that it constituted the remaining part of the ash-free biomass when C, H, N and S contents had been subtracted. All samples were measured in four replicates and outliers excluded.

Chemical Formula of Lignin

The actual chemical formula of lignin was determined based on the C, H and O contents measured in the lignin samples. However, since some of the C, H and O originated from cellulose and hemicellulose the contribution from these were subtracted before estimating the actual chemical formula for lignin. C content of lignin was calculated by isolating C_{lignin} in the following equation

$$C_{sample} = C_{lignin} + C_{cellulose} + C_{hemicellulose},$$

where C_{sample} was the C content (by weight) measured in the given sample, and $C_{cellulose}$ and $C_{hemicellulose}$ were the C contents (by weight) originating from cellulose and hemicellulose, respectively. $C_{cellulose}$ and $C_{hemicellulose}$ were estimated from the monosaccharide compositions. For the calculations, the elemental composition used for cellulose was 44% C, 6% H and 49% O (by weight) and for hemicellulose components xylose/

Table 3 Concentrations of carbon (C), hydrogen (H), nitrogen (N), sulphur (S) and oxygen (O) in the lignin residues measured on dry samples and calculated on ash-free basis

	C (%)	H (%)	N (%)	S (%)	O (%)
OS	66.2 ± 1.9 a	5.9 ± 0.2 a	0.01 ± 0.01 d	0.02 ± 0.002 f	28.0 ± 2.1 c
POP	59.4 ± 1.6 bcd	5.7 ± 0.2 a	1.08 ± 0.13 bc	0.11 ± 0.003 e	33.7 ± 1.7 b
KRAFT	63.9 ± 1.4 ab	5.8 ± 0.2 a	0.62 ± 0.16 c	1.75 ± 0.012 a	27.9 ± 1.4 c
WS-hPhE	60.3 ± 4.1 bc	5.9 ± 0.4 a	1.98 ± 0.16 a	0.23 ± 0.002 c	31.6 ± 4.4 bc
WS-hE	56.6 ± 1.3 cde	6.2 ± 0.1 a	1.95 ± 0.34 a	0.22 ± 0.002 c	35.0 ± 1.6 ab
WS	52.9 ± 2.5 e	5.9 ± 0.4 a	1.16 ± 0.17 b	0.15 ± 0.003 d	39.9 ± 2.8 a
CS	55.0 ± 0.4 de	6.0 ± 0.1 a	2.26 ± 0.34 a	0.26 ± 0.001 b	36.5 ± 0.8 ab

Values are presented as mean ± standard deviation, and different lowercase letters in each column indicate groups of statistically significant different values according to Tukey's HSD test (p value <0.05). WS corresponds to wheat straw lignin residues; CS corresponds to corn stover lignin residue, while POP corresponds to poplar lignin residue. hE designates lignin residue from a high-enzyme dosage process, while hP designates lignin residue from a high pretreatment severity process. Sample details of lignin residues are found in Table 1

Table 4 Contents of main mineral ash-forming elements

	Na (%)	Mg (%)	Al (%)	Si (%)	P (%)	Cl (%)	K (%)	Ca (%)	Fe (%)
OS	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	0.02 ± 0.002 f	0.02 ± 0.00 e	0.004 ± 0.001 e
POP	0.37 ± 0.01 b	0.03 ± 0.00 d	0.04 ± 0.000 d	0.12 ± 0.003 e	0.10 ± 0.002 d	0.01 ± 0.001 d	0.1 ± 0.00 e	0.91 ± 0.01 c	0.07 ± 0.002 c
KRAFT	1.97 ± 0.04 a	0.04 ± 0.00 cd	0.01 ± 0.003 e	0.03 ± 0.005 f	0.00 ± 0.001 e	0.01 ± 0.000 d	0.2 ± 0.00 d	0.02 ± 0.00 e	0.01 ± 0.000 e
WS-hPhE	0.34 ± 0.00 b	0.03 ± 0.01 cd	0.04 ± 0.002 cd	5.64 ± 0.00 a	0.12 ± 0.002 c	0.06 ± 0.002 c	0.3 ± 0.00 c	0.52 ± 0.01 d	0.06 ± 0.002 d
WS-hE	0.37 ± 0.00 b	0.05 ± 0.01 bc	0.05 ± 0.002 c	4.36 ± 0.01 b	0.19 ± 0.004 b	0.08 ± 0.000 b	0.4 ± 0.01 b	0.51 ± 0.00 d	0.06 ± 0.001 d
WS	n.d.	0.07 ± 0.01 b	0.08 ± 0.000 b	3.83 ± 0.06 c	0.09 ± 0.003 d	0.06 ± 0.001 c	0.3 ± 0.00 c	1.27 ± 0.01 b	0.12 ± 0.000 b
CS	0.05 ± 0.01 c	0.32 ± 0.00 a	0.46 ± 0.01 a	3.58 ± 0.01 d	0.26 ± 0.005 a	0.23 ± 0.008 a	0.6 ± 0.00 a	1.32 ± 0.00 a	0.34 ± 0.000 a

Values are presented as mean ± standard deviation, and different letters in each column indicate groups of statistically significant different values according to Tukey's HSD test (p value <0.05). WS corresponds to wheat straw lignin residues; CS corresponds to corn stover lignin residue, while POP corresponds to poplar lignin residue. hE designates lignin residue from a high-enzyme dosage process, while hP designates lignin residue from a high pretreatment severity process

n.d. not detected

arabinose and acetic acid in dehydrated form were 45% C, 6% H, 48% O and 56% C, 7% H and 37% O (by weight), respectively. C_{lignin} was converted from grammes per 100 g sample into moles per 100 g sample using the molar mass of C (12.01 g/mol) and thereafter scaled to C_9. The calculation was performed analogously for H and O yielding a lignin chemical formula for each sample ($C_9H_xO_y$). S could in some cases be included in such lignin chemical formula, since S can be incorporated in the lignin structure during processing. However, S in biorefinery lignin is primarily expected to be present in proteins and not as a structural component of lignin, so in the lignin chemical formulas calculated in this work, S was not included.

