In-depth structural characterization of the lignin fraction of a pine-derived pyrolysis oil

M.B. Figueirêdoa, R.H. Venderboschb, H.J. Heeres*a, P.J. Deussa,*

a Department of Chemical Engineering, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, the Netherlands
b Biomass Technology Group BV, Josink Esweg 34, 7545 PN, Enschede, the Netherlands

ARTICLE INFO

Keywords:
Pyrolytic lignin
Lignin
Chemical characterization
Pyrolysis oil
Pine

ABSTRACT

Pyrolytic lignin (PL) is the collective name of the water-insoluble fraction of pyrolysis oils produced from the fast pyrolysis of lignocellulosic biomass. As the name suggests, PL is composed by fragments derived from lignin, which is the largest natural source of aromatic carbon. Its valorization is of major importance for the realization of economically competitive biorefineries. Nonetheless, the valorization of PL is hindered by its complex structure, which makes the development of tailored strategies for its deconstruction into valuable compounds challenging. In this work, we provide an in-depth analysis of the structural composition of PL obtained from a commercially available pine-derived pyrolysis oil obtained at 500 °C (Empyro B.V., the Netherlands). Molecular weight distribution and thermal stability were accessed by GPC and TGA, respectively, and the monomers present in the PL (≈ 15 wt%) were identified and quantified by chromatographic analyses (GCxGC–FID, GCxGC/TOF–MS, GC–MS and HPLC). Together with FTIR, Py-GC–MS, TAN, elemental analysis and various advanced NMR techniques (13C NMR, 31P NMR, 19F NMR, HSQC NMR, HMBC NMR), structural features of the PL oligomers were elucidated, revealing a guaiacyl backbone linked by alkyl, ether, ester and carbonyl groups, with none of the typical native lignin linkages (i.e. β–O–4, β–β, β–5) present. Furthermore, 72.3 % of the oxygen content in PL could be assigned to specific motifs by the quantitative analyses performed, and oligomeric models were proposed based on the obtained information. We expect that this characterization work can support future research on the development of valorization pathways for PL, allowing the feasible conversion of this promising feedstock into valuable biobased products with a wide range of possible applications, e.g. fuels, materials and specialty chemicals.

1. Introduction

Lignocellulosic biomass is a widely available source of renewable carbon and a promising feedstock for the replacement of petro-based fuels, (intermediate) chemicals and materials through the conversion in so-called biorefineries [1–4]. Lignin, one of its three main building blocks, corresponds to 10–40 wt% of biomass dry solids [3] and consists of a rich aromatic structure that, upon an efficient depolymerization, can yield value-added (functionalized) aromatics with several possible applications (e.g. fuels, fine chemicals, materials). There are many processes to isolate lignin from biomass (e.g. organosolv, pyrolysis, acid hydrolysis), in which the conditions and chemicals used vary substantially. Similarly, the properties of the lignins obtained from these processed vary significantly [5]. Therefore, a detailed characterization of each specific lignin type is crucial to develop efficient upgrading strategies.

Fast pyrolysis stands out as an attractive primary thermochemical process to liquefy biomass due to the flexibility of feedstock and process conditions, relatively low cost and high energy conversion efficiency [6,7]. For instance, yields of up to 75 wt% [6] of pyrolysis oil (also known as pyrolysis liquid and bio-oil) can be achieved. Pyrolysis oils can be easily fractionated by the addition of water, which leads to the separation of a sugar (aqueous) fraction and a water-insoluble fraction comprised mostly of lignin-derived aromatic fragments [8], typically referred to as pyrolytic lignin (PL). This allows for the two fractions, intermediates in a pyrolysis oil biorefinery scheme, to be processed independently by strategies tailored to their nature and inherent properties into a wide range of valuable products, e.g. alkylphenolics [9,10], biofuels [4,11,12], hydroxymethylfurfural (HMF) [13], as well as feedstocks suitable for co-feeding traditional refineries [14,15].

Despite promising initial results from studies on PL upgrading to monomers [16–23] and materials [24–26], its structural complexity
makes further processing overall challenging. For instance, the thermal decomposition of native lignin during pyrolysis involves various pathways, including competitive and/or consecutive reactions (e.g. dehydration, condensation, demethylation, ether cleavage) throughout a wide temperature range [27,28]. This leads to both chemical and size heterogeneity in the obtained PL. Furthermore, the biomass source also significantly influences PL composition. For example, while softwoods form mainly guaiacols-based PLs, hardwoods are decomposed in both guaiacols and syringols units [29]. The existing literature on PL characterization, albeit scarce, has given some insight on its structure. For instance, it has been reported that PL consists mainly of trimers and tetramers of HGS (hydroxyphenyl, guaiacyl and syringyl) units, as a result from the high pyrolysis temperatures leading to thermally driven depolymerization reactions [30–33]. New types of inter-unit linkages different from the typical alkyl-aryl-ether in native lignins are formed, particularly carbon-carbon linkages and saturated aliphatic side chains [20,30,31,34–36]. Some PL studies reported alkyl-aryl-ether linkages in PL as a result of the thermal ejection of (less modified) lignin oligomers [37]. Thermal splitting during pyrolysis is claimed to generate unconjugated carbonyl groups and C=C double bonds [30,33,34], while the amount of methoxy groups and aliphatic hydroxy groups decreases substantially in comparison with native lignin [33]. Fig. 1 shows PL molecular structures (obtained from different biomass sources) as proposed in the literature.

