Elucidation of lignin derived oligomers for predictive thermodynamic modeling in fast pyrolysis bio-oils

Zur Erlangung des akademischen Grades einer

DOKTORIN DER INGENIEURWISSENSCHAFTEN

von der KIT-Fakultät für Chemieingenieurwesen und Verfahrenstechnik des Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

M. Sc. Ana Cristina Corrêa de Araujo
aus Niterói (RJ), Brasilien

Tag der mündlichen Prüfung: 18.12.2025

Erstgutachter: Prof. Dr. Nicolaus Dahmen

Zweitgutachter: Prof. Dr. Alberto Wisniewski Jr.

Acknowledgements

Here, I would like to express my gratitude to everyone who contributed to this PhD work, for their support, guidance, and friendship were essential and deeply appreciated.

First and foremost, I would like to thank Prof. Dahmen and Dr. Axel Funke for giving me the opportunity to pursue my PhD at the Institute of Catalysis Research and Technology (IKFT), Karlsruhe Institute of Technology (KIT). I am sincerely grateful for their supervision, continuous support, and for the trust they placed in me throughout this journey.

I am especially thankful to Frederico for his friendship, mentorship, and for all his help with the Aspen Plus simulations and VLE modeling, our many discussions were truly valuable. My gratitude also goes to Myriam for her friendship, collaboration, and all the support on the lignin-derived oligomers topic; her insights made a real difference in this project.

To my fellow officemates from the "hot office" – Naiara, Eugen, and Carola – thank you for the friendship, good humor, and countless fun moments we shared. You made workdays lighter and this experience much more enjoyable. I also want to thank all my other IKFT colleagues (Gabriela, Henri, Bing, George, Fabrizio, Ecrin, Mahsa) and friends for their kindness, teamwork, and for making the institute a nice place to work.

I would also like to thank Daniel, Norbert, and the entire Python team for conducting the pilot scale lignin pyrolysis experiments and for their efforts in adapting the unit for lignin operation. I also thank Manuel Brislinger, my first master's student, for his valuable help with the advanced distillation curve (ADC) experiments.

A very special thanks goes to Dr. Klaus Raffelt and Dr. Caroline Schmitt for promoting the Brazil–Germany collaboration and for giving me the opportunity to take part in scientific exchanges and research visits. This experience allowed me to meet inspiring people, broaden my scientific perspective and build valuable professional connections.

As a part of this collaboration, I had the opportunity to work with Prof. Alberto Wisniewski Jr. from UFS on the UHRMS analysis of the lignin samples, and I am deeply grateful to him for his valuable contribution. I would also like to thank Wenes for performing the analyses and helping with the interpretation of the results. The outcomes of this collaboration were essential for the study of lignin oligomers and contributed significantly to this thesis.

In addition, this PhD was carried out within the framework of the Flexi-Green Fuels project, funded by the European Union's Horizon 2020 research and innovation program, under Grant Agreement No. 101007130. I am thankful to all project partners for the collaboration, stimulating discussions, and great times during our meetings. Special thanks go to the team at

AUTH – Kostas and Antigoni – for the productive collaboration and all the great work that came out of our joint efforts.

Finally, I would like to express my deepest gratitude to my parents, whose love, encouragement, and belief in me have been the foundation of all my achievements. I am also immensely thankful to my friends, both in Germany and in Brazil, for their constant support, laughter, and companionship throughout these years.

Preamble

This dissertation follows a paper-based format. Accordingly, part of the content presented here has been previously published or submitted for publication. Some sections therefore contain text identical or very similar to that of the original manuscripts, with minor adaptations to ensure consistency in formatting, citation style, figures, and layout throughout the thesis.

The introductory Chapter 1 provides an overview of the scientific background and motivation of the work, including a literature review on lignin valorization, fast pyrolysis, and phase-equilibrium modeling. This structure avoids unnecessary repetition of introductory or theoretical material across the individual chapters.

The main body of the thesis is organized around four research papers, each representing one of the core studies conducted during the PhD. Chapter 2 presents pilot-scale lignin fast-pyrolysis experiments in a twin-screw reactor; Chapter 3 focuses on the comprehensive analytical characterization of lignin-derived bio-oils; Chapter 4 investigates vapor–liquid equilibrium behavior and surrogate modeling of the pyrolytic lignin fraction; and Chapter 5 extends the analysis to liquid–liquid equilibrium studies. The final chapter integrates the findings and discusses their implications for the design and optimization of lignin fast pyrolysis and separation processes of fast pyrolysis bio-oils in general.

List of publications

Peer-reviewed original publications included in this thesis:

Assessing phase behavior of pyrolytic lignin in fast pyrolysis bio-oil via advanced distillation curve

Ana C. C. Araujo, Frederico G. Fonseca, Nicolaus Dahmen, Axel Funke

Biomass and Bioenergy, 2026, 209, 109009. DOI: 10.1016/j.biombioe.2026.109009

Pilot scale fast pyrolysis of lignin in a twin-screw reactor

Ana Cristina Corrêa de Araujo, Axel Funke, Antigoni Margellou, Konstantinos Triantafyllidis, Nicolaus Dahmen

Biomass and Bioenergy, 2025, 200, 108041. DOI: 10.1016/j.biombioe.2025.108041

Liquid–Liquid Equilibrium Prediction in Fast Pyrolysis Bio-Oil Systems: A Framework for Incorporating Bio-Oil Complexity

Myriam Rojas Salas, Frederico Gomes Fonseca, Ana Cristina Corrêa de Araujo, Manuel García-Perez, Axel Funke

Energy & Fuels, 2024, 38, 19, 18769–18780. DOI: 10.1021/acs.energyfuels.4c03387

In-preparation manuscript included in this thesis:

Comprehensive characterization of lignin bio-oils from alkali and Kraft lignin with integrated ^{1}H-NMR spectral modeling

Ana C. C. Araujo, Myriam Rojas, Wenes R. Silva, Antigoni G. Margellou, Konstantinos S. Triantafyllidis, Alberto Wisniewski Jr., Axel Funke

Peer-reviewed original publications not included in this thesis:

Solvolysis and Mild Hydrogenolysis of Lignin Pyrolysis Bio-Oils for Bunker Fuel Blends

Antigoni G. Margellou, Fanny Lasngschwager, Christina P. Pappa, Ana C. C. Araujo, Axel Funke, Konstantinos S. Triantafyllidis

Energies, 2025, 18, 14, 3683. DOI: 10.3390/en18143683

Abstract

The transition toward a sustainable energy system requires the efficient conversion of renewable carbon resources into fuels and chemicals. Lignin, a major component of lignocellulosic biomass, represents the largest renewable source of aromatic carbon but remains underutilized in current biorefinery concepts. Among the available conversion routes, fast pyrolysis offers a promising pathway for lignin valorization, yielding liquid intermediates known as fast pyrolysis bio-oils (FPBOs). However, the complex composition of FPBOs – particularly the presence of high molecular weight pyrolytic lignin (PL) oligomers – poses challenges for process design, product upgrading, and phase equilibrium modeling.

The main objective of this PhD work was to establish a link between the molecular level characterization of lignin-derived oligomers and their thermodynamic behavior, enabling the development of predictive models for process optimization. To address this goal, the study combined experimental, analytical, and modeling approaches in four complementary steps.

The first part of the work focused on the pilot scale fast pyrolysis of lignin in a twin-screw reactor, emphasizing stable operation under the severe melting behavior of lignin. Process performance was evaluated, providing insights into product yields and enabling continuous operation of lignin feeds. The second part dealt with the comprehensive chemical characterization of lignin derived FPBOs, employing FTIR, NMR, GC–MS, and high-resolution Orbitrap mass spectrometry (UHRMS) to elucidate structural diversity and functional group distribution. Based on these results, ^{1}H-NMR spectra were modeled using data from synthetic lignin oligomers (SLOs), linking experimental observations to representative molecular structures.

In the third stage, the vapor–liquid equilibrium (VLE) behavior of the PL fraction was investigated through advanced distillation curve (ADC) experiments and thermodynamic modeling. A range of structures, from dimers to tetramers, were evaluated as model compounds, providing guidance on the molecular representation of the heavy fraction. Finally, the liquid–liquid equilibrium (LLE) of FPBOs was examined to assess phase separation phenomena and the predictive accuracy of the proposed surrogate approach.

Overall, this thesis presents an integrated experimental and modeling framework for understanding and representing the complex phase behavior of lignin oligomers in FPBOs. The results contribute to a more fundamental description of pyrolysis products, supporting the design and optimization of fractional condensation and separation processes aimed at producing tailored, stable bio-oil fractions suitable for further upgrading into fuels and phenolic chemicals.

Zusammenfassung

Der Übergang zu einem nachhaltigen Energiesystem erfordert die effiziente Umwandlung erneuerbarer Kohlenstoffressourcen in Brennstoffe und chemische Wertstoffe. Lignin, ein Hauptbestandteil der lignozellulosischen Biomasse, stellt die größte erneuerbare Quelle aromatischen Kohlenstoffs dar, wird jedoch in gegenwärtigen Bioraffineriekonzepten bislang nur unzureichend genutzt. Unter den verfügbaren Umwandlungsrouten bietet die Schnellpyrolyse einen vielversprechenden Ansatz zur Valorisierung von Lignin, da hierbei flüssige Zwischenprodukte entstehen, die als Fast-Pyrolyse-Bioöle (FPBOs) bezeichnet werden. Die komplexe Zusammensetzung dieser FPBOs – insbesondere der hohe Anteil hochmolekularer pyrolytischer Lignin-Oligomere (PL) – stellt jedoch eine besondere Herausforderung für die Prozessauslegung, die Produktaufwertung und die Modellierung des Phasengleichgewichts dar.

Ziel der vorliegenden Arbeit war es, eine Verbindung zwischen der molekularen Charakterisierung ligninbasierter Oligomere und deren thermodynamischem Verhalten herzustellen, um die Entwicklung prädiktiver Modelle für die Prozessoptimierung zu ermöglichen. Zur Erreichung dieses Ziels wurden experimentelle, analytische und modellgestützte Ansätze in vier aufeinander aufbauenden Arbeitsschritten kombiniert.

Im ersten Teil der Arbeit wurde die Schnellpyrolyse von Lignin im Pilotmaßstab in einem Doppelschneckenreaktor untersucht. Dabei lag der Schwerpunkt auf der Gewährleistung eines stabilen Betriebs unter den ausgeprägten Schmelzeigenschaften von Lignin. Die Prozessleistung wurde umfassend bewertet, wodurch Einblicke in Produktausbeuten gewonnen und der kontinuierliche Betrieb von Lignin ermöglicht wurden. Der zweite Teil der Arbeit befasst sich mit der umfassenden chemischen Charakterisierung der aus Lignin gewonnenen Bioöle. Hierzu wurden FTIR, NMR, GC–MS und hochauflösende Orbitrap-Massenspektrometrie (UHRMS) eingesetzt, um die strukturelle Vielfalt und funktionelle Gruppierungen zu erfassen. Auf Basis dieser Ergebnisse wurden ^{1}H-NMR-Spektren unter Verwendung von synthetischen Lignin-Oligomeren (SLOs) modelliert, um experimentelle Beobachtungen mit repräsentativen Molekülstrukturen zu verknüpfen.

Im dritten Teil wurde das Dampf-Flüssig-Gleichgewichtsverhalten (VLE) der PL-Fraktion anhand von Advanced-Distillation-Curve-(ADC)-Experimenten und thermodynamischer Modellierung untersucht. Dabei wurde eine Reihe von Modellverbindungen – von Dimeren bis zu Tetrameren – bewertet, um eine geeignete molekulare Repräsentation der schweren Fraktion zu identifizieren. Abschließend wurde das Flüssig-Flüssig-Gleichgewicht (LLE) von FPBOs

analysiert, um Phasentrennungsphänomene zu charakterisieren und die Vorhersagegenauigkeit des vorgeschlagenen Surrogatansatzes zu bewerten.

Insgesamt präsentiert die vorliegende Dissertation einen integrierten experimentellen und modellgestützten Ansatz zum Verständnis und zur Beschreibung des komplexen Phasenverhaltens von ligninabgeleiteten Oligomeren in FPBOs. Die Ergebnisse leisten einen Beitrag zu einer vertieften Beschreibung der Pyrolyseprodukte und unterstützen die Auslegung und Optimierung von Fraktionierungs- und Kondensationsprozessen, die auf die Herstellung maßgeschneiderter, stabiler Bioölfraktionen zur weiteren Aufwertung in Brennstoffe und phenolische Chemikalien abzielen.

Table of Contents

Chapter 1. Introduction

This chapter contains content of the following publications:

Pilot scale fast pyrolysis of lignin in a twin-screw reactor

Ana Cristina Corrêa de Araujo, Axel Funke, Antigoni Margellou, Konstantinos Triantafyllidis, Nicolaus Dahmen

Biomass and Bioenergy, 2025, 200, 108041. DOI: 10.1016/j.biombioe.2025.108041

Liquid–Liquid Equilibrium Prediction in Fast Pyrolysis Bio-Oil Systems: A Framework for Incorporating Bio-Oil Complexity

Myriam Rojas Salas, Frederico Gomes Fonseca, Ana Cristina Corrêa de Araujo, Manuel García-Perez, Axel Funke

Energy & Fuels, 2024, 38, 19, 18769–18780. DOI: 10.1021/acs.energyfuels.4c03387

1.1 Context and motivation

Rising global demand for sustainable energy and chemicals has increased the focus on renewable resources to reduce fossil fuel dependence [1]. At the same time, greenhouse gas emissions are the main driver of climate change and have become a key focus of international climate policy [2]. Carbon neutrality goals, including those set in the Paris Agreement and the European Green Deal, call for a deep transformation of the energy and materials sectors through the adoption of renewable and climate-neutral resources [3,4]. Reaching these targets requires deep cuts in fossil fuel use across all sectors. This includes not only decarbonizing energy systems but also ensuring a sustainable supply of renewable carbon for material and chemical production [5].

Biomass, particularly lignocellulosic biomass, has emerged as a leading candidate due to its widespread availability and potential to generate diverse bio-based products [6]. In contrast to other renewable sources like wind or solar, biomass provides not only energy but also a sustainable carbon feedstock for the production of chemicals and materials [7]. This versatility enables conversion into heat, power, liquid fuels, and a broad portfolio of bio-based chemicals, supporting the transition to a sustainable economy across multiple sectors.

Of the three main components of lignocellulosic biomass – cellulose, hemicellulose and lignin – the latter remains the least exploited [8,9]. Lignin is a complex phenolic and aromatic polymer, and it is particularly noteworthy due to its high carbon content and unique chemical structure, making it a promising renewable source of fuels and aromatic chemicals [9]. It is mainly produced as a waste in the pulp and paper industry, during the Kraft process, generating over 50 million tons per year, most of which is simply combusted to recover energy [10]. In addition, lignin is also expected to be produced as a byproduct in the new biorefineries, which produce biofuels usually through processes aimed at the valorization of biomass sugars [11]. Therefore, lignin valorization is essential for developing sustainable biorefineries and advancing the transition towards a bio-based economy.

Despite its advantages, lignin remains one of the most difficult biomass components to valorize. Its heterogeneous structure and the poorly resolved nature of many of its conversion products can pose major challenges for both processing and downstream upgrading [12]. In this context, a deeper understanding of lignin's structural features and the composition of its products is crucial for developing effective processing routes toward sustainable fuels and chemicals.

1.2 Lignin: structure, composition and sources

Lignin, alongside cellulose and hemicellulose as a key structural component of biomass, is an amorphous and irregular biopolymer, primarily composed of three monomers: *p*-coumaryl alcohol (H-unit), coniferyl alcohol (G-unit), and sinapyl alcohol (S-unit) (Figure 1) [13]. These monomers differ in their degree of methoxylation, which directly affects the overall structure and composition of the lignin polymer.

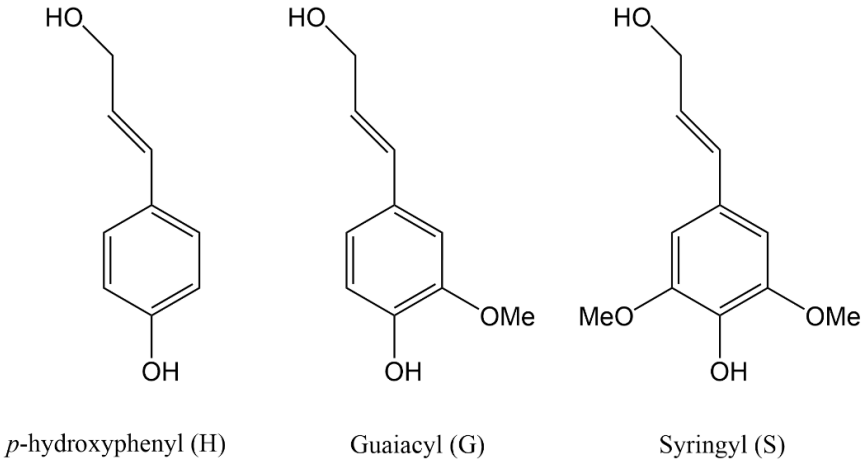


Figure 1. Lignin monomers

The monomers are linked through a variety of ether and carbon-carbon (C-C) bonds, forming a highly branched network (Figure 2) [13]. Among these, ether bonds account for the majority of linkages in native lignin, typically comprising over two-thirds of the total [14]. The β-O-4 aryl ether bond is the most abundant linkage in native lignin and plays a central role in defining its polymeric structure (Table 1) [14,15]. Other linkages include α-O-4 and 4-O-5 ether bonds, as well as various C-C bonds such as β-5 (phenylcoumaran), β-β (resinol), β-1 (diarylpropane), and 5-5 (biphenyl) structures [14,15].