FTIR Spectroscopy

The dry, milled lignin samples were applied onto a Nicolet 6700 FTIR spectrometer (Thermo Fisher Scientific, Waltham, USA) equipped with a temperature adjustable ATR accessory (Golden Gate diamond ATR accessory, Thermo Fisher Scientific, Waltham, USA) equilibrated at 30 °C. ATR-FTIR spectra were obtained between 1360 and 750 cm^{-1} at 2 cm^{-1} resolution and averaging 100 scans per sample. A background spectrum was recorded immediately before analysis of the samples at the same parameters but with 200 scans per sample. Preprocessing of the spectra involved background correction by subtracting the minimum value in each spectrum and normalisation to the maximum absorbance of each spectrum.

The main components of the biorefinery lignin residue FTIR spectra were expected to be amorphous silica, cellulose and lignin. Therefore, multiple linear regression was performed to model the full spectra of lignin samples by linear combination of the pure spectra of cellulose (Avicel PH-101, obtained from Sigma-Aldrich, St. Louis, USA), amorphous silica (99.8% purity, obtained from Sigma-Aldrich, St. Louis, USA), a purified softwood kraft lignin named BioChoice (> 90% lignin, obtained from Domtar, Montreal, Canada) and hardwood organosolv lignin (OS). The estimated coefficients of BioChoice and OS were subsequently utilised to generate the modelled lignin spectra, which were linear combinations of the BioChoice and OS spectra. These two high purity lignin types were selected for modelling the lignin FTIR spectra, since impurities in less pure lignin samples could cause interferences in the spectra. Furthermore, BioChoice and OS originate from different processes and different biomass feedstocks, so it was expected that they could be used to model different lignin FTIR spectra by linear combination.

PCA

Principal components analysis (PCA) was performed of lignin FTIR spectra by singular value decomposition using R

Table 5 Overall organic chemical formulas were calculated directly from C, H and O contents in Table 3

	Overall organic chemical formula	Estimated lignin chemical formula
OS	$C_9H_{9.5}O_{2.9}$	$C_9H_{9.5}O_{2.8}$
POP	$C_9H_{10.4}O_{3.8}$	$C_9H_{9.7}O_{3.3}$
KRAFT	$C_9H_{9.7}O_3$	$C_9H_{9.6}O_{2.9}$
WS-hPhE	$C_9H_{10.5}O_{3.5}$	$C_9H_{10.3}O_{3.4}$
WS-hE	$C_9H_{11.7}O_{4.2}$	$C_9H_{11.1}O_{3.6}$
WS	$C_9H_{12}O_{5.1}$	$C_9H_{10.9}O_{4.3}$
CS	$C_9H_{11.7}O_{4.5}$	$C_9H_{10.6}O_{3.5}$
Average	$C_9H_{10.8}O_{3.9}$	$C_9H_{10.2}O_{3.4}$

Estimated lignin chemical formulas were calculated by subtracting the C, H and O contents expected to originate from cellulose and hemicellulose constituents before converting from mass to molar basis. Chemical formulas were presented on a C_9-basis. WS corresponds to wheat straw lignin residues; CS corresponds to corn stover lignin residue, while POP corresponds to poplar lignin residue. hE designates lignin residue from a high-enzyme dosage process, while hP designates lignin residue from a high pretreatment severity process. Sample details of lignin residues are found in Table 1

statistical software (3.0.2) [20]. Lignin FTIR spectra were preprocessed prior to PCA by background removal, normalisation and mean centering.

Results and Discussion

Chemical Composition of Lignin Samples

The chemical composition of the lignin samples varied as expected due to the difference in purity of the various lignin samples (Table 2). Technical lignin residues from the pulping process had a high lignin content of 87 and 73% for OS and KRAFT, respectively. Similar lignin contents (73%) were observed for the POP lignin sample, which was also from a low ash wood biomass feedstock. WS and CS lignin samples had lower lignin contents of 53 and 49%, respectively, due to higher ash contents in the original feedstock and the presence of residual carbohydrates after processing. Lignin contents of biorefinery lignin residues depended on the processing parameter, so high severity pretreatments and high-enzyme dosage enzymatic hydrolysis (WS-hPhE) significantly increased the lignin content to 71%, which was comparable with the lignin content of the KRAFT lignin sample. High-enzyme dosage applied to wheat straw pretreated at low pretreatment severity (WS-hE) also increased the lignin content but only to 60% (Table 2).

Glucan contents for KRAFT and OS were both below 0.2%. Biorefinery lignin residues contained between 3.5 and 22% glucan, again being governed by the severity of pretreatment and the enzyme dosage. Expectedly, WS-hPhE with high pretreatment severity and high-enzyme dosage had the lowest glucan

content (3.5%), while the WS lignin residue which had been subjcted to low pretreatment severity and low-enzyme dosage had the highest glucan content (22%). POP, CS and WS lignin residues had glucan contents in the same order of magnitude (13–22%) since the aim was to achieve similar glucan conversion after fermentation for these biomasses.