As illustrated in Fig. 1, the reported PL structures vary substantially in terms of linkages. Inconsistencies are likely derived from limitations of the analyses performed, as the characterization of lignin oligomers is not trivial and typically requires the use of advanced analytical procedures. Fortunately, recent developments on NMR techniques have provided unprecedented qualitative and quantitative information regarding the structure of technical lignins [39–43]. Therefore, new motifs and insights can be obtained, being these necessary to further evolve our understanding on the major occurring chemical motifs and structures in pyrolytic lignins.

In detail, the (only) structure proposed for a pine-derived PL (A) mainly contains native linkages (i.e. β-O-4, β-β, β-5) which are not expected to resist to pyrolysis conditions [30,38]. Thus, this structure definitely requires an update. This served as motivation for us to perform an in-depth characterization of a PL obtained from a commercially available pine-derived pyrolysis oil (Empyro B.V., the Netherlands). Macromolecular properties were assessed by gel permeation chromatography (GPC), thermogravimetric analysis (TGA) and elemental analysis, and the monomeric fraction was quantified and identified by chromatographic analyses (GCxGC–FID, GCxGC/TOF–MS, GC–MS and HPLC). Together with FTIR, pyrolysis gas chromatography with mass detection (Py-GC–MS) and total acid number (TAN) analyses, advanced NMR techniques (13C NMR, 31P NMR, 19F NMR, HSQC NMR, and HMBC NMR) were performed to elucidate the structural features present on the PL oligomers and allow for a precise oxygen balance. Finally, the gathered information was used to provide a detailed overview of the structural characteristics of PL, allowing us to suggest further refined chemical structures of the (major) oligomeric fragments.

2. Materials and methods

2.1. Chemicals

The pine-derived pyrolysis oil was supplied by Biomass Technology Group (BTG, Enschede, the Netherlands). The pyrolysis oil is commercially available and was produced at 500°C in a rotating cone reactor [44] (capacity of 5 ton/h) by Empyro B.V. (Hengelo, the Netherlands) and filtered (0.03 wt% of solids removed). Tetrahydrofuran (THF), toluene, deuterated dimethyl sulfoxide (DMSO–d6), cyclohexanol, pyridine, chromium(III) acetylacetonate, 4-(trifluoromethyl)phenylhydrazine, 2-chloro-4,4,5,5-tetramethyl-1,3,2-dioxaphospholane, deuterated chloroform (chloroform-d) and di-n-butyl ether (DBE) were purchased from Sigma-Aldrich. 1-methyl-4-(trifluoromethyl)benzene was purchased from TCI Europe N.V. All chemicals in this study were used as received.

Fig. 1. Previously proposed structures of PL oligomers. (A) G-based tetramers from pine sawdust PL [37] (B) Pentamer from red oak PL [20] (C) Oligomers from beech wood PL [30] (D) Oligomer from switchgrass PL [38].
2.2. PL extraction

The PL fraction of the pine-derived pyrolysis oil was obtained by fractionation with water. Pyrolysis oil (100 g) was added dropwise to Milli-Q water (150 g) at room temperature under vigorous stirring. The water-soluble fraction was removed and another portion of fresh water (100 g) was added to the insoluble fraction, followed by vigorous stirring for 30 min and subsequent removal of the water-soluble fraction. Finally, the insoluble fraction was centrifuged for 15 min to yield 32.6 wt% of PL for analysis. See extraction scheme in Fig. 2.

2.3. Analysis of PL

The detailed characterization of the PL was obtained by performing a series of techniques that provided information on the molecular weight (MW) distribution (GPC), thermal stability (TGA), identification and quantification of monomers (GCxGC/TOF–MS, GCxGC–FID, GC–MS, HPLC), water content (Karl Fischer), TAN analysis, structural and quantification of monomers (GCxGC/TOF–MS, GCxGC–FID, weight (MW) distribution (GPC), thermal stability (TGA), identification of a series of techniques that provided information on the molecular size of the PL fraction for characterization.

Fig. 2. Water fractionation of pyrolysis oil to yield the PL fraction for characterization.