Table 1. Types and Frequency (%) of linkages in Lignin molecule [14–17]

Linkage	Softwood lignin	Hardwood lignin	Grass lignin
β-O-4	45-50	60-62	74-84
α-O-4	6–8	3-11	5-11
4-O-5	4-8	1-7	–
β-5	9-12	3-11	5-11
β-β	2-6	3-12	1-7
β-1	1-9	1-7	–
5–5	5-18	3	–

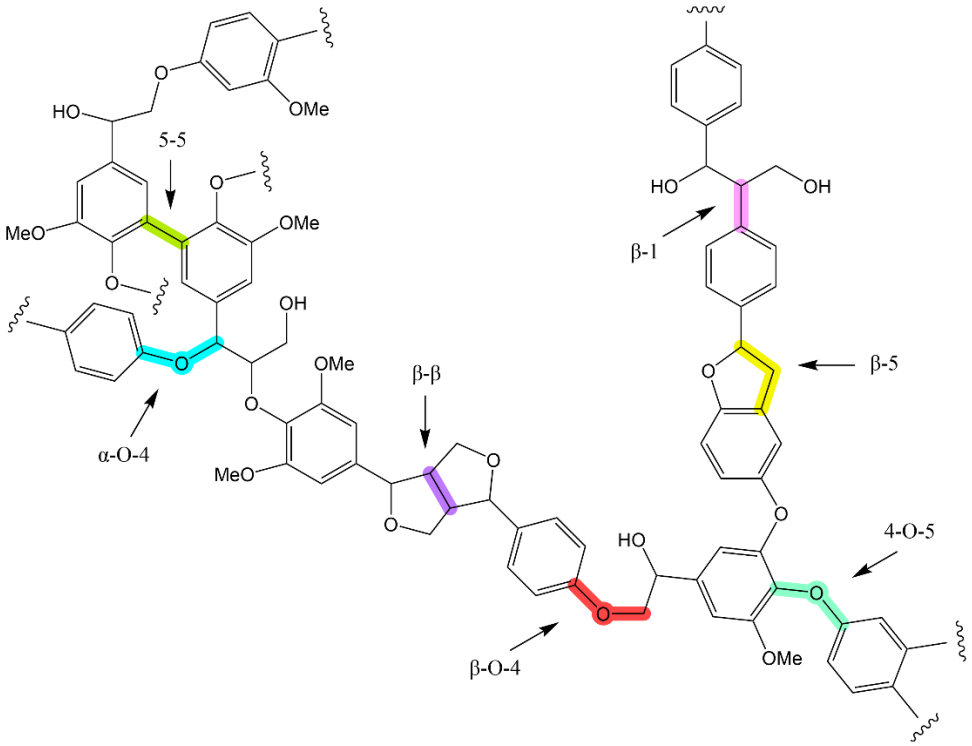


Figure 2. Lignin structure and main linkages, adapted from [15].

The composition, structure, and molecular characteristics of lignin vary significantly depending on the plant source. Softwoods contain the highest lignin content, typically ranging from 25-35 wt.%, followed by hardwoods (15-25 wt.%) and grasses (10-20 wt.%) [17,18]. In terms of monomer composition, softwood lignins consist almost exclusively of guaiacyl (G) units, leading to a more condensed polymer structure with a greater proportion of carbon-carbon linkages [18]. Hardwood lignins, in contrast, contain both guaiacyl (G) and syringyl (S) units, with a higher abundance of β-O-4 linkages and generally lower degrees of condensation [18]. Grasses exhibit all three types of units (G, S, and H), resulting in a more heterogeneous lignin structure [18]. These compositional differences also influence the average molecular weight of lignin, with softwood lignins typically exhibiting higher molecular weights than those derived from hardwoods [19].

The structure of technical lignin is shaped not only by the type of biomass but also by the conditions of its extraction process. During separation, both lignin-carbohydrate and internal lignin linkages can be broken [16]. Ether linkages, particularly the dominant β-O-4 bonds, are especially prone to cleavage due to their lower bond dissociation energy, whereas more resilient carbon-carbon bonds typically survive the process and remain in the final structure [15]. In many cases, this initial fragmentation is followed by repolymerization or condensation reactions, particularly under harsh thermal or alkaline conditions, resulting in significant structural modification [20].

Various separation processes have been developed to isolate lignin from lignocellulosic biomass. Each method applies different chemical and thermal conditions, which ultimately shape the structure and reactivity of the isolated lignin [21]. In alkaline processes such as Kraft and soda pulping, lignin is solubilized under high pH and elevated temperatures, which promotes cleavage of ether bonds and often leads to significant condensation [21]. While Kraft lignin contains sulfur due to the use of sulfur-based reagents, soda lignin is sulfur-free, making it more attractive for applications sensitive to sulfur contamination [21]. Organosolv methods employ organic solvents (e.g., ethanol, acetone) under acidic or neutral conditions, generally resulting in lignins with lower molecular weights and less structural degradation [21,22]. Hydrolysis lignins, typically obtained from acid-catalyzed pretreatment steps in bioethanol production, are often highly condensed due to severe processing conditions and limited solubility [23]. Variations in lignin extraction processes lead to a broad spectrum of technical lignins with different condensation degree, molecular weights and functional group distributions, factors that must be carefully considered in downstream applications.

1.3 Lignin conversion routes

The effective conversion of lignin into fuels, chemicals, and materials is a critical step toward the full valorization of lignocellulosic biomass. Various conversion strategies have been explored, including thermochemical, catalytic, and biological approaches.

Thermochemical approaches represent the most developed routes for the valorization of lignin, with pyrolysis and hydrothermal liquefaction (HTL) being the two most widely studied. Both enable the depolymerization of the lignin macromolecule and the conversion into liquid bio-oils [21,24]. Fast pyrolysis is carried out at moderate temperatures with rapid heating rates and short vapor residence times, conditions that suppress extensive repolymerization and char formation and thus favor the recovery of liquid products [25]. In comparison, hydrothermal liquefaction takes place in hot compressed water, typically at 250–450 °C and under elevated pressures, conditions that promote the solubilization and depolymerization of lignin [26].

Catalytic approaches such as hydrogenolysis and solvolysis, typically promoted by metal or acid catalysts, enable more selective cleavage of ether bonds, especially the abundant β-O-4 linkages, under comparatively mild reaction conditions [21,27]. These strategies are often combined with reductive or oxidative environments to tune product distribution, leading to higher selectivity toward phenolic monomers or stabilized intermediates [27]. Catalytic depolymerization is therefore regarded as a complementary route to thermochemical methods,

offering potential pathways to tailored phenolic building blocks that can be further integrated into biorefinery concepts for fuels, chemicals, and materials [28].

More recently, biocatalytic depolymerization using fungi or bacterial enzymes (e.g., DyP-type peroxidases, laccases) is gaining attention due to its selectivity and milder conditions, although scalability remains limited [27]. Additionally, novel approaches such as microwave-assisted, supercritical fluid, and ionic liquid treatments are being investigated to improve depolymerization efficiency and product control [21,27].

Despite significant progress, several challenges continue to hinder the efficient and scalable conversion of lignin. A first difficulty is the intrinsic heterogeneity of its structure and functional groups, which vary widely depending on botanical origin and isolation process, leading to inconsistent reactivity and product distributions [29]. Under thermochemical conditions, the aromatic and condensed nature of lignin promotes extensive repolymerization and char formation, reducing liquid yields. Catalytic methods face additional hurdles, including catalyst deactivation by impurities (e.g., sulfur in Kraft lignin) and by char deposition, as well as limited selectivity when targeting specific monomers [30,31]. Biological approaches, while highly selective, are still limited by low conversion rates and scalability [27]. Finally, across all routes, process integration and economic viability remain major bottlenecks, as the complexity of lignin-derived product mixtures makes separation and valorization challenging [29].

1.4 Fast pyrolysis of lignin: fundamentals and challenges

Among the available conversion routes, fast pyrolysis has emerged as a promising strategy for lignin valorization due to its scalability, process flexibility, and ability to produce a liquid bio-oil rich in phenolic compounds. Fast pyrolysis is the process defined by rapid heating rates, short vapor residence times, and moderate operating temperatures (typically 450–550 °C) [25]. It produces three product streams: a liquid fraction (fast pyrolysis bio-oil, FPBO), a solid char, and non-condensable gases, with the operating conditions favoring the formation of liquids over the other products [25]. While traditionally applied to lignocellulosic biomass, its application to isolated lignin presents distinct challenges and opportunities.

Compared to whole biomass, lignin pyrolysis exhibits significantly different product distributions. Lignocellulosic feedstocks, typically wood, yield ~75% bio-oil, ~12% char, and ~13% gas [25], whereas lignin pyrolysis leads to lower liquid yields, around ~35%, and higher char formation, up to ~45%, due to its highly aromatic and thermally resilient nature [18]. Besides, the composition of the pyrolysis bio-oil also shows clear differences: while lignin pyrolysis oil tend to be enriched in phenolic and aromatic oxygenates, lignocellulosic biomass

pyrolysis oil yields a larger share of light oxygenated species (e.g. furans, aldehydes, ketones) derived from carbohydrate decomposition [32].

The thermal decomposition of lignin takes place over a considerably broader temperature range (160-900 °C) compared to polysaccharides such as cellulose and hemicellulose, which typically decompose between 220-400 °C [18]. Within this range, two broad regimes can be distinguished: primary pyrolysis reactions (200–400 °C), which mainly release 4-substituted guaiacols and syringols with unsaturated side chains; and secondary reactions (>400 °C), during which guaiacols and syringols further transform into catechols, cresols, and phenols, C–C bond cleavage becomes more pronounced, and condensation pathways give rise to coke and polycyclic aromatic hydrocarbons (PAHs) [33].

This decomposition behavior helps to explain the main difficulties associated with lignin fast pyrolysis. Low liquid yields are typical, while high char formation (Figure 3A) not only reduces the amount of recoverable oil but can also lead to operational issues such as fouling of equipment and poor heat transfer [34]. In addition, lignin's tendency to soften and melt before it fully decomposes can cause plugging and agglomeration in feeding and reactor systems (Figure 3B), making continuous operation more challenging than with lignocellulosic biomass [34,35].

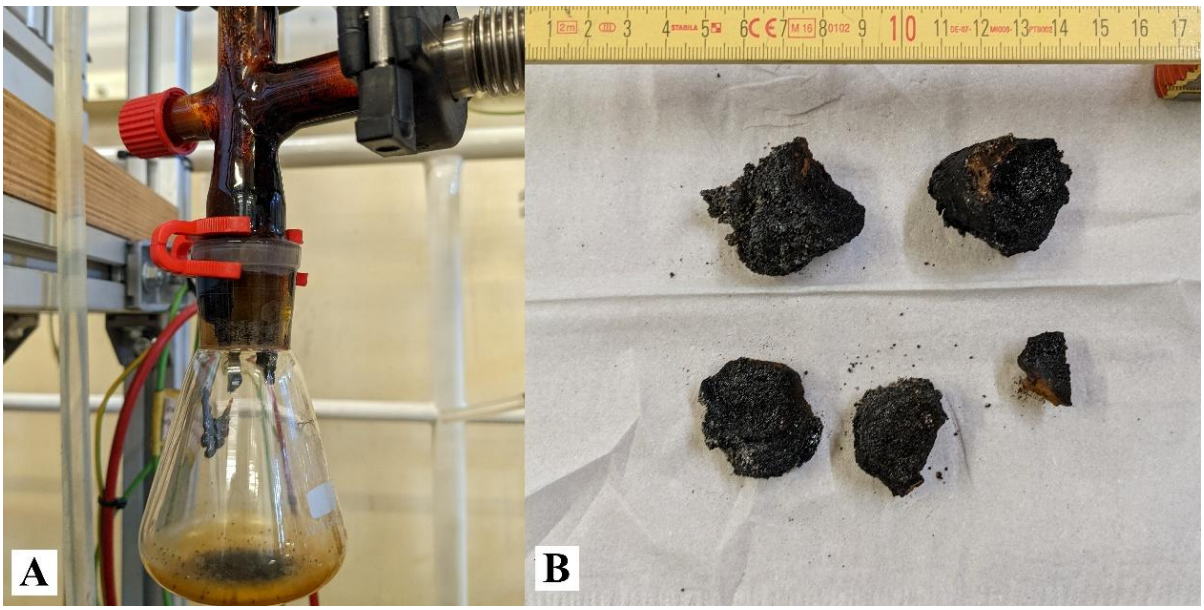


Figure 3. A) Tar deposits formed in the condenser; **B)** Melted lignin agglomerates removed from the reactor after a lignin fast pyrolysis experiment at IKFT.

Given these challenges, several strategies have been explored to improve the handling and processing of lignin during fast pyrolysis. One approach to mitigate melting issues involves the pretreatment of lignin with additives, such as calcium hydroxide or mineral clay [37,38]. Such pretreatments have been shown effective in preventing lignin from melting. However, their mechanism of action is not well understood, and they can lead to undesirable side effects, such as increased ash production. In addition, pretreatments can be expensive.

Another strategy to address such challenges is to enhance pyrolysis units by using equipment specifically designed for lignin conversion. Feeding systems can be optimized to work with lignin. Feeding screws can be cooled to prevent lignin from melting, but they only have been proved effective to a certain extent [39,40]. Slug injectors have been tested as feeders for lignin pyrolysis and proven more efficient in preventing melting and blockage [41,42]; the system works by using a pneumatically activated pinch valve to release small amounts of lignin into a tube, forming a slug, then intermittent pulses of inert gas, controlled by a solenoid valve, propel the slug into the reactor.

Reactor selection and design plays a curtail role in the processing of lignin, besides, they also have a great impact on the efficiency, yield, and scalability of the pyrolysis process. Fluidized beds are the most commonly used reactors for lignin fast pyrolysis experiments, but this choice can present challenges for the process [43]. Agglomerates can disrupt the fluidization of particles within the reactor, leading to poor mixing, uneven temperature distribution and in worst cases blockage of the reactor [39]. To overcome problems with bed agglomeration, Gooty et al. [41] and Pienihakkinen et al. [44] inserted a mechanical stirrer inside the fluidized bed to break up the agglomerates formed. Besides fluidized bed, a small number of studies is reported for other types of reactors [43]. The fixed bed reactor has shown promising results but this is only for experiments performed in small scale [45,46]. The pyrolysis centrifuge reactor was successful in pyrolyzing lignin, with rotational forces inhibiting the formation of agglomerates inside the reactor [47]. However, this reactor type may face difficulties when scaling up since its process is surface area dependent. The entrained flow reactor also proved capable to pyrolyze lignin but shows the disadvantages of having poor heat transfer and being hard to scale up [48].

Auger reactors, though untested for lignin pyrolysis, could potentially be a viable option. Widely applied across slow, intermediate and fast pyrolysis [49–51], they are capable of handling diverse feedstocks, including challenging materials like what straw and *Miscanthus*, validated at Technology Readiness Levels (TRL) 6-7 [52,53]. The twin-screw auger reactor, in particular, may be a good choice due to their enhanced mechanical mixing, which makes them more suitable for sticky and cohesive materials, such as lignin [49]. The screw-driven transport

and mixing inherently features mechanical break up, which has been proven necessary for processing lignin for pyrolysis in fluidized bed reactors [49]. At the Institute of Catalysis Research and Technology (IKFT), a twin-screw reactor unit in the "Python" process development unit, is available for testing [54].

1.5 Lignin fast pyrolysis bio-oil (FPBO): Composition and properties

Fast pyrolysis bio-oil (FPBO) is a dark, viscous liquid formed by condensation of the vapors generated during thermal decomposition of biomass. It is a chemically complex emulsion containing several hundred oxygenated species that span a broad range of molecular weights and functionalities [55]. Its composition includes water, light oxygenated compounds such as acids, alcohols, aldehydes, and ketones, monomeric phenols, sugars, and heavier fraction of oligomers, often referred to as pyrolytic lignin [55,56].

As the pyrolysis of lignin and lignocellulosic biomass differs, so do the compositions of their resulting oils. Lignin-derived FPBO is characterized by a high content of phenolic compounds, including guaiacols, syringols, and their alkylated or demethoxylated derivatives, as well as a considerable fraction of heavier oligomers [47]. In addition to phenolics and oligomers, lignin FPBO also includes water and light oxygenated species, though these are generally less abundant than in oils from whole biomass [47]. In contrast, FPBO from lignocellulosic feedstocks contains a broader distribution of light oxygenates originating from carbohydrate decomposition, including hydroxyacetaldehyde, acetic acid, acetol, furans, and anhydrosugars such as levoglucosan, in addition to a smaller but still significant fraction of phenolic compounds and heavier oligomers [55]. These compositional differences are illustrated in **Table 2**, which show main functional groups identified by GC–MS for lignin and wood derived oils obtained under identical fast pyrolysis conditions [47].

These compositional differences are directly reflected in the properties of the resulting oils. Lignin FPBO generally contains less water (10–20 wt.% compared to 15–30 wt.% for biomass FPBO) and is less acidic (pH 3.5–4.5 vs 2.5), while its higher content of phenolic monomers and oligomers translates into higher heating values (25–30 MJ/kg) and greater viscosity [57,58]. In contrast, the abundance of light oxygenates and sugars in biomass-derived FPBO contributes to its higher water content, stronger acidity, and lower heating values (16–19 MJ/kg), as well as lower viscosity [58]. These property contrasts are summarized in Table 2.

Table 2. Comparative physicochemical properties and functional group composition of lignin and wood bio-oils [40,44,47,55,57,58].

Property	Lignin bio-oil	Wood bio-oil
Water content (wt.%)	10-20	15-30
pH	3.5-4.5	2.5
HHV (MJ/kg)	25-30	16-19
Viscosity (cP at 50 °C)	100-600	40-100
Elemental composition (wt.%):		
C	62-69	54-58
H	6.5-7.5	5.5-7.0
O	23-28	35-40
Main functional groups (% peak area):		
Acids	6.15	25.88
Ketones	2.55	13.3
Furans	1.01	8.61
Glycerol	2.91	-
Levoglucosan	2.55	7.10
Phenols	2.89	0.48
Guaiacols	27.4	11.4
Syringols	18.3	10.3

1.6 Characterization of lignin oligomers

The characterization of FPBO composition is essential for understanding its reactivity, stability, and potential for valorization. Conventional analytical approaches such as Fourier transform infrared spectroscopy (FTIR), gas chromatography–mass spectrometry (GC–MS), and nuclear magnetic resonance (NMR) have been widely employed for bio-oil analysis [59]. Among these, GC–MS remains the most commonly used method because of its relatively low cost and broad availability. However, GC–MS relies on volatility for separation, and thus can only provide reliable identification for low-molecular weight compounds. As a result, typically no more than 40 wt.% of the bio-oil can be identified by GC–MS [60,61]. The remaining fraction is composed of heavier, less volatile compounds that cannot be vaporized, leaving much of the oil chemically unresolved.

This limitation is particularly significant for lignin-derived oils, where a large fraction of the material consists of high-molecular weight species such as oligomers. Although these heavy fractions are poorly detected by GC–MS, they are of critical importance as they originate largely from lignin decomposition and account for a substantial portion of the oil's mass.

To address these limitations, advanced analytical methods have been developed to investigate the unresolved high-molecular-weight fraction. Ultra-high-resolution mass spectrometry (UHRMS) has recently emerged as a transformative tool in this field, offering unmatched sensitivity and accuracy for characterizing highly complex mixtures [62,63]. Instruments such as Orbitrap and Fourier transform ion cyclotron resonance (FTICR) allow detection of thousands of individual molecular species, with precise mass determination and molecular formula assignment [63,64]. FTICR provides the highest mass resolution currently available, reaching resolving powers of several million, which makes it particularly suitable for analyzing complex mixtures like pyrolysis bio-oils. Orbitrap, while somewhat lower in resolution, combines high mass accuracy with greater accessibility and robustness, making it also a valuable tool for analysis of lignin-derived oils [65].