Contents of hemicellulose constituents (xylose, arabinose and acetate) were also lower in OS and KRAFT lignin samples (0.4 and 1.5%, respectively) than in biorefinery lignin residues (from 3.0% for WS-hPhE to 7.8% for WS). Ash contents of biorefinery lignin residues primarily depended on the biomass feedstock used; hence, POP (poplar as biomass feedstock) had a low ash content of 3.4% compared to 13% for CS (corn stover as biomass feedstock) and 10–13% for biorefinery lignin residues with wheat straw as feedstock. Ash contents were low for both OS and KRAFT (0.0 and 4.1%, respectively).

Elemental Composition

The organic fraction of the biomass as expected mainly consisted of C for all lignin residues (Table 3). Highest C contents were measured in KRAFT and OS (64 and 66%, respectively). The remaining lignin samples had C contents between 53 and 60% with the samples having high lignin contents (POP and WS-hPhE) correspondingly having the highest C contents. H contents were 5.7–6.2% with no statistically significant difference between samples (ANOVA test, p value = 0.327). N contents were relatively low in OS (0.01%) and KRAFT (0.6%) and between 1.1 and 2.3% for the
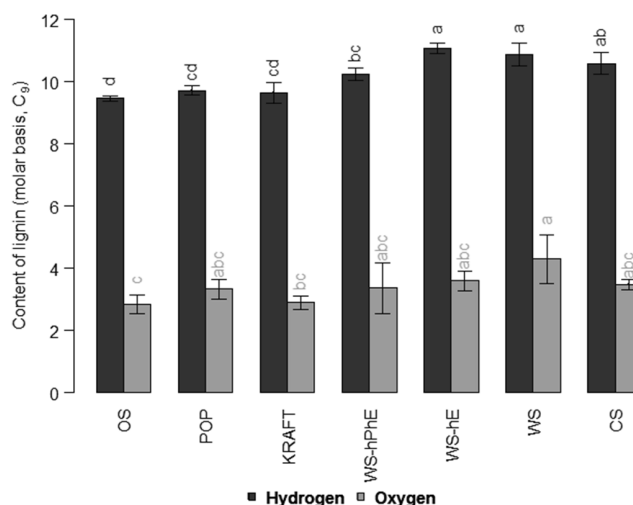

Fig. 1 Estimated molar contents of hydrogen and oxygen in lignin calculated for C_9. *WS* corresponds to wheat straw lignin residues; *CS* corresponds to corn stover lignin residue, while POP corresponds to poplar lignin residue. *hE* designates lignin residue from a high-enzyme dosage process, while *hP* designates lignin residue from a high pretreatment severity process

Fig. 2 Model FTIR spectrum used for modelling lignin spectra. BioChoice lignin was selected as purified softwood kraft lignin and organosolv lignin (OS) as a hardwood lignin. Band characteristics for lignin were annotated with G for band characteristics for guaiacyl units and with S for band characteristics for syringyl units. No letter designation means that the band is not characteristic for neither guaiacyl nor syringyl units. Band assignments are presented in Table 6
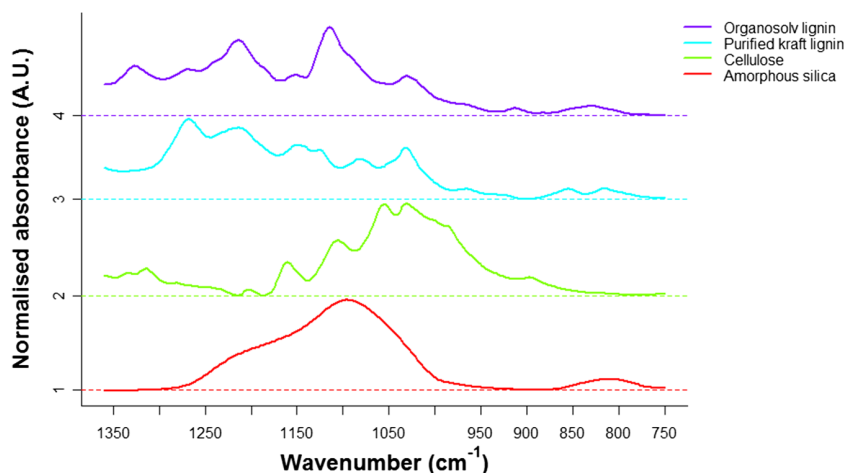
remaining lignin residues. S contents were very low in OS (0.02%) and biorefinery lignin residues (0.16 and 0.26%). KRAFT contained high amounts of S (1.8%) due to utilisation of sulphide during processing of the lignocellulosic biomasses. O contents correlated as expected with carbohydrate contents in the lignin samples, i.e. highest O contents were found in CS and WS (36 and 40%, respectively) compared to 32–35% in WS-hPhE, POP and WS-hE. The lowest O contents were accordingly measured in OS and KRAFT (27.9–28%).

Si was the primary ash-forming constituent in wheat straw lignin residues (WS, WS-hE and WS, hPhE), with Si contents of 3.5–5.6% (Table 4). Thereby, Si constituted 65–80% of all ash-forming elements in wheat straw lignin residues. Si in biomass is present as silica (SiO_2) [21], which is important to keep in mind when performing FTIR analysis, since silica has some significant IR absorption bands overlapping with those of lignin [6, 8]. Ca was also highly abundant in wheat

straw lignin residues (0.5–1.3%), since calcium hydroxide was used for pH adjustment in the process. K is highly abundant in wheat straw biomass feedstock but were found in low amounts relative to Si (0.3–0.4%), since it was solubilised during pretreatment [22]. The ash of KRAFT was predominantly Na (1.97%, which was 85% of all ash-forming elements), which originated from sodium hydroxide and sodium sulphide used in the kraft process. Poplar is a low ash lignocellulosic biomass feedstock, so most of the mineral elements measured in POP was Ca (0.91%), which was added in the biorefining process. OS almost contained no ash, but the small amounts found were K and Ca (0.02 and 0.02%, respectively).