For analysis by gas chromatography (GC), the PL sample was diluted around 20 times with 500 ppm solution of DBE (internal standard) in THF. GCxGC/TOF–MS analysis was performed on an Agilent 7890B system equipped with a JEOL AccuTOF GCv 4 G detector and two capillary columns, i.e. a RTX-1701 capillary column (30 m × 0.25 mm i.d. and 0.25 μm film thickness) connected by a solid state modulator (Da Vinci DVLS GC ³) to a Rxi-5Sil MS column (120 cm × 0.15 mm i.d. and 0.15 μm film thickness). Quantification of GCxGC-FID main groups of compounds (e.g. aromatics, alkanes, alkylphenolics) was performed by using an average relative response factor (RRF) per component group in relation to an internal standard (di-n-butyl ether, DBE). GC–MS analysis was performed on a Hewlett-Packard 5890 gas chromatograph equipped with a RTX-1701 capillary column (30 m × 0.25 mm i.d. and 0.25 μm film thickness) and a Quadrupole Hewlett-Packard 6890 MSD selective detector attached. Helium was used as carrier gas (flow rate of 2 mL/min). The injector temperature was set to 280°C. The oven temperature was kept at 40°C for 5 min, then increased to 250°C at a rate of 3°C/min and held at 250°C for 5 min.

The HPLC analytical device consisted of an Agilent 1200 pump, a Bio-Rad organic acids column Aminex HPX-87H, a Waters 410 differential refractive index detector and a UV detector. The mobile phase was 5 mM aqueous sulfuric acid at a flow rate of 0.55 mL/min. The HPLC column was operated at 60°C. An extra water extraction step (proportion of 1–10 of PL and water, mixed for 1 h in an ultrasonic bath) was performed prior to analysis to solubilize the residual polar compounds. Calibration curves of the targeted molecules (i.e. levoglucosan, acetic acid, glycoaldehyde and formic acid) were built to provide an accurate quantification and were based on a minimum of 4 data points with excellent linear fitting (i.e. R² > 0.99).

The water content in the PL was determined by Karl Fischer titration using a Metrohm 702 SM Titriton titration device. About 0.01 g of the PL sample was injected in an isolated glass chamber containing Hydranal (Karl Fischer solvent, Riedel de Haen). The titrations were carried out using the Karl Fischer titrant Composit 5 K (Riedel de Haen). The analysis was performed at least three times and the average value is reported.

The TAN titration method was performed with a Metrohm 848 Titriton plus apparatus equipped with a Metrohm 6.0262.100 electrode. Between 0.05 − 0.09 g of sample was dissolved in 30 mL of an acetone-water 1:1 solution, and titration with a 1.0 M KOH solution was performed until the solution reached the first endpoint, i.e. point in which strong acids are neutralized [45]. The TAN calculation is depicted in Eq. 1, where C0 is the KOH concentration in the solution, m is the weight of the sample used for titration, V1 is the volume of titrant required for a blank experiment (mL) and V2 is the volume of titrant required for the titration of the PL sample (mL). Three measurements were performed and the average value is reported. The correlation between the TAN and the mmol of carboxylic acid groups/g of PL is based on a stoichiometric
Table 1
Elemental composition of the PL used in this study.

| Element (wt%, dry basis) | Value  |
|-------------------------|--------|
| Carbon                  | 65.3   |
| Hydrogen                | 6.6    |
| Nitrogen                | < 0.01 |
| Oxygen                  | 28.1   |
| Sulfur                  | < 0.01 |

ratio of neutralization of 1:1 and is given in Eq. 2 (56.11 mg/mmol is the molecular weight of KOH).

\[
TAN = \frac{V_2 - V_1}{V_1} \times \frac{C_{\text{OH}} \times 56.11}{m_i} = \frac{\text{mg KOH}}{\text{g oil}}
\]

\[
\text{mmol carboxylic acid groups} = \frac{TAN \times (V_2 - V_1) \times C_{\text{OH}}}{56.11} \times \frac{m_i}{\text{g of PL}}
\]

An attenuated Total Reflection Infrared (ATR-IR) spectrometer was used for the FTIR measurement. Around 1–2 drops of sample were placed on the sample unit (Graseby Specac Golden Gate with diamond top) and IR-spectra were obtained using a Shimadzu IR-Tracer-100 FT-IR spectrometer with resolution of 4 cm⁻¹ and 64 scans.