Beyond molecular identification, UHRMS provides structural insights through derived parameters such as double bond equivalents (DBE) and heteroatom class distributions, which help describe the degree of aromaticity, unsaturation, and oxygenation [65]. When combined with complementary techniques, such as FTIR, NMR and GC-MS, the UHRMS analysis form a robust basis for the structural elucidation of lignin-derived oligomers, the fraction broadly termed "pyrolytic lignin" [64,66,67].

1.7 Fractionation strategies for FPBO

To improve FPBO's properties and broaden its applicability, different fractionation and upgrading techniques have been explored to reduce oxygen content, enhance stability, and increase its calorific value, as well as isolate valuable compounds for potential commercialization [68,69].

Fractional condensation (Figure 4) is a commonly employed strategy, separating pyrolysis vapors into distinct fractions based on their dew points [70]. It is a practical approach that can directly be integrated into the pyrolysis vapor recovery system, and for this reason is has gained attention as a cost-effective alternative for tailoring bio-oil fractions [70].

Staged condensation systems can effectively split FPBO into fractions with distinct compositions. Early work with multi-stage setups showed that the first condensers generally capture heavier, water-insoluble fractions rich in pyrolytic lignin, while later stages concentrate

lighter oxygenates and water [57,71]. Increasing the number of condensation steps improves the differentiation of fractions, although the separation of individual species remains challenging due to overlapping volatilities and functional group effects [71]. Nonetheless, fractional condensation provides a means to control water content, acidity, viscosity, and chemical distribution among fractions, offering a route to tailored product streams with improved stability or targeted chemical functionalities [72,73].

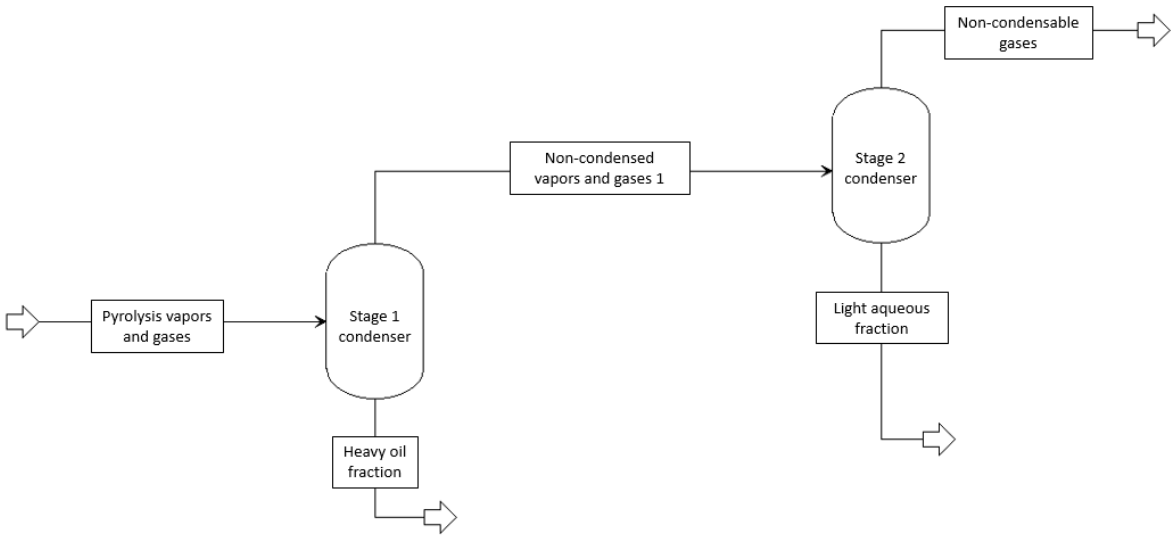


Figure 4. Fractional condensation scheme.

Liquid-liquid extraction offers another method for FPBO fractionation, enabling the separation into different chemical groups according to their polarities and improving the stability of the resulting fractions [74,75]. This technique separates compounds based on functionalities, such as water washing to remove sugars, or light organics to remove phenolics [76,77]. Interesting applications for extracted fractions range from upgrading to biofuels, fermentation/sugar chemistry, extraction of pyrolytic lignin and others).

Figure 5 shows a liquid-liquid extraction scheme commonly used to fractionate FPBO, starting with removing extractives and solids (Adapted from [78]). By adding water, it is possible to tip the balance in the emulsion to produce two main fractions. The water-soluble fraction (WS in Figure 5) is mostly comprised of oxygenated compounds (aldehydes, ketones, carboxylic acids, and sugars, 70-80 wt.% dry), while the water-insoluble fraction (WI in Figure 5) concentrates the majority of the phenolics present in the original mixture, as well as a significant fraction of the furanes [79]. A second extraction stage frequently employs organic solvents, such as dichloromethane (DCM in Figure 5), chloroform, or diethylether, and can further elucidate the complexity of each fraction [80]. While the WS soluble fraction has not

received great attention, the extraction products of the WI fraction can be attributed to small phenolics, mono and anhydrosugars, and hybrid molecules (WIS-DS), while lignin-derived oligomers (LO) populate the altogether insoluble phase (WIS-DIS) [80–82].

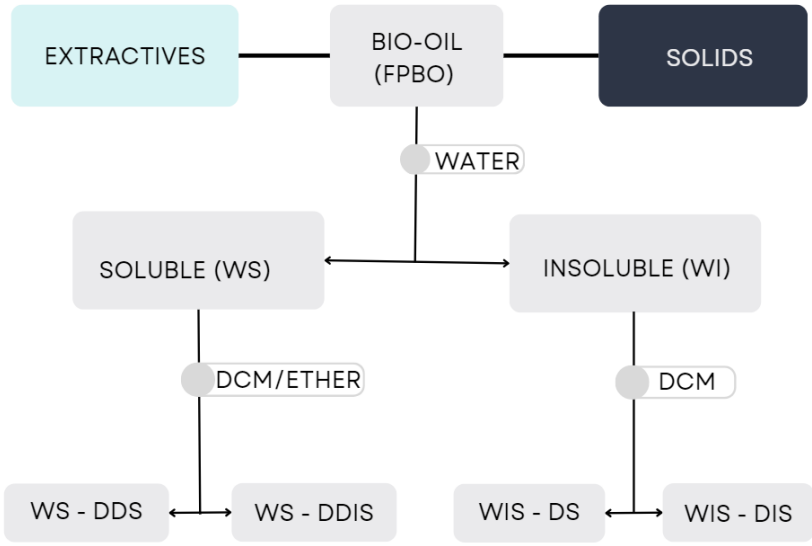


Figure 5. Solvent fractionation scheme. DCM: Dichloromethane (CH_2Cl_2); WS-DDS: Water soluble-DCM+DEE (Diethyl ether, $C_4H_{10}O$) soluble; WS-DDIS: Water soluble-DCM+DEE insoluble; WIS-DIS: Water-insoluble, DCM insoluble; WIS-DS: Water-insoluble, DCM soluble. Adapted from Oasmaa et al., (2003) [78]. Copyright 2003 American Chemical Society.

Advancing separation techniques such as fractional condensation and liquid–liquid extraction further requires the support of predictive modeling [68]. The complex phase behavior of FPBO stems from its diverse molecular composition and the intermolecular interactions that determine how species distribute between phases [68]. Understanding and simulating these interactions is therefore essential for accurately predicting phase equilibria and guiding process design and optimization.

1.8 Modeling approaches for lignin oligomers

Modeling the behavior of FPBO is essential to predict their phase behavior, guide separation strategies, and support process design [83]. Reliable predictions require an adequate description of the FPBO's diverse composition as well as accurate representation of its intermolecular interactions; however, the intrinsic chemical complexity of FPBO poses major challenges [84]. This complexity arises from the presence of hundreds of compounds with varying molecular weights, polarities, and functional groups, which lead to non-ideal behavior, limited availability

of thermophysical data for many of these species, and the existence of numerous unidentified compounds [84].

To address the compositional complexity of FPBO, many studies employ surrogate mixtures composed of representative compounds or functional groups that capture the main chemical features of the oil [73,84]. Volatile components can often be identified and quantified through GC–MS/FID analysis, allowing their direct inclusion in such mixtures [59]. In contrast, the heavy molecular weight fraction, commonly referred to as pyrolytic lignin (PL), is far less understood, as its structural diversity and limited volatility hinder direct chemical characterization [61].

Although less understood, the presence of the PL has a pronounced impact on the physicochemical properties of FPBO, including viscosity, phase stability, solubility, and reactivity [79,85]. These properties, in turn, affect downstream processing steps such as upgrading, separation, and utilization in fuel or chemical applications. Besides, it has been shown that the representation of this fraction is particularly important for improving the predictive capability of phase equilibrium models [86,87]. Therefore, accurately modeling this fraction is essential to achieve reliable phase behavior predictions of FPBO.

Multiple representations of the nature of the PL fraction have been proposed, but there is little consensus towards its modeling. Fonseca and Funke [84] discuss several proposals and their effectivity in modeling this fraction in vapor-liquid equilibrium (VLE) calculations in their study. Several structures have been proposed for the modeling of the PL fraction, based both on insights from advanced analytical characterization and computational modeling [86,88,89].

Among various efforts reported in the literature, Fonts et al. [88] employed a surrogate-based approach using molecular structures proposed by Terrell et al. [64], which originated from high-resolution FT-ICR-MS analysis of lignin. These surrogates were used to estimate a range of thermophysical properties for the heavy fraction by combining ab initio and group-contribution methods, providing essential data for future modeling efforts. Manrique et al. [89] applied a complementary strategy, using density functional theory (DFT) to generate and assess the thermodynamic stability of dimeric to tetrameric lignin oligomers, yielding computationally derived molecular candidates and associated thermochemical properties. Ille et al. [86] investigated smaller phenolic dimers as potential surrogates to model the activity coefficient of water in pyrolysis oils; 3,3′,4-biphenyltriol was identified as the most suitable compound, a finding later validated by Krutof and Hawboldt [90] in distillation and vapor–liquid equilibrium modeling.

Beyond proposed surrogate molecules based on analytical and/or computational approaches, recent studies have also explored the use of synthetic lignin model compounds [91]. Such synthetic lignin oligomers (SLOs) are laboratory-made molecules designed to reproduce specific structural features of lignin, allowing controlled experiments to investigate their reactivity, solubility, and phase behavior [92]. Moreover, these compounds enable direct measurement of key physicochemical properties, providing a robust experimental basis for improving thermodynamic correlations and validating modeling approaches [92].

A summary of some of these molecular structures proposed for representing the PL fraction is presented in Figure 6.

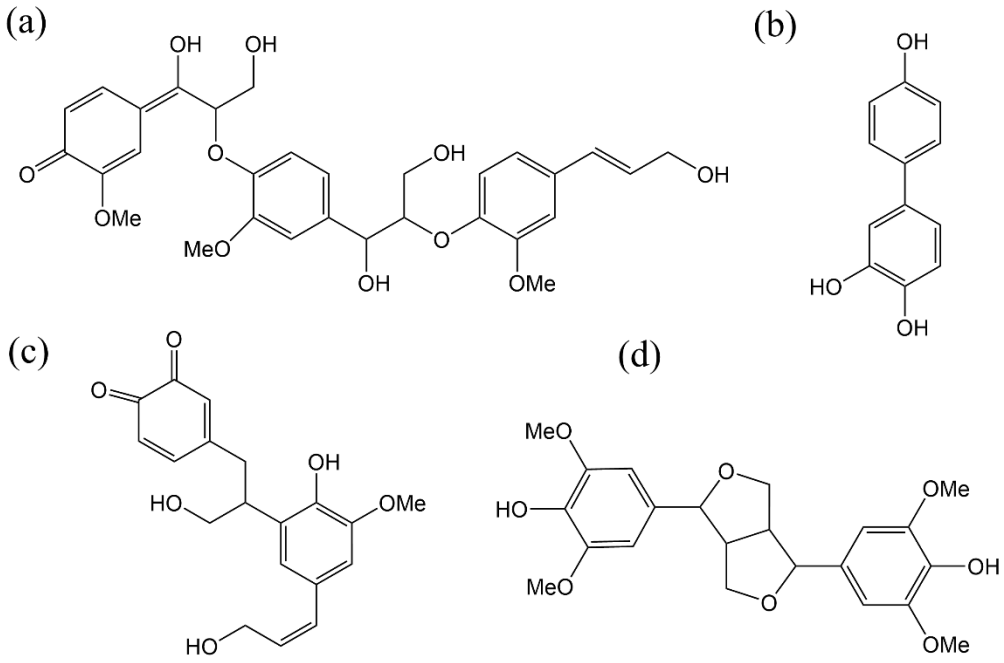


Figure 6. PL surrogate molecules. (a) PL surrogate used by Fonts et al. [88] to model the low molecular mass PL; (b) 3,4,4-biphenyltriol used by Ille et al. [86] and Krutof and Hawboldt [90]; (c) computational derived PL candidate from Manrique et al. [89]; (d) example of SLO [92].

In addition to lignin-derived oligomers, other high-molecular-weight fractions also contribute to the complexity of lignocellulosic pyrolysis oils. These include dehydrated sugar-derived structures, often referred to as humins, as well as hybrid oligomers formed through condensation reactions between carbohydrate and lignin derived intermediates [32,88]. Such components further influence the stability, solubility, and overall phase behavior of FPBO, underscoring the need for comprehensive approaches that account for all oligomeric fractions in thermodynamic and process modeling [79].

1.9 Research objectives and thesis scope

Against this background, the main objective of this PhD work is to establish a molecular representation of the lignin derived oligomers in fast pyrolysis bio-oil for predictive modeling of phase equilibria. To achieve this, the work combines fast pyrolysis experiments, advanced analytical characterization, and thermodynamic modeling in a stepwise approach. The following research questions guide the study:

1. How can stable and continuous lignin fast pyrolysis be achieved to produce representative bio-oils?

   This question is addressed in Chapter 2, which investigates the fast pyrolysis of lignins in a twin-screw reactor at pilot scale. The focus lies on enabling continuous operation despite the severe melting behavior of lignin, and on evaluating the resulting mass and energy balances, product yields, and bio-oil composition.
2. What is the molecular composition of lignin-derived bio-oils, and how can the heavy fraction be described?

   Chapter 3 provides a comprehensive characterization of lignin bio-oils, integrating conventional (FTIR, GC–MS, NMR) and advanced (Orbitrap UHRMS) analyses. To bridge the gap between identifiables and the unresolved heavy fraction, modeling of the ^{1}H-NMR spectra was performed using model molecules to evaluate their ability to represent the latter.
3. How can the phase behavior of pyrolytic lignin (PL) be experimentally assessed and thermodynamically represented?

   Chapter 4 focuses on the vapor–liquid equilibrium (VLE) behavior of the PL fraction in FPBOs. Experimental data from advanced distillation curve (ADC) of lignin FPBOs are used to validate surrogate mixture based approach in VLE models. The suitability of different structures (from dimers to tetramers) is assessed to determine their ability to represent the PL in predictive simulations.
4. Can the proposed structures be used to describe liquid–liquid equilibria (LLE) in FPBOs?

   This question is examined in Chapter 5, where the framework developed for VLE is expanded to describe LLE behavior in FPBOs.

Finally, Chapter 6 summarizes the findings and discusses how the experimental and modeling results contribute to a comprahensive representation of lignin derived oligomers in FPBOs. Together, the outcomes of this work contribute to a better understanding of lignin oligomers and their role in the behavior of FPBOs.

Chapter 2. Pilot scale fast pyrolysis of lignin in a twin-screw reactor

This chapter is based on the following publication:

Pilot scale fast pyrolysis of lignin in a twin-screw reactor

Ana Cristina Corrêa de Araujo, Axel Funke, Antigoni Margellou, Konstantinos Triantafyllidis, Nicolaus Dahmen

Biomass and Bioenergy, 2025, 200, 108041. DOI: 10.1016/j.biombioe.2025.108041

2.1 Introduction

As discussed previously in the general introduction in Chapter 1, lignin poses specific challenges during fast pyrolysis due to its tendency to melt and agglomerate, which can cause severe feeding and operational issues. Different mitigation strategies and reactor configurations have been proposed, yet ensuring stable and continuous operation, especially at larger scales, remains difficult. Addressing this challenge is essential to enable the valorization of lignin through fast pyrolysis.

In this context, the objective of this chapter is to evaluate the performance of the twin-screw reactor for fast pyrolysis of lignin in a pilot scale unit in terms of yield, efficiency, and operational stability. To address the challenges related to melting and agglomeration, modifications were made to the design of a pre-existing reactor. Specifically, the positions of the biomass and heat carrier inlets were reversed to prevent the lignin from melting on the screw. For the study, two types of lignin were tested: Indulin AT, a Kraft lignin representing the traditional technical lignin from the pulp and paper industry, and Miscancell, a sulfur-free alkali lignin derived from *Miscanthus*, representative of the lignin streams entering the market through alternative biorefinery routes beyond the Kraft process. To validate the yields of the pilot runs and the chemical composition of the bio-oils produced, comparative micro-pyrolysis experiments were conducted. The bio-oils produced were analyzed to determine their physicochemical properties and chemical composition.

2.2 Material and Methods

2.2.1 Materials

Two types of lignin were used for the fast pyrolysis experiments. Indulin AT lignin, a purified Kraft lignin from pinewood, free from hemicellulose, manufactured by Ingevity, North Charleston, S.C. USA and supplied by DKSH Switzerland Ltd. And, Miscancell lignin, a sulfur-free lignin extracted from Miscanthus via alkaline pretreatment, produced, manufactured, and supplied by Exegi IP Management B.V.

2.2.2 Micro-pyrolysis experiments

Micro-pyrolysis experiments (Py/GC-MS) were carried out in a Multi-Shot Micro-Pyrolyzer (EGA/PY-3030D, Frontier Laboratories, Japan) connected to a gas chromatographer-mass spectrometer system (GCMS-QP2010, Shimadzu) for fast screening of lignin pyrolysis. For each experiment, a sample of 1 mg of lignin was loaded in a stainless-steel cup, which was

quickly dropped into the hot reactor/furnace and pyrolysis was conducted at 500 °C for 12 s. The interface temperature between the micropyrolyzer and GC was set to 320 °C. Helium (99.999%) was used as the carrier gas at a flow rate of 156 ml/min. The analysis was performed using a capillary column (Ultra Alloy-5, 30 m × 0.25 mm and 0.25 μm film thickness), injector split ratio of 1:150 and column flow 1 ml/min, according to the following temperature program: initial temperature 40 °C, hold for 4 min and heating (5 °C/min) up to 300 °C, hold for 7 minutes. Both the injector and the detector temperatures were set at 300 °C. The mass spectra were recorded in the range of m/z=45-500 with a scan speed of 5000 amu/s. Identification of mass spectra peaks was achieved using of the scientific library NIST11s. The derived compounds were classified and categorized into 16 groups: mono-aromatics (AR), aliphatics (ALI), phenols (PH), acids (AC), esters (EST), alcohols (AL), ethers (ETH), aldehydes (ALD), ketones (KET), polycyclic aromatic hydrocarbons (PAH's), sugars (SUG) nitrogen compounds (NIT), sulfur compounds (SUL), oxygenated aromatics (OxyAR), oxygenated phenols (OxyPH) and unidentified compounds (UN).