Chemical Formula of Lignin

The generally high levels of impurities in the lignin samples made characterisation of lignin challenging. When determining the chemical formula for lignin samples, the high levels of

Fig. 3 Examples of a FTIR spectrum and the spectrum modelled by linear combination of the four model spectra (Fig. 2). Modelled lignin spectrum was obtained by subtracting the model spectrum representing amorphous silica and cellulose, respectively, from the full spectrum of lignin samples. *WS* = low severity pretreatment and low-enzyme dosage wheat straw lignin residue; *WS-hE* = low severity pretreatment and high-enzyme dosage wheat straw lignin residue; *WS-hPhE* = high severity pretreatment and high-enzyme dosage wheat straw lignin residue
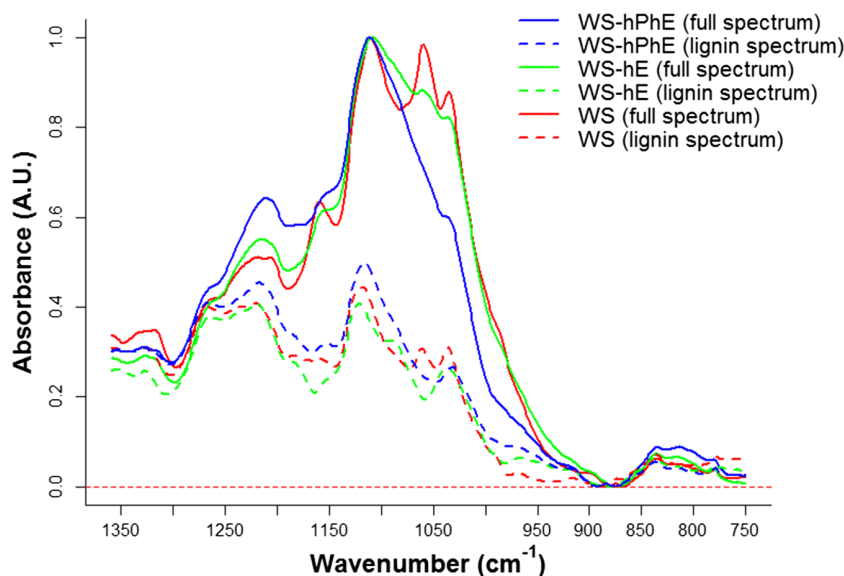
Fig. 4 Lignin FTIR spectra for two lignin samples from the pulping industry. KRAFT is softwood kraft lignin, and organosolv lignin (OS) is a hardwood lignin. Band characteristics for lignin were annotated with *G* for band characteristics for guaiacyl units and with *S* for band characteristics for syringyl units. Bands annotated with *n* were characteristics for neither guaiacyl nor syringyl units. Band assignments are presented in Table 6
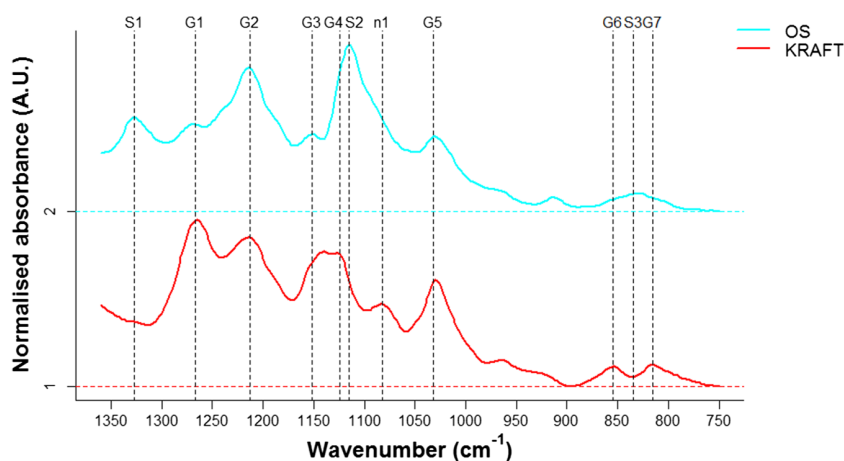
cellulose and hemicellulose made a significant contribution to the C, H and O contents, which had to be corrected for when estimating the chemical formula for lignin in the lignin samples.

However, since the concentrations of cellulose and hemicellulose constituents were known from Table 2, the C, H and O contents measured could be corrected for the C, H and O originating from cellulose and hemicellulose. Other components such as polyphenols/pseudolignin, proteins and fats, were not corrected for, but they were expected to be present at low contents relative to lignin and thereby not influence the measured chemical formulas as much as, e.g. carbohydrates.