Py-GC–MS analysis was performed in a Tandem μ-Reactor (TMR) from Frontier Lab (Rx-3050TR) equipped with a single shot sampler (PY1 – 1040). A carrier gas inlet was connected on the top of the TMR, providing the gas flow to the GC–MS. The entire system was attached by a docking station on top of the GC–MS and connected by an injection needle through a rubber septum. Before the experiment, the system was pressurized to 150 kPa with an inert carrier gas (Helium) to check for leakage. After the leak check, the pressure was set back to 50 kPa and the system was heated to 500 °C. A stainless steel cup filled with PL was attached to the sample injector and the cup was dropped into the TMR. Different NMR experiments were performed on a Bruker NMR spectrometer (600 MHz) at 293 K using a standard 90° pulse, and the spectra were processed and analyzed using MestReNova software, refer to the Supplementary Information for integration details. Sample preparation involved the dissolution of the PL in DMSO-d₆ (25 wt%). Heteronuclear single quantum correlation (HSQC) NMR and heteronuclear multiple bond correlation (HMBC) NMR spectra were acquired with the following parameters: 11 ppm sweep width in F2 (¹H), 220 ppm sweep width in F1 (¹³C) and 8 scans. ¹H NMR spectrum was acquired using a sweep width of 11 ppm and 8 scans. ¹³C NMR spectrum was acquired using a relaxation delay of 5 s, sweep width of 220 ppm and 2048 scans. Hydroxyl content analysis was performed through ³¹P NMR following a procedure elsewhere [41], using cyclohexanol as the internal standard. The ³¹P NMR spectrum was acquired using a relaxation delay of 10 s and 512 scans. Carboxyl content analysis was performed through ¹⁹F NMR following a procedure described elsewhere [39], using 1-methyl-4-(trifluoromethyl) benzene as the internal standard. The ¹⁹F NMR spectrum was acquired using a relaxation delay of 3 s and 256 scans.

Elemental analysis (C, H, N) was performed using a EuroVector EA3400 Series CHN-O analyzer with acetonilide as the reference. The oxygen content was determined by difference. The analysis was carried out at least in duplicate and the average value is reported.

3. Results and discussion

To obtain the PL (water-insoluble) fraction from its parent pyrolysis oil, a water extraction procedure was performed (vide supra, Section 2.2). The PL was obtained as a viscous dark brown liquid (32.6 wt%) yield. A residual water content of 8.2 wt% of PL was obtained through Karl Fischer analysis. This relatively high water content likely causes the product to be in the liquid state [17,46]. The dry yield of PL was calculated to be 29.9 wt% of PL based on pyrolysis oil, which is in the 25 – 30 wt% range reported for lignin content in pine wood [47–49]. A TGA analysis under air confirmed the absence of ash in the PL fraction (Fig. S1).

Different types of analysis were performed on the obtained PL to get further insight into its composition and structural characteristics, which are ordered and divided over the following sections. The macromolecular properties such as MW distribution (GPC), thermal stability (TGA) and elemental composition are discussed in Section 3.1. The monomers present in the PL are identified and quantified in Section 3.2, and the global chemical features observed by NMR techniques, FTIR, TAN and Py-GC–MS analyses are discussed in detail in Section 3.3. The study is concluded with an overview of the PL structure (Section 3.4), which includes a structural proposal for the PL oligomeric fraction.

3.1. Macromolecular properties

The elemental composition of the PL is shown in Table 1. Due to the use of pine wood as the biomass source, amounts of nitrogen and sulfur are negligible, being beneficial in further catalytic upgrading processes as these elements may have a negative impact on catalytic performance [50,51]. Furthermore, when having high quality fuels as the final application, environmental regulations require low contents of nitrogen and sulfur to avoid harmful emissions during combustion [52].

Fig. 3 shows a comparative Van Krevelen plot of the PL used in this study, highlighted as Pine (commercial), and various PLs reported in literature. The plot clearly shows a significant variation on H/C (1.0–1.3) and O/C (0.25 – 0.4) molar ratios, which can be related to the pyrolysis conditions applied, PL extraction procedure and the biomass source used.

TGA results for the PL are shown in Fig. 4, in which the non-volatile residue was of around 20 wt%, being in line with the literature [16,18]. The sharp peak at around 100 °C is expected due to the presence of residual water, and indeed a weight loss of 8.8 wt% is observed in the heating range of 95–130°C, similar to the water content obtained by Karl Fischer analysis (vide supra). Overall, the PL thermal decomposition profile can be divided into three main stages [33,54,55]: i) volatilization of residual water and low MW compounds (< 160 °C); ii) thermal cracking of labile C–O bonds (particularly in methoxy side groups) and dissociation of C≡O functional groups, with extensive formation of gaseous products (i.e. CH₄, CO, CO₂) and volatile monomers (160–280 °C); iii) thermal cracking of more stable C–C bonds and other inter-unit linkages (280–500 °C). At temperatures above 500 °C, the flattened signal likely indicates recondensation of aromatics during analysis [56].