2.2.3 Fast pyrolysis pilot unit

For the lignin fast pyrolysis experiments, the "*Python*" Processing Development Unit, operated at the Institute of Catalysis Research and Technology at the Karlsruhe Institute of Technology, was used. A flowsheet of the plant can be seen in Figure 7. The unit operates at a biomass feed rate of 10 kg/hr and features a twin-screw mixing reactor that ensures thorough mixing of solid biomass feedstock with a pre-heated heat carrier. The reactor design accommodates a range of feedstocks with varying bulk densities, including lignin powders, without compromising feeding consistency. The residence time of biomass and heat carrier within the reactor is between 10–15 seconds at a rotational frequency of 2 s^{-1} [93]. The heat carrier (steel beads) is recirculated using a bucket elevator to enable continuous flow and consistent thermal conditions, maintaining a heat carrier-to-biomass mass ratio of 100:1. The temperature of the steel beads leaving the reactor, measured as the 'reactor temperature,' is maintained using an electric heater. An inert gas flow (N_2) of 1 m^3/h is supplied during the experiments to maintain an inert atmosphere. Pyrolysis gases, vapors, and char fines are removed from the top of the pyrolysis reactor, with particles separated at reactor temperature in two serial cyclones. The primary product, fast pyrolysis bio-oil, is recovered using a quenching system typically operated at around 90 °C, followed by an electrostatic precipitator. This organic-rich condensate (ORC) is cooled and recirculated to quench the incoming vapor, avoiding the use of an additional quenching medium. In a second condensation step, aqueous

condensate (AC) is recovered in a tubular heat exchanger, operated at around 20 °C. Vapors are condensed by direct contact with recirculated and cooled aqueous condensate in a countercurrent flow. Non-condensable gases are analyzed online by gas chromatography and subsequently vented by a compressor. Gas yields are determined by introducing a tracer gas (neon) with controlled mass flow rate to the reactor. Neon is also quantified by gas chromatography every 15-20 min and the mean of these values are used to determine the gas flow rate over the experiment. The pyrolysis reactor operates under slight underpressure conditions (pressure difference <2 kPa), and the condensation process is similarly under slight underpressure. Ethylene glycol is used as the entrainer material for the first quenching system, while water is used for the second; the mass balance to the condensates refers to the mass of the final product discarding the initial mass of entrainer. Non-condensable gases are analyzed online by gas chromatography (Emerson, Daniel Modell 700) and subsequently vented by a compressor. A detailed description of the experimental procedure can be found elsewhere [94].

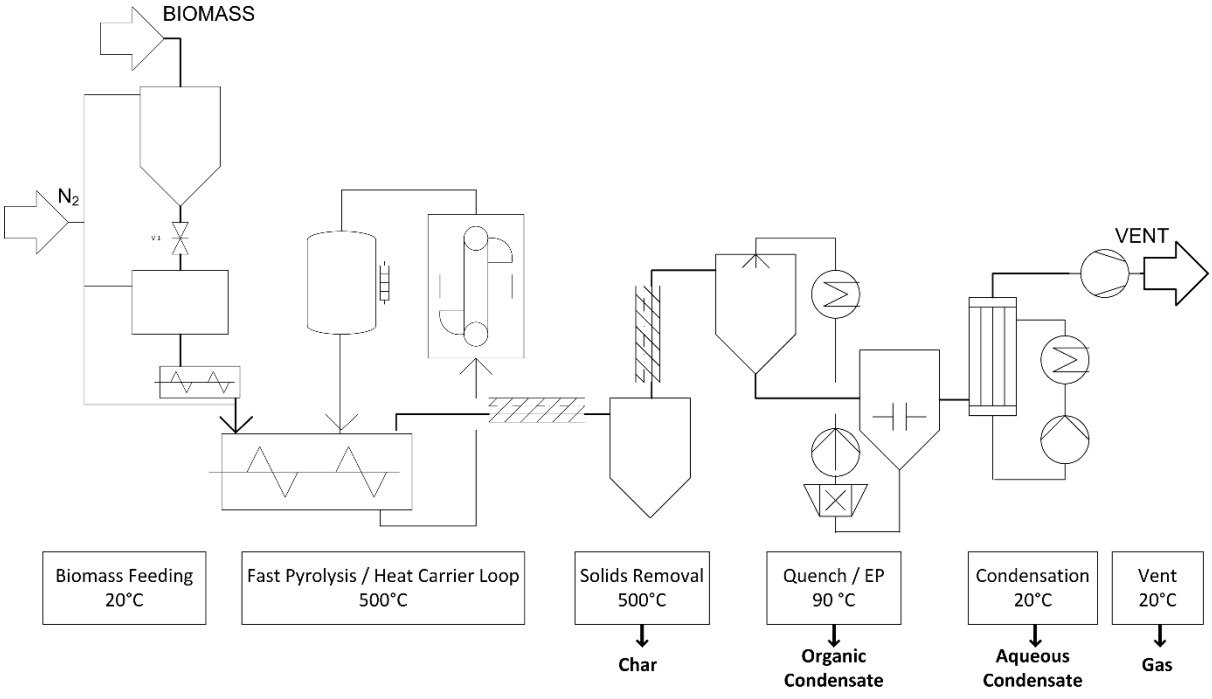


Figure 7. Flow scheme of Python fast pyrolysis unit [50].

In the original setup of the *Python* unit [50], the feedstock enters the twin-screw reactor at the cold inlet of the screw, while the heat carrier is added later by dropping it onto the feedstock transported by the screws, ensuring fast mixing and minimizing heat losses to the reactor material and the sweeping gas. Although this is not an issue for most biomasses tested in this

unit, this arrangement has proven prone to heavy plugging when processing lignin, due to agglomerates forming on the screw shafts. Severe plugging problems were observed in a test run with Indulin AT lignin, leading to complete blockage and shutdown of the unit after just a few minutes of feeding. To mitigate these issues, the feedstock and heat carrier inlets were reversed, as shown in Figure 8. By dropping the feedstock onto the moving bed of hot steel particles, it is expected that agglomerates will form within the moving bed rather than on the shaft of the screws, where the mechanical agitation could break them up.

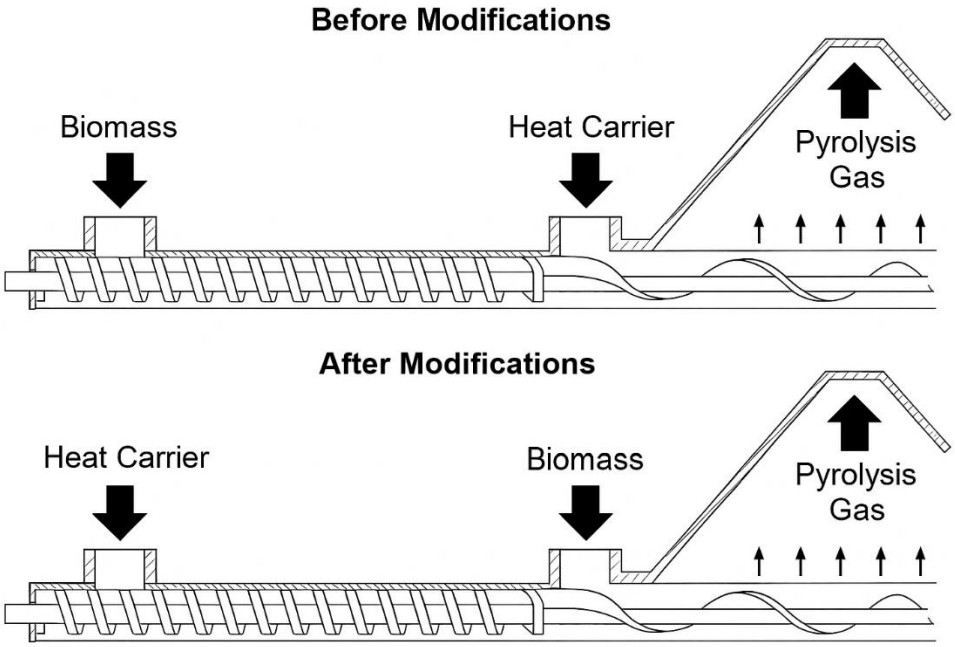


Figure 8. Scheme of changes to reactor setup to reduce agglomeration at the screw shaft

Several other hardware modifications were made to facilitate the use of lignin, including adjustment of the gas outlet to avoid fine lignin particles bypassing the agitated bed of heat carrier. Previously, the gas outlet covered the entire length of the reactor from the heat carrier inlet to the solid particle outlet, posing a high risk for shortcut flow of biomass particles if they were inserted instead of the heat carrier particles. The gas outlet was moved to the end of the auger reactor and reduced to a pipe with a small cross-section (around 60 mm). All experiments reported in this manuscript have been conducted with this modified setup.

2.2.4 Analytical methods

Characterization of the lignin feedstocks and pyrolysis products is crucial for understanding the pyrolysis process and the resulting products. The feedstocks and products were

characterized using various techniques. Total solids, moisture, and ash content were determined gravimetrically according to the NREL protocols (NREL/TP-510-42621 and NREL/TP-510-42622) by heating air-equilibrated samples at 105 °C and 575 °C, respectively. Elemental analysis of C/H/N/S content was performed using an Eurovector EA3100 Series CHNS-O analyzer. The oxygen content was calculated by difference. Higher Heating Values (HHV) were computed according to Demirbaş [95] using Eq. (1), where C, H, O, and N, refer to the mass fractions of carbon, hydrogen, oxygen and nitrogen in the material, respectively. Molecular weight of lignins were determined by Gel Permeation Chromatography (GPC), using a Shimadzu Corporation instrument with Shodex KF-801, KF-802.5, and KF-803 columns. THF was used as eluent at 0.5 mL/min, with columns and injection system maintained at 42 °C. Results were given relative to polystyrene standards within the calibration range (162–500.000 Da). Prior to the measurements, lignin were dissolved in THF (1–2 mg/mL) and filtered.

Thermal degradation profiles and residual mass of the lignin feedstocks were determined via thermogravimetric analysis (TGA/DTG). The samples were heated from room temperature to 950 °C at a heating rate of 10 °C/min, under N_2 gas at a flow rate of 50 mL/min. Fourier Transform Infrared Spectroscopy (FTIR) spectra were obtained using a Perkin-Elmer spectrometer with the KBr pellet method, covering a range of 4000 to 400 cm^{-1}. Inorganic materials were analyzed via Inductively Couple Plasma (ICP) spectroscopy according to the DIN EN ISO 11885 standard. Modified Karl Fischer titration was used to determine water content in the pyrolysis condensates. Gas chromatography-mass spectrometry/flame ionization detector (GC-MS/FID) analysis of the pyrolysis condensates were performed by the Thünen Institute (Hamburg, Germany), for identification and quantification of components, using internal standards and calibration procedures. A detailed description of the applied method can be found in Windt et al. [96].

Energy balances were calculated based on the mass yields of the pyrolysis products and their respective HHVs. The HHVs of the lignin, ORC and char were estimated according to Eq. (1), using the Demirbaş [95] correlation, and the HHV of the gas was calculated from the gas composition using the standard combustion enthalpy for each species. The aqueous condensate (AC) was excluded from the energy balance due to its high water content and negligible energy contribution.

$$HHV = (33.5[C] + 142.3[H] - 15.4[O] - 14.5[N]) \times 10^{-2} \quad (1)$$

2.3 Results and Discussion

2.3.1 Lignin characterization

The physicochemical properties of the lignin feedstocks are presented in Table 3, highlighting the differences between Indulin AT and Miscancell lignin. The moisture and ash content in the Miscancell lignin are significantly higher compared to Indulin AT, while elemental analysis indicates that Indulin AT has a higher carbon content than Miscancell, which corresponds to higher lignin content and greater purity, with less ash and fewer cellulose/hemicellulose impurities. Indulin AT also shows the presence of sulfur, expected as it is resulting from the Kraft process, whereas Miscancell is sulfur free, as advertised. On the other hand, Miscancell shows a higher content of nitrogen, which could be attributed to the higher protein content of *Miscanthus* compared to pinewood [97]. Molecular weight values, determined through gel permeation chromatography, suggest that the Indulin AT lignin is composed of larger fractions, while Miscancell is composed of smaller molecules. The similar polydispersity index (PDI) values indicates that both lignin show a broad distribution of molecular weights.

Table 3. Physicochemical properties of lignin feedstocks.

Lignin Sample	Moisture (wt. %)	Ash[1] (wt. %)	C (wt. %)	H (wt. %)	N (wt. %)	S (wt. %)	O[2] (wt. %)	M_N (g/mol)	M_W (g/mol)	PDI
Indulin AT	4.4	2.4	67.7	6.0	1.0	1.7	21.2	1318	4377	3.3
Miscancell	13.4	14.1	55.8	6.7	2.9	0.0	20.5	844	2587	3.1

[1] Ash content and elemental analysis data are given on dry lignin basis. [2] Calculated by difference.

Due to the elevated ash content found in the Miscancell lignin, an inorganic material analysis was conducted. The results of ICP analysis reveal that Miscancell lignin contains more calcium compared to *Miscanthus* samples values from the literature [53,98]. This could be related to the specific *Miscanthus* type used as feedstock for the lignin extraction, or it could be a consequence of the extraction process itself, where the calcium content is increased as it is concentrated in the lignin fraction, while the overall biomass is reduced to about one-third (see Table 4). The high silicon content present in the sample can be attributed to the characteristics of the original biomass. Miscanthus, classified as grass, is known to exhibit high amounts of silicon both through natural uptake and from residual soil or mineral debris during harvesting [98,99]. These sources may contribute to the elevated levels observed and can persist through

the lignin isolation process. The content of the other compounds analyzed are consistent with values for *Miscanthus* biomass [53,98].

Table 4. Inorganic components determined in Miscancell lignin.

Element		Miscancell (ppm)
Aluminum	Al	886
Calcium	Ca	6220
Iron	Fe	851
Potassium	K	6070
Magnesium	Mg	526
Sodium	Na	326
Phosphorus	P	732
Sulphur	S	1430
Silicon	Si	51700
Titanium	Ti	51

The thermal decomposition profile of both feedstocks is depicted in the thermogravimetric analysis (TGA) and differential thermogravimetric analysis (DTG) curves shown in Figure 9

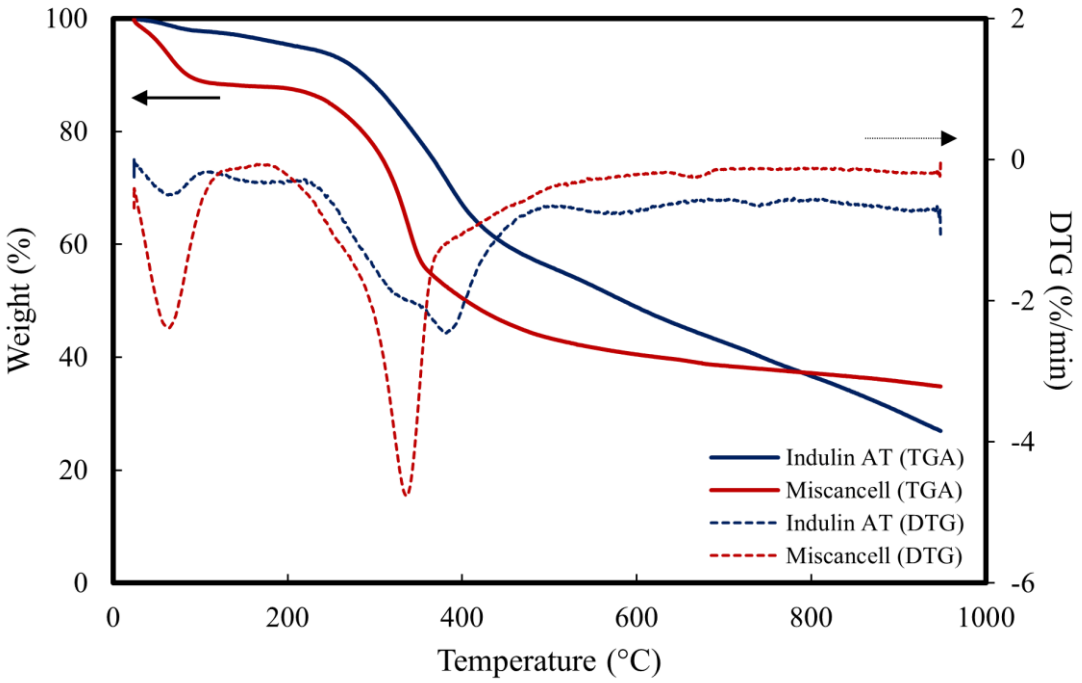


Figure 9. TGA and DTG curves of Indulin AT and Miscancell lignin.

The initial weight loss of 2.4% for the Indulin AT lignin and 12% for Miscancell lignin at temperatures up to 120°C in Figure 9 is related to the evaporation of physically adsorbed water molecules (moisture), the numbers are consistent with their moisture content shown in Table 3. The steep weight loss of 39% for Indulin AT and 46% for Miscancell in the range of 120 °C to 500 °C is attributed to the decomposition of the lignin. The Indulin AT curve resembles a typical lignin decomposition curve, with a wide temperature range and a small decomposition rate, whereas the Miscancell curve shows a high decomposition rate in a rather small temperature interval. The narrow degradation temperature range of Miscancell lignin could be related to its lower molecular weight compared to the Indulin AT [100]. Additional factors, such as its degree of condensation, interlinkages within the structure, and higher inorganic content – which can potentially act as catalysts – may also influence its decomposition behavior [100–102]. Moreover, the small shoulder around the temperature of 250 °C on the Miscancell DTG curve could indicate the presence of hemicellulose impurities in its composition. The residual char produced from the Indulin AT (27%) and Miscancell (35%) were consistent with typical lignin values [103]. The higher residual mass of Miscancell lignin can be correlated with the higher content of inorganic materials, in accordance with the higher ash content.

The structural properties of the lignins were determined via FTIR. In the spectra shown in Figure 10, the characteristic functional groups were identified.