Overall organic chemical formulas for the lignin samples and the estimated lignin chemical formulas are presented in Table 5. As expected, chemical formulas for lignin were much more alike after correction for contents of residual carbohydrates. As seen in Fig. 1, WS and WS-hE

lignin residues had similar H contents in the lignin chemical formula (10.9 and 11.1, respectively), which was expected, since they both originated from wheat straw biomass and had been subjected to low severity pretreatment. CS had been subjected to the same low severity pretreatment and had a H content in its lignin chemical formula similar to that of WS and WS-hE (10.6), even though CS was a corn stover lignin residue. High severity pretreatment caused WS-hPhE lignin to have lower H content in its lignin chemical formula compared to WS and WS-hE (10.3), indicating formation of double-bonds or concentrating aromatic rings by cleavage of single-bonds such as ethers. OS, KRAFT and POP also had low H contents in their chemical formulas (9.5–9.7), either due to the processing or because they originated from wood species instead of grasses. O contents in the chemical formulas were all 3–4 with no clear trends (Fig. 1).

Table 6 Assignment of IR bands in lignin spectra of three selected lignin samples

Band no.	Detected band (cm^{-1})	OS	KRAFT	WS	WS-hE	WS-hPhE	CS	POP	Assignment
S1	1327	x		(x)	x	(x)	x	x	C–O stretch of syringyl ring: 1325–1328 cm^{-1} [6, 27]
G1	1267	x	x	x	x	x	(x)	x	C–O stretch of guaiacyl ring: 1265–1269 cm^{-1} [6, 27, 28]
G2	1213	x	x	x	x	x	(x)	x	C–O stretch of guaiacyl ring: 1214–1221 cm^{-1} [6, 27]
G3	1151	x	x			x	(x)		Possibly C–H stretch of guaiacyl ring: 1140–1145 cm^{-1} [6, 27, 28]
G4	1124		x	x	x	(x)	x	(x)	Aromatic C–H deformation of guaiacyl ring: 1125 cm^{-1} [6]
S2	1115	x		x	(x)	x	(x)	(x)	Aromatic C–H deformation of syringyl ring: 1110–1116 cm^{-1} [6, 27]
n1	1082		x		x		x		C–O stretch of secondary alcohols and aliphatic ethers: 1086 cm^{-1} [27, 28]
G5	1032	x	x	x	x	x	x	x	C–O deformation in primary alcohol, C=O stretch, aromatic C–H deformation in guaiacyl ring: 1025–1033 cm^{-1} [27, 28]
G6	854		x				x		C–H bending of guaiacyl units: 855–858 cm^{-1} [6, 27]
S3	835	x		x	x	x	x	(x)	C–H bending of syringyl units: 829–835 cm^{-1} [6]
G7	816		x	(x)			x		C–H bending of guaiacyl units: 817 cm^{-1} [27]

Band no. refers to bands detected in Fig. 4

Fig. 5 Extracted lignin FTIR spectra of three wheat straw lignin residues, which had been subjected to different processing conditions. *WS* = low severity pretreatment and low-enzyme dosage wheat straw lignin residue; *WS-hE* = low severity pretreatment and high-enzyme dosage wheat straw lignin residue; *WS-hPhE* = high severity pretreatment and high-enzyme dosage wheat straw lignin residue
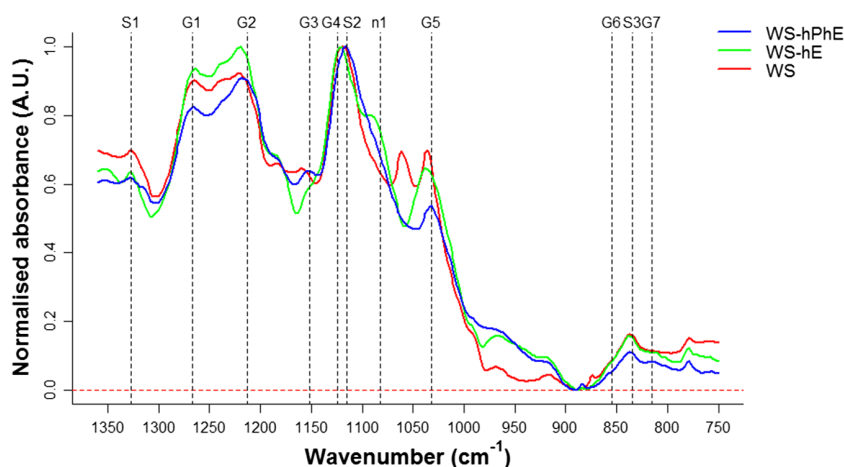
The average chemical formula for lignin in the lignin samples was $C_9H_{10.2}O_{3.4}$, which was close to the lignin chemical formula reported for isolated lignin ($C_9H_{11.3}O_{3.7}$) [23]. However, by the method proposed in this study, no isolation procedure, which might alter the structure and composition of lignin, was needed.

As a supplement for determination of the lignin chemical formula using the method proposed in this study, one could determine the methoxyl group contents in the lignin samples and present the chemical formula in a methoxy form as done by Vázquez et al. [23]. This may provide further information about the lignin structure; however, in any case, further elucidation of the lignin structure requires more advanced procedures such as NMR [9] or combinations of multidimensional NMR and coherent anti-Stokes Raman scattering analyses [3].

Structural Assessment of Lignin by FTIR

FTIR characterisation of non-purified lignin samples can be challenging, since both cellulose and amorphous silica (which were present in high concentrations in most lignin samples in this study, Tables 1 and 2) have many bands overlapping with lignin. This problem could, however, be overcome by first modelling the FTIR spectra by a linear combination of model spectra and then decomposing the spectra. For doing this, two model FTIR spectra of the main contaminants were obtained by analysing pure amorphous silica and cellulose (Fig. 2). Two very different types of lignin were needed for representing the lignin structure and for this a hardwood and a softwood lignin was selected. Purified softwood kraft lignin (BioChoice) with a lignin content above 90% and OS, being a hardwood organosolv lignin, were analysed and used as FTIR model spectra 3 and 4 (Fig. 2). A linear combination of the model spectra 3 and 4 was expected to make a good representation of the lignin structure in the lignin samples.