GPC analyses provided information regarding the MW distribution of the PL. Fig. 5 shows the results, which are in line with values related to pine-derived PLs [16–18,53] and ratify that lignin in the pine wood is thermally depolymerized during pyrolysis, as the weight average molecular weight (Mₙ) of PL is much lower than those of other types of technical lignins (i.e. Kraft and Alcell [18,57]). By integrating specific areas of the GPC data [34], the Mₙ range comprising monomers, dimers and trimers was shown to correspond to 42 % of the distribution, while tetramers, pentamers and hexamers corresponded to 33 % and larger fragments (> hexamers) corresponded to 25 %. While the proportion of smaller compounds was within the same range as for other PLs from different biomass sources, the pine-derived PL used in this study was overall richer in intermediate fragments (rather than large ones, i.e. > hexamers) [34]. The presence of small molecules in the PL allows for the identification and quantification of monomers via chromatography. In the next section, techniques and results will be discussed in detail.

3.2. Monomeric fraction

To determine the small molecules present in the PL, both HPLC and different types of GC analyses were performed. Through HPLC, the four main residual water soluble compounds were identified and quantified,
i.e. levoglucosan, glycolaldehyde, acetic acid and formic acid (Table 2). GCxGC–FID was performed to estimate the amounts of monomers per group of chemical functionalities, following a procedure previously reported by our group [17,58,59]. For instance, this technique provides a straightforward separation of the organic compound classes typically found in biomass-derived liquids (see Fig. S2 for the PL chromatogram).

Table 2 shows the integration results. Phenols and guaiacols (i.e. methoxyphenols) form the bulk of the PL monomeric fraction, followed by residual polar compounds. Included in the latter group that corresponds to 3.9 wt% of PL are acetic acid and glycolaldehyde, which were also quantified individually by HPLC (formic acid is not detectable by GCxGC-FID).

As the amount of GC-detectables in the PL were found to be significant, we were interested in further identifying the main chemical constituents, particularly the (methoxy)phenolics due to their high value and wide range of possible applications [60]. To that end, PL was analyzed by GCxGC/TOF–MS and GC–MS. The GCxGC/TOF–MS PL chromatogram with its main peaks identified is shown in Fig. 6 (see Fig. S3 and Table S3 for a complete overview). Together with minor amounts of ketones and furans, a range of substituted (methoxy)phenols was observed (among others 4-methyl, 4-ethyl and 4-propyl guaiacols, alkylphenols, catechols and vanillin), being in line with previous chromatographic studies of PL and whole pyrolysis oils of various origins [35,46,61,62]. While G-type biomasses (such as the pine wood used in this study) release guaiacols due to thermal cracking...
pathways occurring during pyrolysis, further demethoxylation reac-
tions of the guaiacols lead to their respective phenols [28]. In some
monomers, an unsaturated propyl side chain was present (i.e. eugenol),
being derived from the dehydration of the OH group at the positions α
and γ of the β–O–4 linkage [63]. Other monomers have carbonyl and
ester groups in their side chains. A few naphthalenes were also iden-
tified by GCxGC, contrasting a previous characterization study of
hardwood PL, where the lack of naphthalenes was specifically high-
lighted [19]. While condensation pathways are indeed not prominent
(only small amounts were observed), they seem to have happen to a
minor extent in the pine-derived PL. This was further confirmed by Py-
GC–MS (vide infra). GC–MS analysis aided further elucidation by con-
firming the main phenolic monomers and identifying low MW com-
pounds such as propanal, which are likely derived from lignin propyl
side chains (Fig. S4).

3.3. Global chemical features

When summing up the residual water (8.2 wt% of PL) and monomer
fraction (12.4 wt% of GCxGC–FID detectables), around 80 wt% of the
PL remains unidentified. Accordingly, this fraction consists of oligomers
that cannot be simply identified by gas-chromatography due to their
higher MW. Thus, other analyses were employed to get better insights
on the PL structure as a whole with the aim aims to shed light on its
global chemical features by means of FTIR, Py-GC–MS, TAN and ad-
vanced NMR techniques.

FTIR spectroscopy ratified the overall phenolic structure of PL (Fig.
S5 and Table S4). Furthermore, the results also show the presence of
C–O, C–O and alkyl chains. These are in line with the monomers
observed in the previous section, suggesting that the PL oligomers are
comprised of similar structural motifs in which the phenolic backbone
is linked by aliphatic C1–C3 chains that might as well contain oxygen in
the form of C–O and C==O.