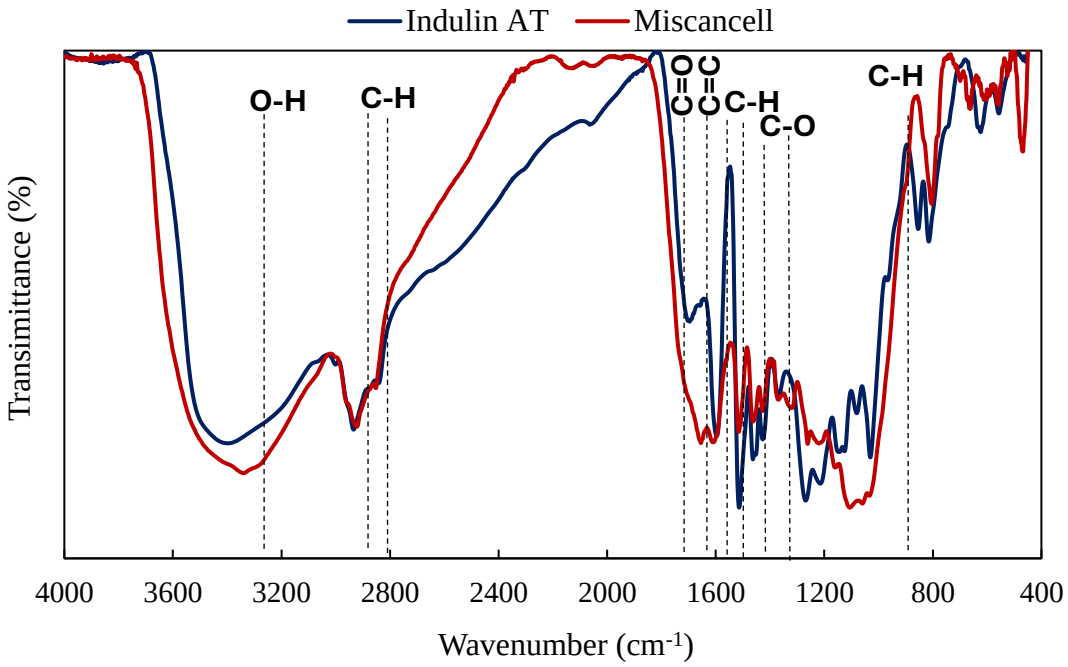


Figure 10. FTIR of Indulin AT and Miscancell.

Indulin AT lignin exhibits a strong band at 853 cm^{-1} due to deformation vibrations of C-H bonds and at 625 cm^{-1} due to sulfonic groups (C-S) remaining after the Kraft pulping process [104]. Other significant differences between the two lignins are observed in the fingerprint region (800-1700 cm^{-1}). Miscancell lignin shows stronger bands at 1320 and 1150 cm^{-1}, attributed to syringyl groups, related to its grass nature; In contrast, Indulin AT lignin displays higher intensity bands at 1270, 1030, and 850 cm^{-1}, attributed to guaiacyl units, which are more prominent due to its softwood origin [105].

2.3.2 Micro-pyrolysis experiments

Micro-scale pyrolysis experiments were performed using the micro pyrolyzer system with on-line GC-MS to investigate the decomposition profiles of lignin and to validate the chemical composition of the bio-oils from the pilot scale runs. The vapor residence time in Py-GC/MS systems is usually much shorter than in larger scale reactors, thus minimizing the secondary reactions which are responsible for char/coke formation, being able to reveal the intrinsic kinetics of lignin fast pyrolysis. On the other hand, the pilot scale reactors refer to relatively longer residence times, allowing (to some extend) secondary reactions thus affecting the products yields (and composition) of liquid oil, gases and char. This combination of scales enables a more complete understanding of lignin pyrolysis, linking intrinsic decomposition behavior with process-scale product distributions. The tests were carried out at 500 °C.

As can be observed in Figure 11, pyrolysis of Indulin AT identified mainly alkoxylated phenols (OxyPH: 75.8%), alkoxylated aromatic compounds (OxyAR: 3.9%), alkylated phenols (PH: 5.6%) while low amounts of sulfur containing compounds (SUL:1.3%) were also formed due to the sulfur impurities remained from Kraft pulping. The product distribution obtained within this study is in accordance with other works focused on the pyrolysis of Kraft lignin [104]. Regarding the alkoxylated phenols, all compounds are substituted with one methoxy group, such Guaiacol, 4-vinylguaiacol, Creosol, Isoeugenol, etc. and can be attributed to the softwood nature of pine. The S/G (syringol to guaiacol) ratio of OxyPH is 0/100. For the Miscancell lignin alkoxylated phenols (OxyPH: 56.7%) and alkoxylated aromatic compounds (OxyAR: 9.4%) were the most abundant. This lignin also produced more nitrogen containing compounds (NIT: 4.8%) compared to Indulin AT due to the higher nitrogen content confirmed via elemental analysis. On the other hand, there is the absence of sulfur containing compounds, since the Miscancell lignin is sulfur free. Furthermore, compounds such as sugars (Levoglucosan), furans (Benzofuran, 2,3-dihydro-), acids (acetic acid) are also observed and can be attributed to the degradation of sugar impurities remained from the lignin extraction

process [105]. The relative abundance of these compounds accounts for approximately 16%, suggesting that the carbohydrate content of Miscancell lignin is relatively low. The presence of carbohydrate impurities in Miscancell lignin is also supported by the elemental analysis (Table 3). The weight ratio of C/H is 11.3 for Indulin lignin (similar also to other typical technical lignins) while the ratio for Miscancell lignin is 8.3. The lower C/H ratio indicates the presence of oxygen-rich, non-lignin constituents such as carbohydrates. Regarding the alkoxylated phenols, the five most abundant compounds, shown in Table 5, are substituted with one (guaiacol, 4-vinylguaiacol, Isoeugenol, etc.) and two methoxy groups (syringol), with respect to the grass nature of *Miscanthus*. These results are in line with other studies on Miscanthus-derived lignins, which also report a predominance of guaiacyl and syringyl compounds as major pyrolysis products [106,107]. The S/G ratio of OxyPH is 33/67.

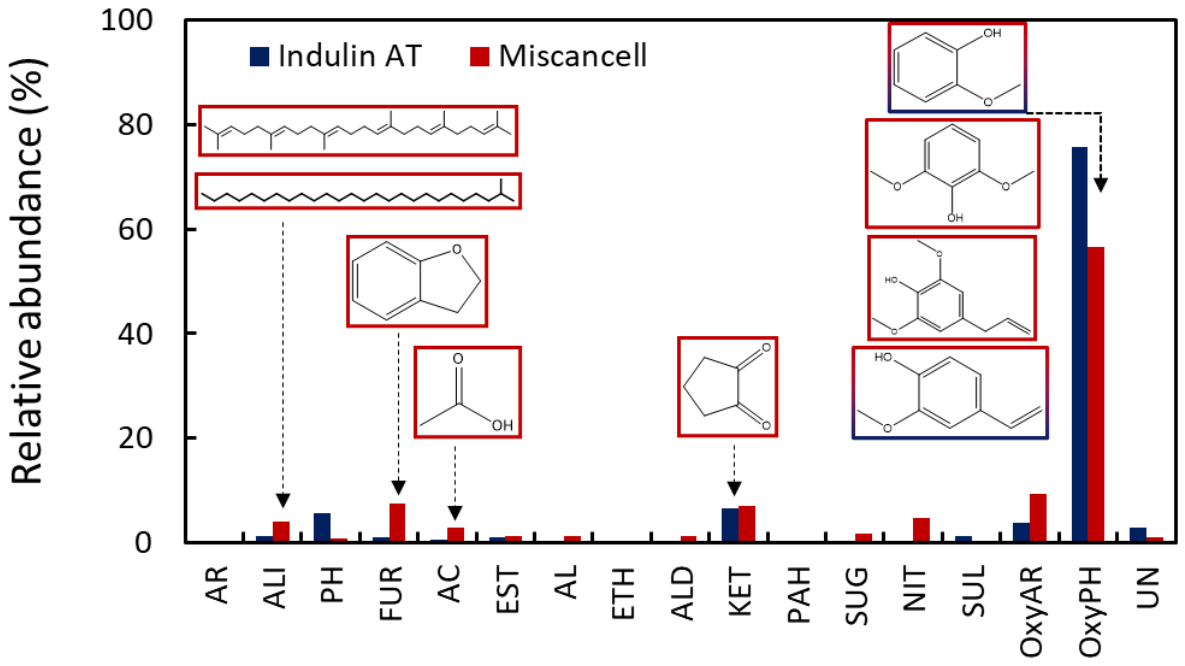
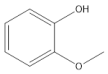
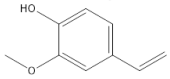
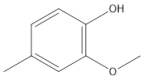
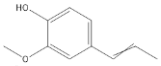
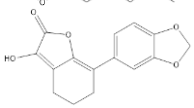
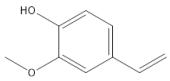
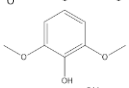
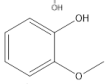
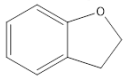
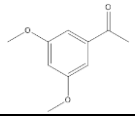


Figure 11. Relative composition of Indulin AT and Miscancell lignin derived bio-oils in the Py-GC/MS system.

Table 5. Five most abundant compounds identified in lignin's pyrolysis vapors.

Compounds	Structure	Category	Relative abundance (%)
	Indulin AT		
Guaiacol		OxyPH	22.9
4-vinylguaiacol		OxyPH	10.9
Creosol		OxyPH	10.5
trans-Isoeugenol		OxyPH	6.7
7-(3,4-Methylenedioxy)-tetrahydrobenzofuranone		KET	6.6
	Miscancell		
4-vinylguaiacol		OxyPH	15.6
Syringol		OxyPH	9.3
Guaiacol		OxyPH	9.2
Benzofuran, 2,3-dihydro-		FUR	7.5
3',5'-Dimethoxyacetophenone		OxyAR	6.6

2.3.3 Fast pyrolysis pilot scale experiments

With the changes to the order of biomass and heat carrier inlets of the reactor completed, the pilot-scale experiments with the new configuration started using the Indulin AT lignin. The modifications successfully reduced blockage inside the reactor and plugging issues were no longer observed in the reactor (see Figure 12 A and B). However, agglomerates started to build up at the inlet of the biomass feeding drop pipe, eventually leading to blockage of biomass supply into the reactor (see Figure 12 C). This is likely due to hot lignin particles being spun off the rotating shaft/heat carrier bed and hitting the relatively cold drop pipe walls. These melted or partially pyrolyzed particles stick to the surface of the pipe, leading to a gradual buildup. The blockage required mechanical breakup, indicating it had solidified and was no longer the original free-flowing material. Since the experiments with the Indulin AT only ran for a short amount of time due to the blockages, steady state and mass balances were not achieved; still some bio-oil samples were recovered for analysis.

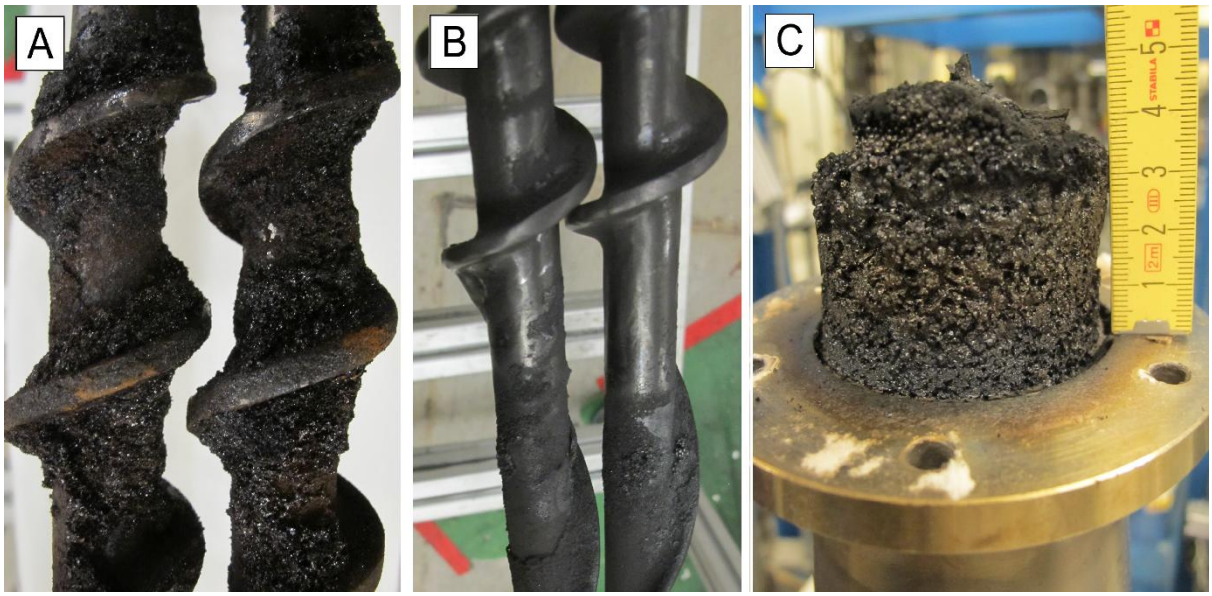


Figure 12. A) Melted lignin on screw shaft from experiment with Indulin AT before modifications on reactor's feeding section; B) Clean screw shaft and C) Built-up in reactor inlet from the experiment with Indulin AT after modifications.

In contrast to Indulin AT, no issues were observed when processing the Miscancell lignin. Visual inspection post-run confirmed the absence of buildup material on the shaft or drop pipe surfaces, as shown in Figure 13 A and B respectively. The differences in processing the two lignins could be attributed to their distinct chemical nature as well as their physicochemical properties, such as molecular weight distribution. Higher molecular weights, which is the case for Indulin AT, have been correlated with faster polymerization and higher char yields, which can contribute to agglomeration and processing challenges [108]. On the other hand, Miscancell lignin, with its lower molecular weight and different structural characteristics, processed smoothly with no agglomeration issues, highlighting how specific structural features and properties can significantly influence pyrolysis behavior. Similarly, hydrolysis lignin has been reported to exhibit no severe melting problems [44,100], further supporting the role of the chemical properties in processing outcomes. These observations show how the lignin structure and properties are important for successful processing, as well as the need for further optimization of feeding systems for efficient pyrolysis of different lignin types.

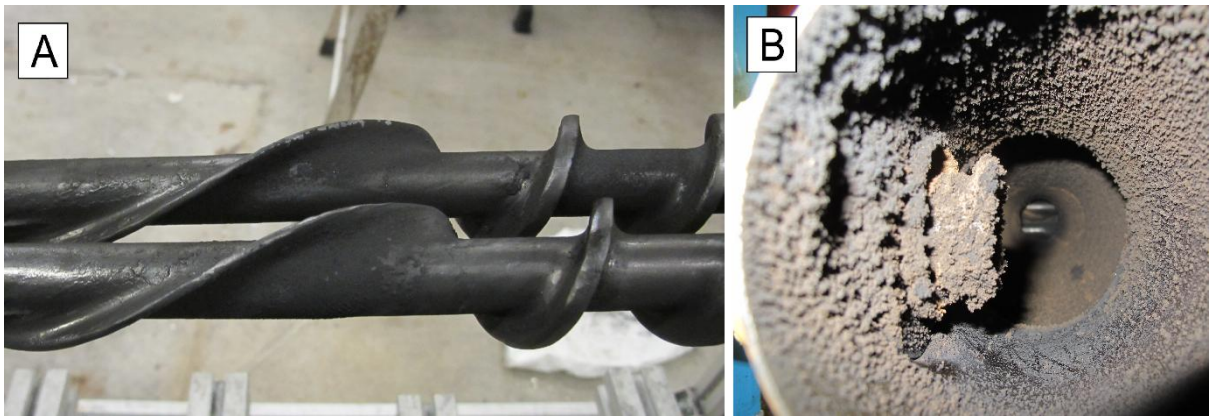


Figure 13. A) Clean screw shaft and B) Interior of biomass drop pipe (showing no blockage) after the experiment with Miscancell lignin.

Two batches of the Miscancell lignin were processed in the *Python* unit. The overall experiments were rated as successful, and the mass balances of the runs are summarized in Figure 14. The ORC and AC yields of the Miscancell lignin experiments (36.5-39.8 wt.% and 7.6-10.7 wt.%) seem reasonably high; however, they are skewed by a surprisingly high solids content that passed the two serial cyclones as well as high water content. Organic liquid yields are 22.9-29.2 wt.% and thus significantly lower that reported by Pienihakkinen et al. [44] (around 40 wt. %). It should be noted that the lignin used in the present study contained a high amount of potassium, which is known to catalyze pyrolysis reactions and consequently leads to lower condensate yields [109]. The larger deviations in char and gas yields can be due to an underestimation of gas yield in the Miscancell batch #1 run. However, higher gas and lower solid yields are not typical for such process. One reason for the low char yield could be in char agglomeration in the heat carrier cycle, leading to unaccounted for char not reaching the cyclone systems. This is typically not observed in this unit but might be promoted to liquefied lignin sticking to heat carrier particles and subsequent charring. In larger scale facilities this would not pose a problem since the char is partially burnt and steady state conditions are achieved over a prolonged period [52]. Short term agglomeration of solids in the heat carrier cycle significantly affects the mass balance for the comparably short duration of these pilot runs (3-4 hours).

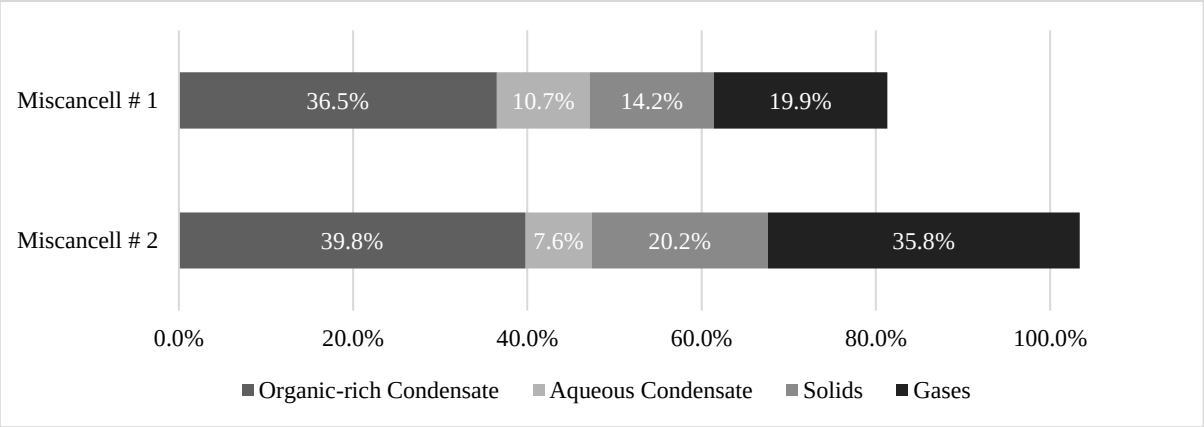


Figure 14. Mass balances of pilot runs with the Miscancell lignin in wt. %.