By multiple linear regressions, the FTIR spectra of lignin samples were modelled using the model spectra (examples seen in Fig. 3). All models were statistically significant ($P < 2.2 \times 10^{-16}$) with adjusted R^2 values above 0.975. By subtracting the model spectra of cellulose and amorphous silica with coefficients according to the estimated model coefficients obtained by multiple linear regressions from the full FTIR spectra of each lignin residue, the modelled lignin

Fig. 6 Extracted lignin FTIR spectra of lignin residues from different biomass feedstocks treated in the Inbicon process. *POP* = poplar lignin residue; *CS* = corn stover lignin residue; *WS* = wheat straw lignin residue
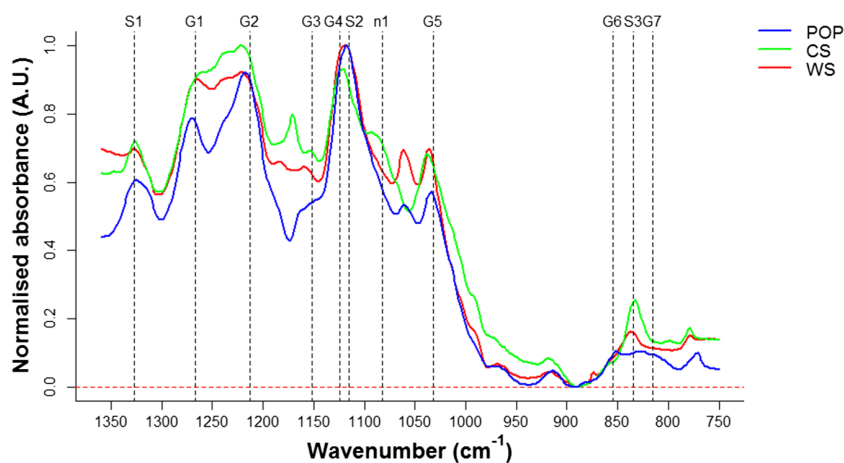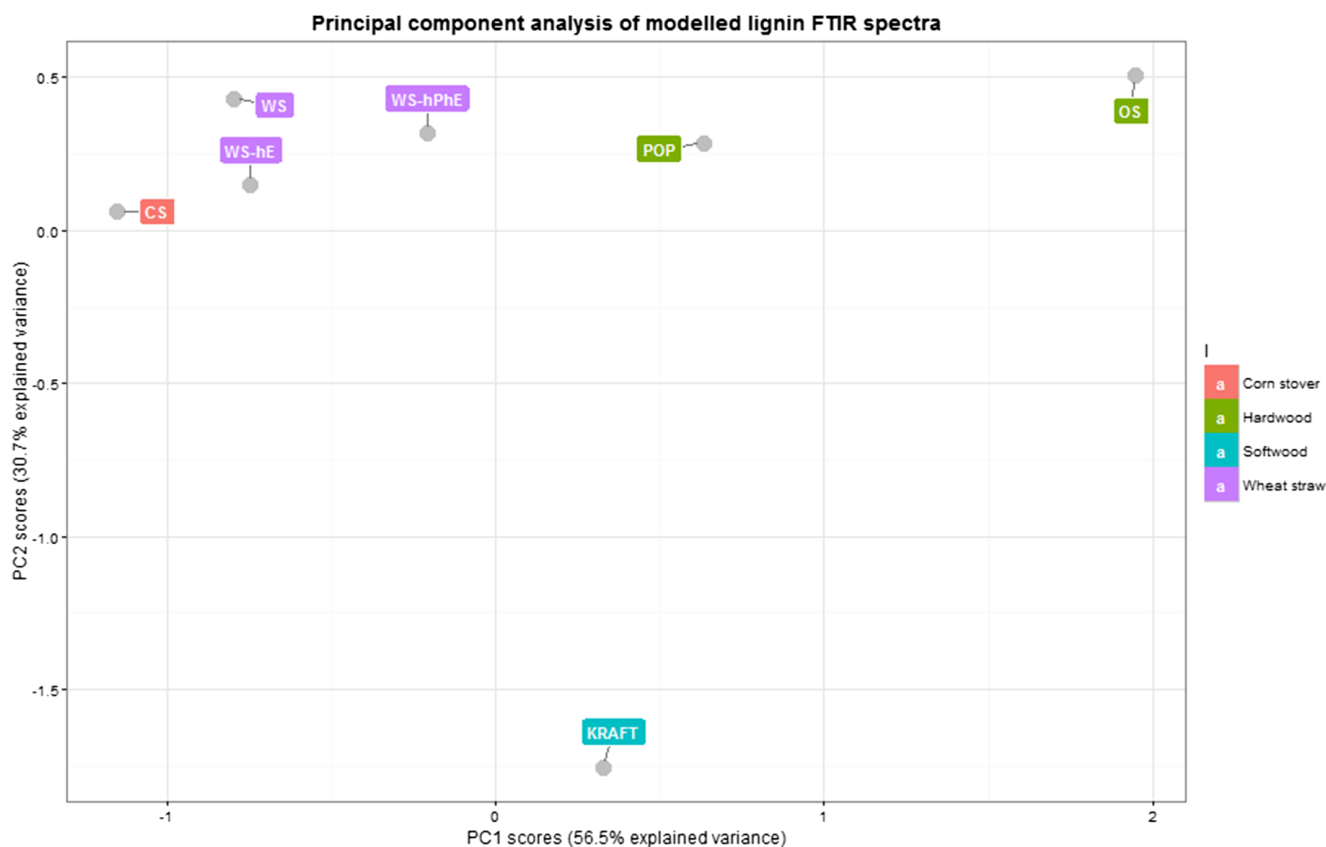
Fig. 7 Scores plot from principal component analysis of lignin FTIR spectra. *WS* corresponds to wheat straw lignin residues; *CS* corresponds to corn stover lignin residue, while *POP* corresponds to poplar lignin residue. *hE* designates lignin residue from a high-enzyme dosage process, while *hP* designates lignin residue from a high pretreatment severity process

spectra were obtained. In Fig. 3, it is seen that the lignin spectra could be extracted from three very different full spectra (WS, WS-hE, WS-hPhE).