Py-GC–MS analysis was also performed to gain more insights re-
garding the PL end groups and interunit linkages, having as reference
the pyrolysis products identified in the obtained chromatogram (Fig. 7).
For instance, the spectrum shows a mixture consisting mostly of aro-
matics and phenolics, with low amounts of methoxyl side groups due to
thermally induced demethoxylation. Some of the main compounds
observed, namely 4-methyl guaiacol, eugenol, isoeugenol and vanillin,
are in line with the results from a previous Py-GC–MS study of pine-
derived PL [46]. This study was unfortunately limited in terms of
identified compounds (< 10). The presence of vanillin and benzoic acid
is likely a result of carboxyl and aldehyde end groups in the PL oligo-
mers. Furthermore, trimers observed at longer retention times suggest
biphenyl interunit linkages in the PL, as well as aromatization pathways
that lead to the formation of polycyclic aromatics (e.g. naphthalenes).
While the presence of biphenyl linkages has been suggested in literature
[19], polycyclic aromatics were not previously reported for PL, prob-
able due to their low occurrence. A previous pyrolysis study using a
synthetic G-based lignin model showed that the formation of polycyclic
aromatics happens extensively at higher pyrolysis temperatures

![Fig. 6. GCxGC/TOF-MS chromatogram with the main PL peaks assigned. (A) Aliphatics and furan monomers and (B) Phenolic monomers. DBE is the IS and BHT is the THF stabilizer.](image-url)
Although anthracene was observed already at 500 °C [64], condensation pathways likely also occur in a minor extent at lower pyrolysis temperatures when a more complex and chemically heterogeneous feedstock such as pine wood is used. Overall, no substantial qualitative differences were observed between the monomers identified in the previous section and the Py-GC–MS results. This indicates that the monomer structures identified are very representative of the subunits present within the PL oligomeric structures.

The $^{13}$C NMR spectrum obtained for PL ratifies its highly aromatic profile, as aromatic linkages correspond to > 50 % in terms of relative $^{13}$C NMR area (Fig. 8). The difference between aromatic C–O linkages (Arom C–O) and methoxy side groups (Arom-OCH$_3$) indicates the existence of other types of C–O bonds between the aromatic rings and oxygen, i.e. C–OH in phenolic units and C–O–C in diaryl structures. The integration results show that aliphatic C–H and C–O bonds represent a significant part of the PL structure (≈ 40 % in terms of relative area). Accordingly, the aliphatic fraction is present in the form of side chains of the phenolic backbone, as well as residual sugars and acids as shown by the identified monomers (vide supra). It is likely that the amount of C=O groups was underestimated, as the quaternary carbon signal is suppressed in this analysis [65,66]. For this reason, other techniques (TAN and $^{19}$F NMR, vide infra) were used to better visualize the acids, ketones and aldehydes present in the PL.

The distribution of the OH groups within the PL structure was assessed in detail via $^{31}$P NMR, see Fig. 9 for the obtained spectrum with the integration results of the assigned regions. For instance, most of the OH content in the PL arise from guaiacyl (G) and phenolic units, yet C-5 substituted structures were also identified. The aliphatic OH content was significant, but instead of being related to β-O-4 bonds (as in the case of lignin types more similar to native lignin, e.g. organosolv), in PL such groups are mostly a result of β-O-4 cleavage during pyrolysis, which leads to propanol side chains. Furthermore, the presence of residual sugars and hydroxylated furans (i.e. HMF) contribute to the aliphatic OH fraction. Such variety on the types of aliphatic OH can be clearly observed when comparing the $^{31}$P NMR spectrum of PL with other lignins [41], which typically show one strong signal on the aliphatic OH region.

To estimate the content of acid groups in the PL, TAN titrations were also performed. The TAN method applied identified the first endpoint of the titration curve, which corresponds to the neutralization of higher dissociation acids (i.e. carboxylic acids) rather than the weakly acidic phenolics [45] (Fig. S6). The averaged TAN of 42.1 mg KOH/g PL (which corresponds to 0.75 mmol KOH/g PL) is in line with previous results reported for PL [67], corresponding to 0.75 mmol of acid groups/g of PL. As 0.4 mmol acid groups/g of PL are related to formic and acetic acids (based on HPLC results, vide supra), an estimated
concentration of 0.35 mmol acid groups/g of PL can be assigned to end
groups present within the oligomeric structure of the PL.

$^{19}$F NMR analysis was performed to identify and quantify the car-
bonyl groups in PL. The distribution of the carbonyl groups in the ob-
tained spectrum is shown in Fig. 10, corresponding to 2.5 mmol/g of PL
in total. Whereas aldehyde and ketones cannot be distinguished [39],
signals related to both aliphatic and conjugated structures are present
in the PL spectrum. In line with the side groups observed in the iden-
tified monomers (vide supra), these results clearly show that a sig-
nificant amount of carbonyl is present (in both aliphatic and conjugated
forms) in the PL oligomers. This chemical functionality was largely
underestimated in all the previous characterization studies of PL
[20,30,37,38].