Carbon balance results support the previous observation that conversion to organic liquids was comparably inefficient (see Table 6). Since yields are in a similar range as other feedstocks with high ash content [50], it is likely that the high lignin ash content is causing these low organic liquid yields. Moreover, it is very likely that there was a problem in gas metering during the first Miscancell run.

Table 6. Carbon balances of lignin fast pyrolysis experiments (wt.% of feedstock carbon).

	Char	ORC[1]	AC	Gases
Miscancell #1	24.3	24.7	4.1	15.6
Miscancell #2	30.8	28.7	4.2	28.7

[1] Excluding solids

The energy distribution calculated is summarized in Table 7. For both Miscancell lignin runs, the energy recovery ranged from approximately 60% to 75% of the initial energy content of the feedstock. The ORC represented the largest energy carrier (around 40% of the energy recovery), followed by char and gases. Run 2 showed an improved energy recovery, which is consistent with its more balanced mass and carbon results. The lower gas energy observed in run 1 is in agreement with a possible underestimation of the gas phase during that experiment.

Table 7. Energy balances of lignin fast pyrolysis experiments (MJ/kg of dry feedstock).

	ORC	Char	Gases	Energy Recovery (%)
Miscancell #1	9.5	2.9	2.1	60
Miscancell #2	10.3	4.4	3.9	75

2.3.4 Product characterization

2.3.4.1 Pyrolysis Condensates

Properties of bio-oils/ORCs from Indulin AT and Miscancell lignin can be seen in Table 8. The Indulin AT bio-oil has a very low water content of only 6.7 wt.%, whereas the Miscancell oil has a higher water content of 12.5 wt.%. The higher water content of the Miscancell lignin ORC may be attributed to the higher moisture content in the feedstock, as well as the decomposition of sugar impurities, which yields more water in its decomposition compared to lignin [105]. In addition, the condensation temperature profiles could also have influenced the results. The temperature of the first condensation stage was aimed at 90 °C for all experiments. This was verified for the two successful runs with the Miscancell lignin; however, due to the short duration of the Indulin AT experiments, the temperature profiles might have encountered more significant fluctuations. Moreover, the composition of pyrolysis vapors and vapor-liquid equilibrium interactions during condensation can impact the amount of water that is condensed [53]. The density of the two oils is quite comparable, with Indulin AT oil at 1.1521 g/ml and Miscancell oil at 1.1073 g/ml, showing no significant difference. The elemental composition of the oils falls within the ranges documented in literature for lignin bio-oils [110,111]. The carbon content of the oils (60.4–68.3 wt.%) aligns with values reported for lignin oils [110,111], and is notably higher than the carbon content typically found in bio-oils from whole biomass, which generally remains below 55 wt.% [112]. Similarly, the oxygen content (23.5–30.7 wt.%) is within the expected range for lignin pyrolysis oils [110,111] and is lower than the 35–40 wt.% commonly observed in biomass-derived oils [112]. Comparing the oils, Indulin AT oil shows a higher carbon and lower oxygen content than Miscancell oil, which correlates to the feedstock characteristics, with Indulin AT being a more purified and richer in lignin, while Miscancell lignin appears to contain sugar impurities. The nitrogen content follows a similar trend, with Miscancell oil containing more nitrogen (2.7 wt.%) than Indulin AT (1.4 wt.%), also in line with the feedstock composition. Surprisingly, the Indulin AT oil does not show the presence of sulfur, which likely ended up in the gases during the pyrolysis process [113]. The HHV values for the organic-rich condensate (ORC) of the lignin oils are in good agreement with those reported in the literature [110,111,114]. The HHV of lignin-derived bio-oils, typically around 30 MJ/kg, is notably higher compared to oils obtained from lignocellulosic biomass (usually around 17 MJ/kg [112,114]). This difference arises because lignin typically has a higher carbon content and lower oxygen content than lignocellulosic biomass, contributing to a greater energy density.

Table 8. Properties of organic-rich condensates (ORC).

Property	Indulin AT	Miscancell
Water content (wt. %)	6.7	12.5
Density at 60 °C (g/ml)	1.1521	1.1073
C[1] (wt. %)	68.3	60.4
H (wt. %)	6.9	6.1
N (wt. %)	1.4	2.7
S (wt. %)	0.0	0.0
O (wt. %)	23.5	30.7
HHV ($MJ \cdot kg^{-1}$)	29.9	25.6

[1] Elemental analysis data given on dry basis.

Table 9 displays the main functional groups of compounds identified and quantified through GC-MS/FID in the pyrolysis condensates. Results are reported as approximate weight percentages. A detailed list of all the components and amounts identified in the analysis is available in the supporting information (Table S1 – Table S4). Regarding the composition of the ORCs, the Indulin AT oil predominantly consist of phenols (6.2 ± 0.4 wt. %) and guaiacols (11.4 ± 3.0 wt. %), with guaiacol, creosol, isoeugenol, and phenol being the most prevalent compounds. There is no presence of syringol, as in the micro-pyrolysis experiments, due to the softwood nature of the biomass, and no sugar compounds in the oil, confirming the purified nature of this lignin. In contrast, the Miscancell lignin ORC shows both the presence of syringol, due to the grass nature of the lignin feedstock (that contains S, G and H units) and of sugar compounds (mainly levoglucosan), also confirming the result of the micro-pyrolysis experiments. Additionally, the Miscancell lignin ORC also exhibits the presence of acids (8.2 ± 0.1 wt. %) and furans (0.4 ± 0.0 wt. %), which are related to the degradation of sugars [105]. The total content of phenolic monomers of the Indulin AT oil is around 18 wt. % which is slightly higher than for the Miscancell oil, which is around 14 wt. %, which again correlates to the more purified nature of the Indulin AT lignin.

These profiles closely match the results obtained from the micro-pyrolysis (Py-GC/MS) experiments, where the same dominant phenolic monomers were identified. The agreement between the two techniques confirms that key structural features of the lignins were retained during fast pyrolysis and reinforces the reliability of the pilot-scale experiments in capturing the representative product distribution.

Table 9. Main compound groups identified by GC-MS in organic-rich condensate (ORC) and aqueous condensate (AC).

	Indulin AT (wt. %)		Miscancell (wt. %)	
Compound	ORC	AC	ORC	AC
Non-aromatic Compounds				
Acids	0	0.3 ± 0.0	8.2 ± 0.1	1.1 ± 0.1
Ketones	0.7 ± 0.0	1.2 ± 0.2	1.4 ± 0.1	0.7 ± 0.1
Heterocyclic Compounds				
Furans	0.1 ± 0.0	0	0.4 ± 0.0	0.1 ± 0.0
Aromatic Compounds				
Benzene	0.4 ± 0.0	0	0.6 ± 0.0	0
Lignin-derived Phenol	6.2 ± 0.4	0.5 ± 0.0	3.9 ± 0.1	0.2 ± 0.0
Guaiacols (Methoxy Phenols)	11.4 ± 3.0	0.2 ± 0.0	6.1 ± 0.5	0.2 ± 0.0
Syringols (Dimethoxy Phenols)	0	0	3.6 ± 0.5	0
Carbohydrates				
Sugars	0	0	2.1 ± 0.1	0
Other Organic Compounds				
N-compounds	1.1 ± 0.2	1.0 ± 0.0	0.2 ± 0.0	0.5 ± 0.1

It is worth noting that, besides the water content, only about 20-27 wt.% of the composition of the ORC could be identified through the GC-MS/FID analysis. This is due to the limitation of volatility of high molecular weight compounds that are often too heavy to be detected by standard GC-MS equipment [115]. This high molecular weight fraction of the bio-oil, often referred to as “pyrolytic lignin”, is composed of oligomers derived from the decomposition of lignin, which explains its significant presence in lignin-derived oils [84,116].

For the AC, the main components identified were non-aromatic organic acids for Miscancell and non-aromatic ketones for the Indulin AT. Small amounts of phenols, guaiacols and nitrogenated compounds were also encountered. These components are typically lower molecular weight compounds with lower boiling point, which are likely to condensate on the second condenser and end up on the AC, which operates near room temperature. In addition, Karl Fischer titration results show that the AC is composed of mainly water, 91.8 ± 1.0 wt. % for the AC of Indulin AT and 92.6 ± 1.0 wt. % for the AC of Miscancell.

2.3.4.2 Char

Pyrolysis chars were analyzed for their elemental composition and ash content, HHV values were calculated based on the elemental composition. Table 10 shows the values obtained for the two Miscancell batches. Analysis shows that the char is rich in ash, which was anticipated given the high ash content of the Miscanthus feedstock. This elevated ash content affects the carbon content, resulting in lower-than-expected levels and also contributes to a lower higher heating value (HHV) compared to both the bio-oil and the feedstock. The low oxygen content, below 10 wt.%, is consistent with reported values for pyrochar from other biomasses pyrolyzed in the unit [117]. The nitrogen content observed in the char was lower than in the bio-oil.

In addition to potential energetic use as a solid fuel, biochar is increasingly regarded as a valuable product for carbon removal and material applications. Lignin-derived chars, in particular, can exhibit high aromaticity and structural stability, making them promising for carbon sequestration and advanced uses [118].

Table 10. Elemental analysis and ash content of Miscancell lignin chars (dry basis).

Char	C (wt. %)	H (wt. %)	O (wt. %)	N (wt. %)	Ash (wt. %)	HHV (MJ/kg)
Miscancell #1	56.3	2.4	9.5	1.5	30.4	20.6
Miscancell #2	58.9	2.3	7.1	1.5	30.3	21.7

2.3.4.3 Pyrolysis Gases

Table 11 presents the composition of the non-condensable gases produced during the Miscancell lignin fast pyrolysis, expressed in volume percent of the total dry gas, and excluding the nitrogen supplied to create the inert environment required for the pyrolysis. The main components identified in the pyrolysis gases were carbon monoxide (CO), carbon dioxide (CO_2) and methane (CH_4), typical results for biomass and lignin fast pyrolysis [117,119]. The higher deviations that can be seen for the run of Miscancell batch #1 could be somehow related to the underestimation of gases during this run. Overall, the produced pyrolysis gases presents the primary components of syngas, making it suitable for energy generation.

Table 11. Lignin fast pyrolysis gas composition (N_2 free, dry basis).

	H_2 (vol %)	CO (vol %)	CO_2 (vol %)	CH_4 (vol %)	C_2-C_5+ (vol %)
Miscancell #1	7.3 ± 1.3	27.7 ± 2.3	45.2 ± 3.7	15.7 ± 1.6	5.7 ± 0.6
Miscancell #2	6.6 ± 0.1	29.4 ± 1.2	45.7 ± 2.6	15.2 ± 0.5	6.1 ± 0.3

2.4 Conclusion

This study investigated the use of a twin-screw reactor for the fast pyrolysis of lignin at a pilot scale, comparing the performance and resulting bio-oils of two types of lignin: Indulin AT, a Kraft lignin, and Miscancell, a *Miscanthus*-derived alkali lignin. The effectiveness of this type of reactor was evaluated based on its ability to process the lignins and the properties and composition of the bio-oils produced.

Initial tests with the original reactor setup, where the feedstock entered at the cold inlet of the screw and the heat carrier was added later, revealed significant challenges with Indulin AT, leading to severe agglomeration inside the reactor. Modifications to the reactor inlet, which reversed the order of biomass and heat carried, successfully minimized blockage inside the reactor. This confirmed the hypothesis that the mechanical forces associated with the twin-screw reactor could break up agglomerates formed during lignin pyrolysis. However, a new issue arose with agglomerates building up at the biomass feeding drop pipe during the pyrolysis of Indulin AT. In contrast, processing of the Miscancell lignin in the modified reactor setup proceeded smoothly without any plugging or significant material deposition, highlighting its potential as a suitable feedstock for fast pyrolysis. The pilot runs with Miscancell lignin yielded little organic liquids; however, the observed yields were in line with results from feedstock with a similar high ash content. The differences in processing behavior between the two lignins can be attributed to their distinct chemical nature and physicochemical properties. Analysis of the oils showed that, Indulin AT produced a phenolic-rich oil, composed of mainly phenols and guaiacols, while the Miscancell lignin generated a broader range of components, including syringols, sugars, and furans, suggesting the presence of sugar impurities. Additionally, the consistency between the oil analysis and micro-pyrolysis experiments provides validation of the pilot-scale runs and its operational reliability.

These findings suggest that the twin-screw reactor design used in the study is effective for breaking up agglomerates during lignin pyrolysis, proving its suitability for this process. Still, further investigation is necessary to optimize feeding conditions for Indulin AT lignin and other lignins with pronounced melting behavior. Ultimately, Miscancell lignin and other lignins that do not exhibit severe melting behavior stand out as promising candidates for conversion via fast pyrolysis, due to their smooth processing and lack of operational issues.

Chapter 3. Comprehensive characterization of lignin bio-oils from alkali and Kraft lignin with integrated 1H-NMR spectral modeling

This chapter is based on the following in-preparation manuscript:

Comprehensive characterization of lignin bio-oils from alkali and Kraft lignin with integrated ^{1}H-NMR spectral modeling

Ana C. C. Araujo, Myriam Rojas, Wenes R. Silva, Antigoni G. Margellou, Konstantinos S. Triantafyllidis, Alberto Wisniewski Jr., Axel Funke

3.1 Introduction

As discussed in Chapter 1, the characterization of fast pyrolysis bio-oils (FPBOs) remains challenging due to their chemical complexity and the presence of a high molecular weight fraction that is difficult to analyze. This fraction is largely composed of lignin-derived oligomers, which represent the heaviest and least understood components of FPBOs. Advanced analytical techniques, in particular ultra-high-resolution mass spectrometry (UHRMS), offer new possibilities for characterizing this fraction in greater molecular detail, as it enables the detection of thousands of individual molecular species, and provides information on elemental composition and unsaturation levels. Such detailed insights can help us understand the chemical nature of this heavy lignin-derived fraction and, ultimately, advance the valorization of FPBOs.

In view of this, this chapter aims to gain a deeper understanding of the lignin-derived oligomers present in fast pyrolysis bio-oils through a comprehensive characterization of bio-oils produced from Kraft and alkali lignin. These bio-oils were selected to enhance the contribution of the lignin-derived fraction while minimizing the influence of carbohydrate-derived compounds. A multi-technique analytical approach was employed, combining FTIR, GC-MS, 2D HSQC NMR, and UHRMS to investigate the composition of the bio-oils and its structural features. In addition, the experimental findings were complemented by modeling of the ^{1}H-NMR spectra using a surrogate mixture based on the characterization results and supported by synthetic lignin oligomers (SLOs), providing effort toward validating the structural insights obtained from the analytical data.

3.2 Materials and Methods

3.2.1 Bio-oil

The lignin bio-oil used in this study was produced via fast pyrolysis using a twin-screw reactor equipped with a fractional condensation system. The process yielded an organic-rich condensate, referred to as the bio-oil, collected at 90 °C. Two types of lignin were utilized: a Kraft lignin derived from pinewood, Indulin AT, manufactured by Ingevity, North Charleston, S.C., USA , and supplied by DKSH Switzerland Ltd.; and an alkali lignin derived from Miscanthus, Miscancell, produced and supplied by Exegi IP Management B.V., the Netherlands. Full experimental details are available in Chapter 2.

3.2.2 Fourier transform infrared spectroscopy (FTIR)

The functional groups of lignin pyrolysis bio-oils were determined via Fourier transform infrared spectrophotometer (FTIR) using a PerkinElmer spectrometer. The bio-oils were mixed with KBr (1 wt.%), ground in a mortar to obtain a homogeneous mixture and pelletized under pressure (10 kPa). The samples were scanned in the wavenumber range of 4000–400 cm^{-1}, with 2 cm^{-1} resolution and 32 scans.

3.2.3 Gas chromatography mass-spectrometry (GC-MS)

The composition of the lignin bio-oils was determined via gas chromatography-mass spectrometry (GC-MS). In addition to the GC-MS/FID analysis reported in Chapter 2, which provided quantitative data on the bio-oils main functional groups, a complementary GC-MS analysis was performed to enhance the detection of heavier compounds. The organic condensates were diluted in methanol (99%, Sigma Aldrich) at a concentration of 2 wt.% and the analysis was performed using a GC-MS, equipped with an HP-5 MS column (30 m × 0.25 mm × 0.25 μm). Identification of mass spectra peaks was based on NIST11s library.

3.2.4 2D HSQC NMR

The 2D HSQC NMR spectra were obtained using a Varian 600 MHz spectrometer (Agilent Technologies, California, CA, USA). 0.1 g of the organic condensates was dissolved in 0.5 mL of DMSO-d6 (99.8%, Deutero GmbH, Kastellaun, Germany), under stirring. The chemical shifts were referenced to the solvent signal (δH/δC:2.50/39.52 ppm). The measurements were performed using a 5 s relaxation delay, spectral widths were from 14 to −2 ppm and from 190 to −10 ppm for the ^{1}H and ^{13}C dimensions respectively, the number of scans were set to 16 and 450 increments (t1) were recorded in the ^{13}C dimension. The spectra were processed using the MestReNova software. Prior to Fourier transformation, FID (free induction decay) signals were apodized with a π/2 sine squared-bell function in both dimensions. An auto-baseline and manual phase correction were also applied to both dimensions. More details regarding the experimental protocol and spectra analysis can be found in a previous publication [120].

3.2.5 Ultra-high-resolution mass spectrometry (UHRMS)

Lignin bio-oils were analyzed by an Exactive HCD Plus system (Thermo Scientific, Bremen, Germany) equipped with an Ion Max API ionization source with HESI probe (Heated Electrospray Ionization) using a Fusion 101 syringe pump (Chemyx, Stafford, TX) with a 500 μL syringe (Thermo Scientific, NJ, USA) at a flow rate of 12 μL min^{-1}. The analysis solutions

were prepared from organic condensates by dissolving approximately 5.0 mg of the organic phase in 1.0 mL of HPLC-grade methanol. This produced an initial solution with a concentration of 5.0 mg mL^{-1}, or 5,000 ppm. From this solution, a 40 µL aliquot was taken and diluted to a final volume of 1.0 mL using methanol, resulting in an analysis solution with a final concentration of 200 ppm. The samples were analyzed in negative ionization modes, the analysis conditions employed in the HESI(−) mode were: spray voltage 4.0 kV, vaporization region heating 100 °C, capillary temperature 300 °C, sheath gas 3 au, auxiliary gas 7 au, and sweep gas 2 au. Mass spectra were acquired in the m/z range of 100 to 700 with a resolution of 140,000 FWHM at m/z 200, and accumulation of 100 micro scans. The final spectrum of the raw bio-oil was obtained by subtracting the solvent spectrum. UHRMS data processing followed previously described by Santos et al. [121], using the following assignment criteria: $C_{3–50}$, $H_{4–100}$, $N_{0–4}$, $O_{0–15}$, double bond equivalent (DBE) −0.5-20, and molecular formula attribution error upon to 3 ppm.