Characteristic lignin bands [24, 25] of the extracted lignin spectra were assigned according to the annotation of peaks in the lignin spectra of two lignin samples from the pulping industry: OS and KRAFT (Fig. 4; Table 6). Many bands were generic for the two types of lignin, especially some of those characteristic for guaiacyl units (G1 1267 cm^{-1}, G2 1213 cm^{-1}, G3 1151 cm^{-1}, G5 1032 cm^{-1}). Syringyl units are absent or only present in very low concentrations in softwood lignin [26, 27], and therefore, FTIR band characteristics for syringyl units were only observed in OS (S1 1327 cm^{-1}; S2 1115 cm^{-1}, S3 835 cm^{-1}; Table 6). Other band characteristics for guaiacyl units (G4 1124 cm^{-1}, G6 854 cm^{-1}, G7 816 cm^{-1}) and for general C–O stretching in secondary alcohols and aliphatic ethers (n1 1082 cm^{-1}) were very clearly observed in the KRAFT lignin spectra and not in the OS spectra, since these bands overlapped with strong bands of syringyl units in OS.

The three wheat straw lignin spectra from Fig. 3 were normalised and characteristic lignin bands were assigned in Fig. 5. WS-hPhE lignin had been subjected to a higher pretreatment severity than WS and WS-hE lignin, and it seemed to have altered the lignin structure. The WS-hPhE lignin spectrum was thus different from the others, especially in terms of ratio between absorbance of the G4/S2 band to the G5 and G1 bands, respectively. Also, S1 was very weak in WS-hPhE. WS-hE and WS had been subjected to similar pretreatment conditions, but a higher enzyme dosage was used for WS-hE meaning that less residual cellulose was expected in the lignin residue. In spite of this difference in the enzymatic hydrolysis, WS and WS-hE had quite similar lignin spectra. However, some differences could be observed (Fig. 5). A shoulder on the G4/S2 peak was observed for the WS-hE lignin spectrum at the n1 band, but more notable was the band at 1060 cm^{-1} in the lignin spectrum of the WS lignin residue. This band is usually assigned to cellulose [7], meaning that the cellulose spectrum was not completely filtered out, but it was no longer a dominant band as in the full spectrum (Fig. 3).

The differences in lignin structure of lignin samples from three different biomass feedstocks are seen in the lignin spectra in Fig. 6. Many of the same bands were present in the spectra of POP, CS and WS (Table 6), but some of the characteristic lignin bands (G1 and G2) were more clearly defined in POP. The ratio between G4/S1 and G5 was also larger for POP compared to the others.

Some distinct differences were also observed between the CS and WS lignin spectra. S1 and S3 bands were very strong in the CS lignin spectrum compared to the WS lignin spectrum. Also, in the CS lignin spectrum, the G2 band was strong relative to the G4/S2 band. A clear band between G2 and G3 was observed in the CS lignin spectrum at 1161 cm^{-1}. This was not found in the spectrum of any of the other lignin samples, and it is characteristic for C=O stretching in esters of GSH lignin [28].

PCA

PCA of the lignin spectra was performed to quantify similarities between the lignin spectra. Principal component 1 and 2 (PC1 and PC2) explained 56.5 and 30.7%, respectively, of the variance in the spectra (Fig. 7). OS and KRAFT had high absolute PC1 and PC2 scores, respectively, indicating that their lignin spectra were uncorrelated, which was also why a linear combination of FTIR spectra of a kraft lignin and an organosolv lignin were well suited for modelling lignin spectra of the various lignin samples. Biorefinery lignin residues were well spread in regards to PC1 scores indicating quite different lignin structures, except for WS and WS-hE (PC1 scores of −0.79 and −0.75, respectively) which were both wheat straw lignin residues with similar pretreatment conditions. POP was located further away from the other biorefinery lignin samples in the PCA scores plot (Fig. 7). This was because of differences in the absorbance of G1 and G5, which had high PC1 loadings (data not shown). POP was closer to OS, which was from another process but also from a hardwood feedstock.

Conclusions

Authentic lignin biorefinery residues were characterised by compositional analysis, elemental analysis and Fourier transform infrared spectroscopy (FTIR) and compared to well-studied benchmark lignin products from the pulping industry (Indulin kraft lignin and organosolv lignin). Compositional data was used for extracting the lignin chemical formula and lignin FTIR spectra from mixed spectra and values of residual carbohydrates and minerals. The average chemical formula of lignin was found to be $C_9H_{10.2}O_{3.4}$. Lignin chemical formulas only varied slightly depending on the biomass feedstock and processing conditions. FTIR lignin spectra showed many of the same characteristic peaks as organosolv and kraft lignin, but the lignin spectra of biorefinery lignin residues varied depending on biomass feedstock (wheat straw, corn stover or poplar) and on pretreatment severity. Lignin residues which had been treated with different enzyme dosages but similar pretreatment severities produced similar FTIR lignin spectra in spite of differences in purities. This verified that impurities

in the spectra were successfully filtered out of the spectra. The suggested method provides a fast and efficient way of analysing lignin structures in impure samples with only minimal sample preparation.