Finally, 2D NMR techniques were employed to assist the complete
fingerprinting of the PL structure. The HSQC NMR spectrum shows
direct C–H linkages existent in the PL, and distinct regions were
identified based on the literature, see Fig. 11. For instance, strong sig-
nals in the aliphatic C–H region ($\delta_C/\delta_H$ 0–45/0–3 ppm) were ob-
served, which include both aliphatic C–H and aliphatic C–H located
next to aromatic rings and carbonyl groups. In detail, a specific signal
related to CH$_2$ groups in diaryl methane was also identified for the first
time for PL [68]. The signals present in the aliphatic C–O region ($\delta_C/\delta_H$
45–105/3–5.5 ppm) indicated the presence of oxygenated aliphatic
chains, ester groups and residual sugars [41,69–72]. Regarding the
latter, typical LCC (lignin-carbohydrate complex) phenylglycoside and
ester bonds were identified, indicating that part of the levoglucosan is
linked to the PL structure. This is in line with the LCC literature [73–75]
and is also suggested by the HSQC NMR spectrum of a pine-derived PL
obtained through an extensive water fractionation procedure reported
elsewhere [17] (thus not expected to contain free sugars), which shows

![Fig. 9. $^{31}$P NMR spectrum of PL and integration results of assigned regions (solvent: chloroform-d).](image)

![Fig. 10. $^{19}$F NMR spectrum of PL and integration results of assigned regions (solvent: DMSO-$d_6$).](image)
identical levoglucosan and LCC signals (Fig. S7). Importantly, none of the typical interunit bonds (e.g. β–O–4, β–β, β–5) found in native lignin was observed, as these are highly prone to cleavage during pyrolysis [19,30,38]. Since the PL used in this study is derived from softwood (pine), the aromatic region ($\delta_{C}/\delta_{H}$ 105–135/6–8 ppm) consists mostly of G-based units [76]. Furthermore, signals related to furan structures [77], aromatic side chains containing C–C double bonds and a range of other oxygenated aromatics (e.g. benzaldehyde, ferulate, cinnamaldehyde) are observed. Additional HMBC NMR analysis (Fig. S8) elucidated long range C–H correlations, and confirmed the presence of esters, ketones and acid groups, as well as linkages between (quaternary) aromatic carbons.

Fig. 11. HSQC NMR spectrum of PL and main representative structures (solvent: DMSO-$d_6$). (Ar–OCH$_3$) Methoxy groups bonded to aromatics (Lg) Levoglucosan; (LCC) Lignin-carbohydrate bonds; (Fe) Ferulate; (Fu) Furans.
3.4. Proposal for the PL structure

The extensive set of qualitative and quantitative analyses performed with the PL provided valuable information regarding its structural features. This included the accurate quantification of oxygen-containing groups (OH by $^{31}$P NMR, COOH by TAN and C=O by $^{19}$F NMR, vide supra), which allowed for an oxygen balance having as reference the oxygen content of PL determined by elemental analysis (see Supplementary Information for the calculations). For instance, 49.5 % of the oxygen content was assigned to hydroxyl groups, 8.5 % was assigned to carboxylic acid groups and 14.3 % was assigned to carbonyl groups, adding up to a total of 72.3 %. Reasons for the unidentified oxygen fraction are related to imprecisions inherent to the analyses, the (known) occurrence of esters and furans, and presence of ether linkages not identified by NMR (e.g. ether side chains in non-phenolic structures).

Fig. 12 shows an overview of the pine-derived PL characterized in this study, including main identified monomers and possible oligomeric structures (the estimated amounts of 15 wt% and 85 wt% are based on dry PL, see Supplementary Information for the calculation). Quantitative analyses provided valuable information on the proportions of chemical functionalities within the PL structure, i.e. aromatic OH/alkyllic OH, aromatic OH/COOH and aromatic OH/C=O ratios, which were used as a reference in the proposed PL structures. $^{13}$C NMR integration results and additional $^1$H NMR analyses (Fig. 59) provided further insights on the carbon and hydrogen distribution (see Supplementary Information for the calculations). For instance, the results suggest that the aromatic backbone contains an average of two free (C–H) positions and one methoxy side group per aromatic ring. In addition to that, phenolic OH side groups, (oxygenated) alkyl chains, ester, diaryl ether (particularly 4–O–5) and C–C (particularly 5–5 and stilbene) interunit linkages are present. Most of the aliphatic hydrogen occurs in the form of CH$_2$, since CH$_3$ groups are mainly attached to an oxygen as in methoxy side groups. The structures proposed in Fig. 12 have an overall elemental composition of 66.1 % C / 5.9 % H / 28 % O, which is indeed similar to the elemental composition results of PL (vide supra).