3.2.6 ^{1}H-NMR modeling

Modeling of the ^{1}H-NMR spectrum of the alkali lignin bio-oil, obtained from the F2 projection of the 2D HSQC experiments, was conducted in the following steps:

First, individual ^{1}H-NMR spectra were generated for a set of compounds selected to form a representative surrogate mixture of the lignin bio-oil. This mixture was based on the characterization results and was divided into two fractions: one composed of volatile GC-MS-identifiable compounds, and another representing the non-volatile high molecular mass fraction. For this second fraction the representatives were chosen so that the chemical formulas from UHRMS results match with SLO structures from the NMR database [122]. The individual spectra were predicted using ACD/Labs NMR Predictor (v2024.2.3), with DMSO-d_6 specified as the solvent to match experimental conditions.

The mixture spectra were simulated in a Python, considering code assuming linear additivity, which means each component's ^{1}H signal adds linearly with its amount, without mixture-induced distortions. Linear (see linear additive model in Eq.2.). For each compound i, on a common ppm grid x:

$$y_{\mathrm{mix}}(x) = \sum_{i=1}^{M} \alpha_i \, s_i(x) \quad (2)$$

where:

$y_{\mathrm{mix}}(x)$is the simulated mixture intensity at chemical shift x; $s_i(x)$is the reference ^{1}H NMR spectrum (intensity vs. ppm) of component i, interpolated onto the common grid; $\alpha_i$is the weighting factor proportional to the amount of component i in the mixture mass-weighted (shape matching): $\alpha_i \propto w_i^{(\mathrm{wt})}$. To apply this model, experimental and theoretical conditions (solvent and resolution) were matched, and it was assumed that there are no inter-component interactions or matrix effects that would alter the spectra.

Deviations were calculated by comparing the normalized simulated spectrum $\tilde{y}_{\mathrm{sim}}(x)$to the normalized experimental spectrum $\tilde{y}_{\mathrm{exp}}(x)$on a common ppm grid. To reduce baseline bias, errors were computed only over a mask $\mathcal{M} = \{x: \tilde{y}_{\mathrm{exp}}(x) \geq \tau\}$(threshold τset in the code; e.g., 0.02). For points $x \in \mathcal{M}$, we report:

Mean Absolute Error (MAE):

$$\mathrm{MAE} = \frac{1}{\mid \mathcal{M} \mid} \sum_{x \in \mathcal{M}} \mid \tilde{y}_{\mathrm{sim}}(x) - \tilde{y}_{\mathrm{exp}}(x) \mid \tag{3}$$

Root Mean Square Deviation (RMSD):

$$\mathrm{RMSD} = \sqrt{\frac{1}{\mid \mathcal{M} \mid} \sum_{x \in \mathcal{M}} (\tilde{y}_{\mathrm{sim}}(x) - \tilde{y}_{\mathrm{exp}}(x))^2}. \tag{4}$$

Point-by-point relative error (RE) and its average (MRE):

$$RE_i = \frac{|\tilde{y}_{\mathrm{sim}}(x) - \tilde{y}_{\mathrm{exp}}(x)|(I^{exp} - I^{sim})}{\tilde{y}_{\mathrm{exp}}(x) I^{exp}} \tag{5}$$

All metrics are optionally expressed as percentages by multiplying by 100.

Finally, during the optimization step, the volatile GC-MS identifiable fraction was held constant while the composition of the high molecular weight fraction was varied. A total of 2,000 random combinations were sampled using a Dirichlet distribution, and the mixture yielding the lowest MAE was selected. The optimized formulation was re-imported into ACD/Labs to obtain a fully predicted spectrum and its Shifts list. The predicted and experimental shifts were compared by first pairing each predicted line to the nearest experimental peak using a two-stage rule: (i) if the prediction falls within any experimental range $[\delta^{\min}, \delta^{\max}]$, match to the nearest center; otherwise (ii) match to the nearest experimental

peak if within a tolerance T(0.03 ppm). For each matched pair we compute the signed error $e_{\mathrm{ppm}} = \delta^{\mathrm{pred}} - \delta^{\mathrm{exp}}$, the absolute error $|\, e \,|_{\mathrm{ppm}}$, and the absolute relative error RE(%) $=|\, e_{\mathrm{ppm}} \,|/|\, \delta^{\mathrm{exp}} \,| \times 100$. Aggregate metrics reported are MAE (ppm), RMSD (ppm), and MRE (%), calculated over all matched peaks. A detailed description of the algorithm can be found in the supporting information (Figure S1).

3.3 Results and Discussion

3.3.1 FTIR results

The FTIR spectra of the lignin pyrolysis bio-oils, shown in Figure 15, reveal distinct structural characteristics of the two samples. The alkali lignin bio-oil exhibited a broader and more intense O-H stretching band around 3300–3400 cm^{-1}, indicating a higher abundance of hydroxyl-containing compounds such as phenols and alcohols. This broadening may also reflect the higher water content of this sample (≈12 wt.%). Additionally, it also showed a stronger absorption in the aliphatic C-H region (2800–3000 cm^{-1}), suggesting a greater contribution of aliphatic side chains. A slightly more pronounced carbonyl C=O stretch near 1680 cm^{-1} also points to a higher content of carbonyl-containing species in this bio-oil, such as acids, ketones and aldehydes. In contrast, the Kraft lignin bio-oil showed sharper and more intense bands in the aromatic ring region (1600 and 1514 cm^{-1}), indicating a more condensed aromatic structure. Furthermore, stronger absorption in the C-O stretching region (1365–1118 cm^{-1}) in Kraft suggests a higher content of guaiacyl and/or syringyl units.

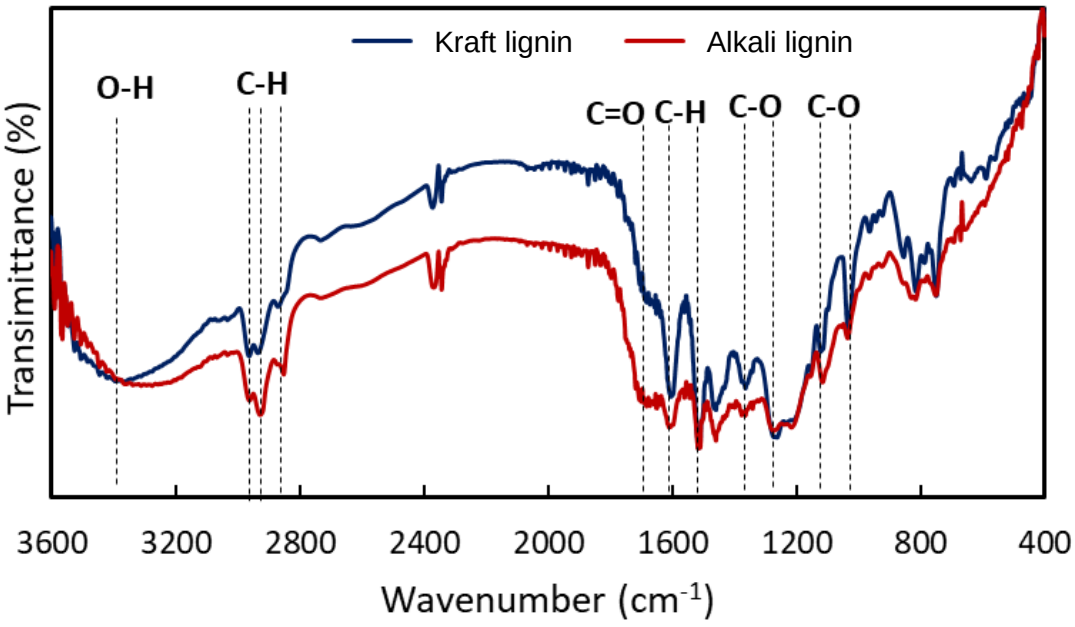


Figure 15. FTIR spectra of Kraft and alkali lignin bio-oils

3.3.2 GC-MS results

The volatile composition of lignin bio-oils was further investigated using a complementary GC-MS analysis, optimized for the detection of heavier species. This analysis enabled the identification of a broad spectrum of compounds, offering a deeper structural insight that complements the quantitative GC-MS/FID data reported previously in Chapter 2. To facilitate interpretation, the identified compounds were grouped into functional classes based on their dominant chemical features: aromatics (AR), aliphatics (ALI), phenols (PH), acids (AC), esters (EST), alcohols (AL), ethers (ETH), aldehydes (ALD), ketones (KET), polycyclic aromatic hydrocarbons (PAH), sugars (SUG), nitrogen-containing compounds (NIT), sulfur-containing compounds (SUL), oxygenated aromatics (OxyAR), oxygenated phenols (OxyPH), and unidentified compounds (UN). The relative distribution of these compound classes is shown in Figure 16. The main identified compounds, with > 2 % peak area), as well as identified oligomers, are listed in Table 12. A more detailed list of identified compounds is provided in Table S5 and Table S6 in the supporting information.

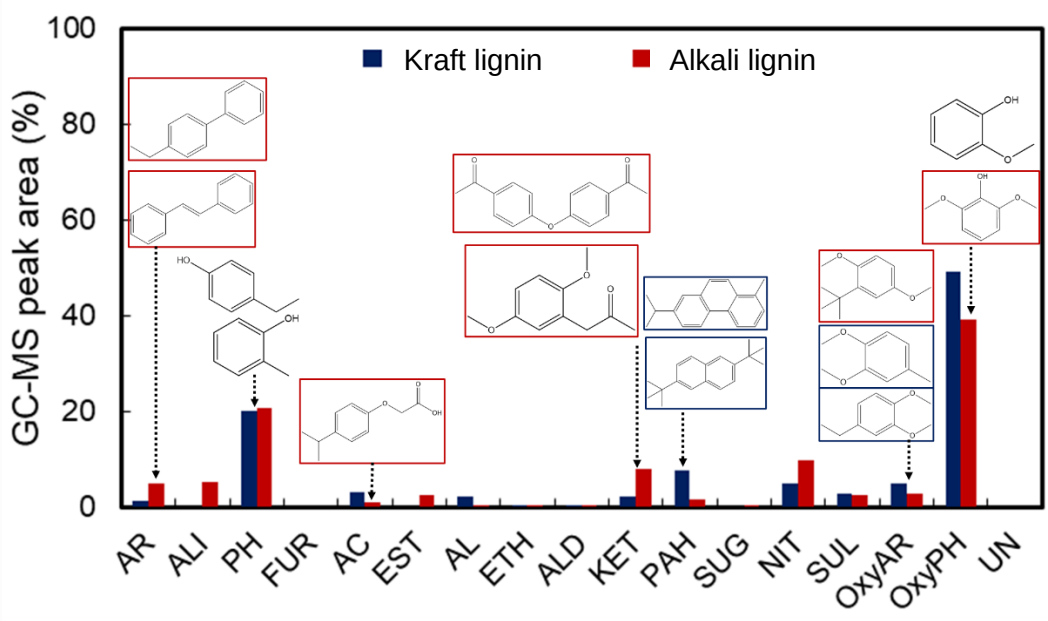


Figure 16. Composition of Kraft and alkali lignin bio-oils. Compounds in blue frame correspond to Kraft lignin bio-oil; while in red frame correspond to alkali lignin bio-oil. Compounds without frame were identified in both bio-oils.

As shown in Figure 16, both bio-oils contain high amounts of OxyPH (40-50%, relative abundance). The Kraft lignin bio-oil exhibited a higher OxyPH content compared to the alkali lignin bio-oil. PH were the second most prominent group (around 20%). Notably, guaiacol was identified as the dominant compound in both bio-oils. While, syringol was absent in the Kraft

lignin bio-oil, which is related to the softwood origin of the biomass, it was present in the alkali lignin bio-oil, reflecting the grass nature of its original biomass.

Beyond these major components, the Kraft lignin bio-oil displayed significant levels of PAHs (7.7%) and OxyAR (5.1%). PAHs higher levels in Kraft lignin bio-oil can be correlated to the high-condensed aromatic nature of Kraft lignin [123]. In contrast, the alkali lignin bio-oil contained higher amounts of AR (5.0%), ALI (5.42%), KET (7.9%), and NIT (9.9%). The higher abundance of ALI and KET suggests a contribution from residual (hemi)cellulosic impurities in the alkali lignin feedstock. Interestingly, small oligomers were detected within these groups in both bio-oils, with examples of these structures highlighted in Figure 16. Among the PAH, compounds containing up two and three aromatic rings were detected. In the AR group, the biphenyl and (E)-stilbene structures were identified. Moreover, a diphenyl ether structure was observed in the KET group.

Table 12. GC–MS identified compounds with > 2 % peak area and all detected oligomers in alkali and Kraft lignin bio-oils. Normalized peak areas are given in percent.

Component	Ret. Time (min)	Group	Alkali lignin (%)	Kraft lignin (%)
Phenol	15.96	PH	2.24	2.68
2-Methylphenol	20.55	PH	1.60	3.24
Guaiacol	21.81	OxyPH	7.93	15.9
4-Ethylphenol	28.24	PH	6.12	0.73
Creosol	28.88	OxyPH	2.64	8.56
Naphthalene, 2,6-bis(1,1-dimethylethyl)-	33.81	PAH	-	3.39
4-Ethylguaiacol	34.65	OxyPH	3.98	6.25
4-Vinylguaiacol	37.08	OxyPH	2.19	2.81
4-Acetylphenyl ether	38.79	KET	1.54	-
Syringol	39.54	OxyPH	3.72	-
trans-Isoeugenol	42.96	OxyPH	6.39	5.27
4-Ethylbiphenyl	50.19	AR	1.88	-
(E)-Silbene	57.55	AR	0.40	-
Retene	69.55	PAH	-	1.62

3.3.3 2D HSQC NMR results

The bio-oils were analyzed using 2D HSQC NMR to investigate their structure and composition, with the spectra displayed in Figure 17. This technique correlates hydrogen and directly bonded carbon atoms, mapping chemical shifts in a two-dimensional spectrum that reveals molecular structures. Widely used in the analysis of complex polymers like lignin, 2D HSQC NMR offers high sensitivity and resolution, enabling the identification of specific functional groups and structural features.

In the middle of each spectrum (δ_H/δ_C=3.0-4.0/55.0-60.0), an intense signal corresponding to the methoxy group of substituted oxygenated phenolic compounds was observed, consistent with the GC-MS analysis. In the same region, cross peaks attributed to aliphatic C-H in Hibbert ketone structure were detected at δ_H/δ_C=3.6-4.2/67.5 and δ_H/δ_C=3.6-4.2/45.0. Additionally, in the bio-oil derived from alkali lignin, signals indicative of benzoic acid derivatives were identified.

In the aromatic region of the spectra, both bio-oils exhibited multiple signals, reflecting their distinct chemical structures. The alkali lignin bio-oil showed signals for guaiacyl (G2, G5, G'6) and syringyl (S2,6, S'2,6) compounds, consistent with its grass origin, while the Kraft bio-oil displayed only guaiacyl (G2, G5) signals, aligning with its softwood origin. Cross peaks for p-coumaryl alcohol motifs (δH/δC = 7.3/128.3), corresponding to alkyl-substituted phenols, were also observed. The S/G/H ratios further confirmed these differences, calculated as 24.6/42.1/33.4 for alkali lignin bio-oil and 0.0/74.3/25.7 for Kraft lignin bio-oil. Additionally, within the aromatic region, signals corresponding to stilbene-type structures were identified, further supporting the findings from GC-MS analysis.

Regarding the interunit linkages, the alkali lignin bio-oil exhibits signals corresponding to β-O-4 aryl ether bonds (δ_H/δ_C=4.88/71.63) and 4-O-5 (δ_H/δ_C=6.4/110), the latter identified in a dimer structure in the GC-MS analysis (4-Acetylphenyl ether). No signals of β-β and β-5 carbon bonds were observed, confirming the enhanced depolymerization of lignin during pyrolysis towards small fragments [124]. The spectra of the Kraft lignin derived bio-oil show no signals assigned to interunit linkages.

In the upper right part of the spectra, in the aliphatic region, cross peaks corresponding to aliphatic compounds (alkanes, alkenes, etc.) were determined as well as to fatty acids or fatty esters impurities. These compounds remained during the extraction of lignin [125], such as hexadecenoic acid, which were also confirmed via GC-MS analysis (see Table S6).