References

1. Thakur VK, Thakur MK, Raghavan P, Kessler MR (2014) Progress in green polymer composites from lignin for multifunctional applications: a review. ACS Sustain Chem Eng 2:1072–1092
2. Azadi P, Inderwildi OR, Farnood R, King DA (2013) Liquid fuels, hydrogen and chemicals from lignin: a critical review. Renew Sust Energ Rev 21:506–523
3. Ragauskas AJ, Beckham GT, Biddy MJ et al (2014) Lignin valorization: improving lignin processing in the biorefinery. Science 344:709–719
4. Ghaffar SH, Fan M (2013) Structural analysis for lignin characteristics in biomass straw. Biomass Bioenergy 57:264–279
5. Vishtal AG, Kraslawski A (2011) Challenges in industrial applications of technical lignins. Bio Resources 6:3547–3568
6. Sammons RJ, Harper DP, Labbé N et al (2013) Characterization of organosolv lignins using thermal and FT-IR spectroscopic analysis. Bioresources 8:2752–2767
7. Tsuboi M (1957) Infrared spectrum and crystal structure of cellulose. J Polym Sci 25:159–171
8. Bertaux J, Froehlich F, Ildefonse P (1998) Multicomponent analysis of FTIR spectra; quantification of amorphous and crystallized mineral phases in synthetic and natural sediments. J Sediment Res 68:440–447
9. Wen J-L, Sun S-L, Xue B-L, Sun R-C (2013) Recent advances in characterization of lignin polymer by solution-state nuclear magnetic resonance (NMR) methodology. Materials (Basel) 6:359–391
10. Guerra A, Filpponen I, Lucia LA et al (2006) Toward a better understanding of the lignin isolation process from wood. J Agric Food Chem 54:5939–5947
11. Lu F, Ralph J (2003) Non-degradative dissolution and acetylation of ball-milled plant cell walls: high-resolution solution-state NMR. Plant J 35:535–544
12. Asikkala J, Tamminen T, Argyropoulos DS (2012) Accurate and reproducible determination of lignin molar mass by acetobromination. J Agric Food Chem 60:8968–8973
13. Mansouri N-EE, Salvadó J (2006) Structural characterization of technical lignins for the production of adhesives: application to lignosulfonate, kraft, soda-anthraquinone, organosolv and ethanol process lignins. Ind Crop Prod 24:8–16
14. Ringena O, Lebioda S, Lehnen R, Saake B (2006) Size-exclusion chromatography of technical lignins in dimethyl sulfoxide/water and dimethylacetamide. J Chromatogr A 1102:154–163
15. Gidh AV, Decker SR, Vinzant TB et al (2006) Determination of lignin by size exclusion chromatography using multi angle laser light scattering. J Chromatogr A 1114:102–110
16. Chakar FS, Ragauskas AJ (2004) Review of current and future softwood kraft lignin process chemistry. Ind Crop Prod 20:131–141
17. Xu F, Sun J, Sun R, et al (2006) Comparative study of organosolv lignins from wheat straw. Ind Crop Prod 23:180–193

18. Carvalheiro F, Silva-Fernandes T, Duarte LC, Gírio FM (2009) Wheat straw autohydrolysis: process optimization and products characterization. Appl Biochem Biotechnol 153:84–93

19. Sluiter A, Hames B, Ruiz R et al (2008) Determination of structural carbohydrates and lignin in biomass determination of structural carbohydrates and lignin in biomass. NREL Lab. Anal, Proced

20. Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. J Comput Graph Stat 5:299–314

21. Le DM, Sørensen HR, Knudsen NO, Meyer AS (2015) Implications of silica on biorefineries - interactions with organic material and mineral elements in grasses. Biofuels Bioprod Biorefin 9:109–121

22. Le DM, Sørensen HR, Knudsen NO et al (2014) Biorefining of wheat straw: accounting for the distribution of mineral elements in pretreated biomass by an extended pretreatment-severity equation. Biotechnol Biofuels 7:141

23. Vázquez G, Antorrena G, González J, Freire S (1997) FTIR, 1 H and 13 C NMR characterization of acetosolv-solubilized pine and eucalyptus lignins. Holzforschung 51:158–166

24. Casas A, Alonso MV, Oliet M et al (2012) FTIR analysis of lignin regenerated from *Pinus radiata* and *Eucalyptus globulus* woods dissolved in imidazolium-based ionic liquids. J Chem Technol Biotechnol 87:472–480

25. Pandey KK (1999) A study of chemical structure of soft and hard-wood and wood polymers by FTIR spectroscopy. J Appl Polym Sci 71:1969–1975

26. Nimz HH, Robert D, Faix O, Nemr M (1981) Carbon-13 NMR spectra of lignins, 8. Structural differences between lignins of hard-woods, softwoods, grasses and compression wood. Holzforschung 35:16–26

27. Scholze B, Meier D (2001) Characterization of the water-insoluble fraction from pyrolysis oil (pyrolytic lignin). Part I. PY–GC/MS, FTIR, and functional groups. J Anal Appl Pyrolysis 60:41–54

28. Faix O (1991) Classification of lignins from different botanical origins by FT-IR spectroscopy. Holzforschung 45:21–28