4. Conclusions

In this work, an in-depth characterization study of a pine-derived PL obtained from a commercially available pyrolysis oil provided qualitative and quantitative information regarding both its monomeric (15 wt%) and oligomeric (85 wt%) fractions. A range of (methoxy) phenolic monomers was identified, followed by minor amounts of naphthalenes and residual polar compounds such as furans and small acids. The PL oligomeric structure was shown to be comprised of a guaiacyl backbone linked by alkyl, ether, ester and carbonyl groups, with none of the typical linkages found in native lignin (i.e. β–O–4, β–β, β–5). Aldehydes and acids are present as end groups, and the occurrence of other structures formed during pyrolysis (i.e. benzofurans, naphthalenes and catechols) is also suggested. Furthermore, LCC bonds were identified, indicating the presence of sugar molecules bonded within the PL structure. Quantitative analyses allowed for an accurate oxygen balance in which 72.3 % of the oxygen content in PL could be assigned to specific motifs. The results allowed for further evolving the understanding of the complex and chemically heterogeneous structure of PL. This can be used to further develop tailored upgrading strategies that support future research on PL processing, ultimately aiming for the production of valuable biobased chemicals and materials.

CRediT authorship contribution statement

M.B. Figueirêdo: Methodology, Validation, Formal analysis, Investigation, Writing - original draft, Visualization, Funding acquisition. R.H. Venderbosch: Conceptualization, Methodology, Resources. H.J. Heeres: Conceptualization, Methodology, Writing - review & editing, Supervision. P.J. Deuss: Conceptualization, Methodology, Formal analysis, Writing - review & editing, Visualization, Supervision.
Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

Financial support from the Science without Borders program (CNpq, Brazil) on the PhD project of MBF is gratefully acknowledged. The TKI Biobased Economy and the Netherlands Enterprise Agency (RVO) are gratefully acknowledged for their financial support of the Lignin2Fuel project under agreement no. TEBE116143. We also thank Léon Rohrbach and Jan Henk Marsman for the analytical support. Hans van der Velde is acknowledged for performing the elemental analyses, and Dian S. Santosa is acknowledged for performing the Py-GC–MS experiment. Arjan Kloekhorst is gratefully acknowledged for providing the extensively washed PL, which was supplied by BTG. We also thank BTG for providing the pyrolysis oil (commercialized by Empyro B.V.) used in this work.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:10.1016/j.jaap.2020.104837.

References

[1] F. Cherubini, The biorefinery concept: using biomass instead of oil for producing energy and chemicals, Energy Convers. Manage. 51 (2010) 1412–1421, https://doi.org/10.1016/j.enconman.2010.01.015.
[2] M. FitzPatrick, P. Champagne, M.F. Cunningham, R.A. Whitney, A biorefinery processing perspective: treatment of lignocellulosic materials for the production of value-added products, Biosour. Technol. 101 (2010) 8915–8922, https://doi.org/10.1016/jbiosour.2010.06.125.
[3] V.K. Ponnamury, D.D. Nguyen, J. Dharmanarayana, S. Shobana, J.R. Bana, R.G. Sarate, S.W. Chang, G. Kumar, A review on lignin structure, pretreatments, fermentation reactions and biorefinery potential, Biosour. Technol. 271 (2019) 462–472, https://doi.org/10.1016/j.biortech.2018.09.070.
[4] G.W. Huber, A. Corma, Synergies between bio- and oil refineries for the production of fuels from biomass, Angew. Chemie Int. Ed. 46 (2007) 7184–7201, https://doi.org/10.1002/anie.200604504.
[5] H. Guo, B. Zhang, Z. Qi, C. Li, J. Ji, T. Dai, A. Wang, T. Zhang, Valorization of lignin to simple phenolic compounds over tungsten carbide: impact of lignin structure, ChemSusChem 10 (2017) 523–532, https://doi.org/10.1002/cssc.201601526.
[6] A.V. Bridgwater, Review of fast pyrolysis of biomass and product upgrading, Biomass Bioenergy 38 (2012) 68–94, https://doi.org/10.1016/j.biombioe.2011.01.048.
[7] T. Dickerson, J. Soria, Catalytic fast pyrolysis: a review, Energies 6 (2013) 514–538, https://doi.org/10.3390/en60100514.
[8] M. Fortin, M. Mohadjer Beromi, A. Lai, P.C. Tarves, C.A. Mullen, A.A. Boateng, Characterization of water-insoluble fraction from fast pyrolysis liquids (pyrolytic lignin). Part IV: catalytic hydrogenation of pyrolytic lignin and milled wood lignin by using TG–FTIR analysis, J. Anal. Appl. Pyrolysis 108 (2014) 78–85, https://doi.org/10.1016/j.jaap.2014.05.017.
[9] M.B. Figueirêdo, et al.