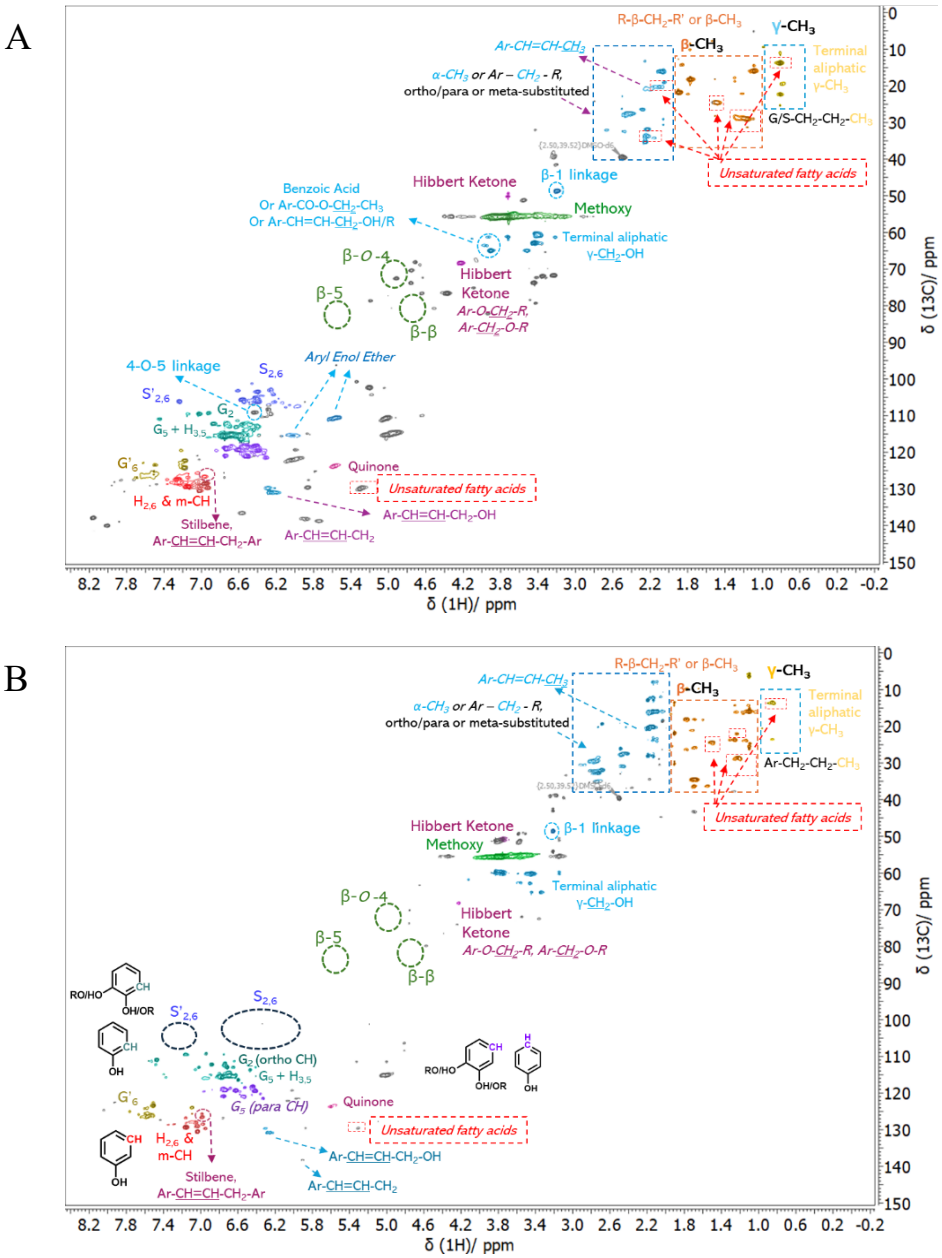


Figure 17. 2D HSQC NMR spectra of A) Alkali lignin and B) Kraft lignin bio-oils

3.3.4 UHRMS results

The FT-Orbitrap analysis was used for a comprehensive characterization of the polar and non-volatiles constituents of bio-oils. With high-resolution accuracy, FT-Orbitrap MS, provides exact mass measurements, enabling the detection of thousands of compounds in the bio-oil. This allows for the precise determination of elemental composition and key structural features, such as the degree of molecular unsaturation. This technique offers a more detailed analysis of bio-oils, effectively complementing the limited scope of GC-MS in identifying complex molecular structures and chemical classes within the bio-oil matrix.

The detected ions with assigned molecular formulas were grouped into chemical classes based on heteroatom composition. Figure 18 shows class histograms for the lignin oils, highlighting the most abundant classes: six oxygenated (O_2 – O_6) and one nitrogen-containing oxygenated class (O_3N_1). In the alkali lignin bio-oil, O_2, O_3, and O_4 dominate, with O_2 being the most abundant, while higher oxygenated classes are less prevalent, indicating a predominance of lower-oxygenated compounds. The O_3N_1 class suggests a small but still measurable nitrogen content. In Kraft lignin bio-oil, O_4 is the most prevalent class, with O_2 and O_3 present in smaller quantities compared to the alkali lignin bio-oil, and no O_3N_1 compounds detected.

Although the intensity of the signal obtained is not directly correlated with the concentration of the compound, but rather with the ionization capacity of the molecule, this effect is minimized in bio-oils, as most compounds are phenolic derivatives with similar ionization behavior. Thus, the high intensity in the O_2 class is mainly attributed to compounds with the molecular formulas $C_8H_{10}O_2$ and $C_7H_8O_2$, which corresponds to creosol and guaiacol, respectively, also identified in the GC-MS analysis and found in highest concentrations in both bio-oils. In the alkali lignin bio-oil, the O_3 class intensity surpasses O_4, which is associated with the syringols present in this bio-oil bio-oil but absent in the Kraft lignin one. Conversely, in Kraft lignin bio-oil, O_4 surpasses O_3. Classes O_5 and O_6 are present in smaller amount for both bio-oils, and O_3N_1 is exclusive to the alkali lignin bio-oil.

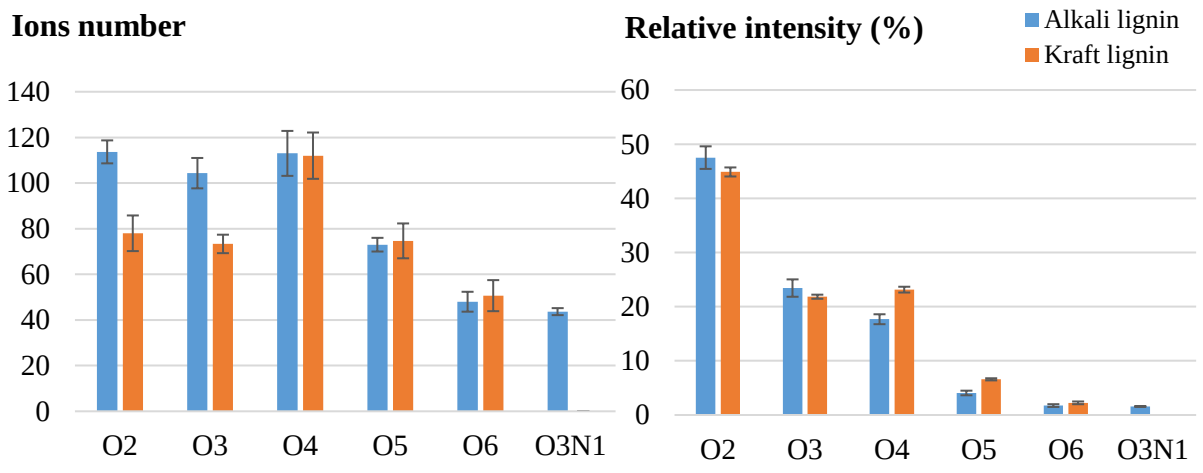


Figure 18. Class histograms of alkali and Kraft lignin bio-oil

To gain a clearer understanding of the chemical composition of the lignin bio-oils, the UHRMS data for the major classes (O_2 – O_6) were analysed via plots commonly used for molecular characterization, including DBE vs. carbon number, Van Krevelen diagrams, and NKM vs. O/C, plotted based on the intensity of signal. Resulting plots are shown in Figure 19.

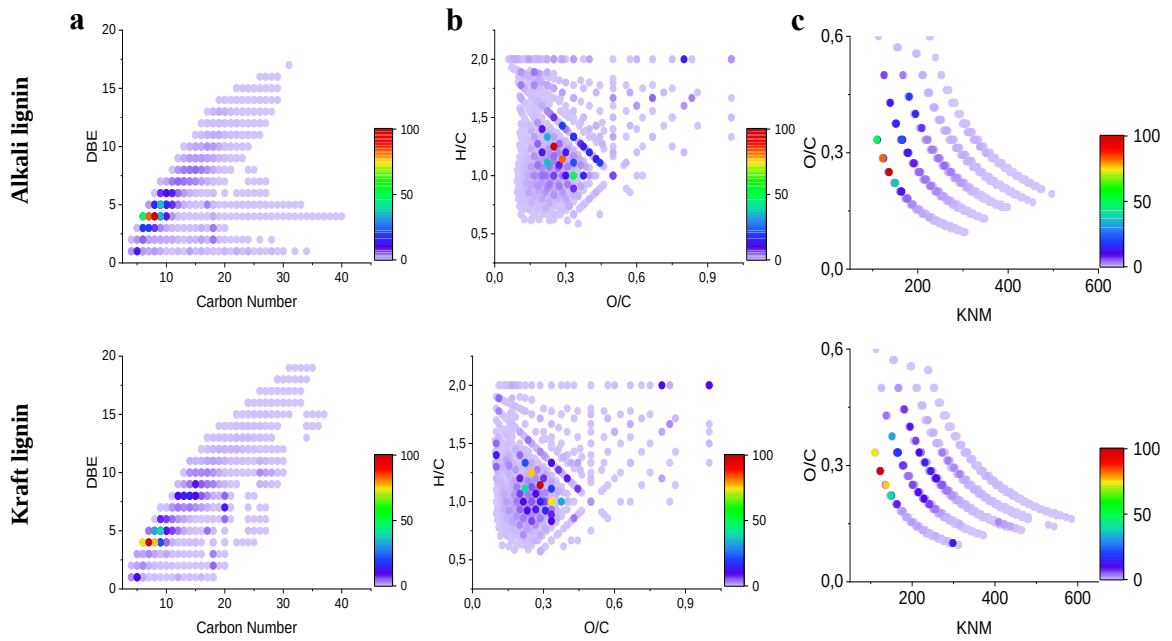


Figure 19. a) DBE vs carbon number, b) Van Krevelen diagram, c) KNM vs O/C of alkali and Kraft lignin bio-oil

The DBE vs. carbon number of the lignin bio-oils can be seen in Figure 19a. The similar trends of increase in DBE and carbon number for both oils suggest the formation of phenolic structures, ranging from monomers to oligomers. In Kraft lignin oil, the increase is more pronounced, with highly unsaturated oligomers extending up to approximately 38 carbons and a DBE of 18. For alkali lignin oil, this trend is less prominent, with oligomers reaching up to about 32 carbons and a DBE of 17, suggesting slightly smaller structures. Additionally, the alkali lignin bio-oil reveals molecules with high carbon numbers and low DBE values, consistent with large, saturated aliphatic compounds, corroborating the GC-MS findings. For both lignins, the region of highest intensity on the plots correspond to compounds with 6–10 carbons and DBE values of 4-6, suggesting a more significant presence of phenolic monomers.

Figure 19b presents the Van Krevelen diagrams of the bio-oils, which share similar features. The H/C ratios range from 0.5 to 1.5, indicate the presence of aromatic compounds alongside some aliphatic hydrocarbons, while O/C ratios range from 0.1 to 0.6, reflect oxygenated phenolic structures. The highest concentration of data points is observed at H/C values of 1.0-1.2 and O/C values of 0.2-0.4, corresponding to moderately oxygenated aromatic compounds. Notably, the Kraft lignin diagram shows a slightly higher data density, suggesting a greater abundance of complex aromatic structures.

The KNM (Kendrick mass defect) vs. O/C diagrams are shown in Figure 19c. Based on the mass range, components were classified as monomers (100-250 Da), dimers (250-450 Da), trimers (400-600 Da), and tetramers (550-750 Da), following ranges established in previous

UHRMS fragmentation studies [126,127]. The monomer range, as in other graphs, is the region with the highest intensity. The monomer and dimer ranges display the widest variety of compounds, with oxygen numbers spanning from O_2 to O_6. The trimer region is narrower, containing only compounds from O_4 to O_6. No tetramers were observed in the alkali lignin oil for these classes, while a small number of O_6 molecules were identified in the Kraft lignin oil.

3.3.5 ^{1}H-NMR modeling

Following the comprehensive characterization of the lignin bio-oils, modeling of the ^{1}H-NMR spectrum of the alkali lignin bio-oil (obtained from the F2 projection of the 2D HSQC NMR) was performed with a focus on the contribution of the heavy fraction. A surrogate mixture was used to represent the bio-oil and this mixture was divided into two parts: (i) the volatile pool identified by GC-MS/FID in Chapter 1, whose concentrations were kept fixed according to the values summarized in Table 13; and (ii) the residual heavy fraction.

Table 13. Volatile composition of alkali lignin bio-oil

Compound	Concentration (wt.%)
Acetic Acid	3.31
Propionic acid	4.97
Cyclopentenone	0.56
Cyclopentenone, 2-methyl	1.32
2-Furanone	0.42
Benzofuran	0.59
Phenol	1.32
P-cresol	1.32
Phenol, 4-ethyl	1.31
Guaiacol	1.84
Guaiacol, 4-vinyl	1.23
Isoeugenol	1.84
Acetoguaiacone	1.23
Syringol	1.64
Syringol, 4-vinyl	1.64
Acetosyringone	0.37
Levoglucosan	2.11
Water	12.25

To represent the heavy fraction, surrogate molecules were selected from a database of synthetic lignin oligomers (SLO) with experimentally determined ^{1}H-NMR spectra [122]. The

selection was guided by the results of the multiple analyses performed in the characterization, which showed that the heavy fraction was primarily composed of dimers and trimers containing two to six oxygen atoms, featuring common lignin inter-unit linkages such as β-O-4 and 4-O-5, included structures like biphenyls and (E)-stilbenes, and exhibited a high abundance of hydroxyl, methoxy, and carbonyl functional groups. Taken together, 13 surrogate structures were selected, with their chemical formulas and DBE validated against the UHRMS data. Details of all tested surrogates are provided in Table S7 in the supporting information. In addition, due to database limitations, compounds with oxygen levels up to O_9 were also included, and one compound not validated by the UHRMS results was considered.

Figure 20 shows the comparison between the experimental ^{1}H-NMR spectrum and the simulations obtained using only the volatile pool (Figure 20A) and the full surrogate mixture including the heavy fraction (Figure 20B).

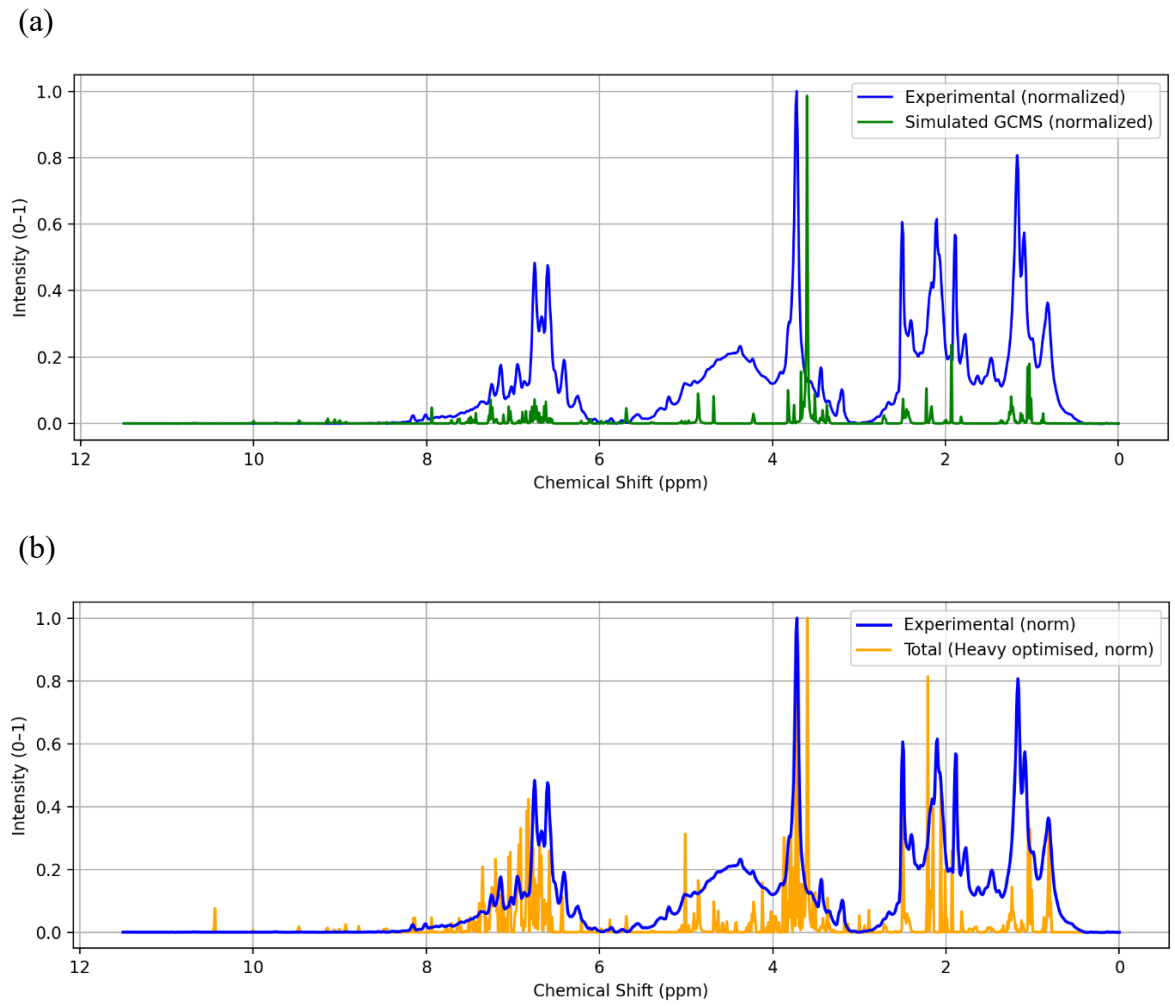


Figure 20. Experimental and simulated ^{1}H-NMR spectra of alkali lignin bio-oil for (a) mixture of volatiles and (b) mixture of volatiles + heavy fraction.

When considering only the volatile pool, the simulated spectrum (Figure 20A) showed extremely poor agreement with the experimental data, with a MAE > 1000%.% (for the spectra comparison). This significant mismatch confirms that a major contribution to the ^{1}H-NMR signal arises from the GC-MS-invisible heavy fraction. Introducing an initial, untuned set of lignin-derived oligomer surrogates significantly reduced the error; however, the MAE remained above 200%, indicating that although the general baseline of the spectrum was captured, the relative proportions of the oligomers were not yet well represented. After optimization of the oligomer composition (Figure 20B), the model converged to MAE = 14.87% and RMSD = 20.23%, demonstrating that accounting for and adjusting the heavy fraction is crucial for reproducing the experimental spectrum of the bio-oil.

The remaining deviations arise mainly from the water resonance (δ 3.6-3.72 ppm), which is highly sensitive to solvent exchange, and from subtle matrix effects that affect line shape and chemical shift in complex mixtures. Additional deviations observed in the δ 4–6 ppm region may arise from residual carbohydrate-derived species, which were outside the scope of the surrogate modeling aside from levoglucosan. The concentrations of the optimized surrogate molecules are listed in Table 14. Three of the 13 tested surrogates yielded concentrations close to zero and therefore were not included in the table. The structures of the optimized SLOs are shown in Figure 21.

Table 14. Optimized representatives for the heavy fraction

Synthetic Lignin Oligomer (SLO)	Chemical Formula	DBE	Optimized conc. (wt.%)	Intensity of signal in UHRMS
SLO-1	$C_{14}H_{14}O_3$	8	2.749	1.80e+07
SLO-2	$C_{16}H_{16}O_4$	9	11.07	1.17e+07
SLO-3	$C_{16}H_{18}O_4$	8	3.785	1.70e+07
SLO-4	$C_{16}H_{18}O_5$	8	9.251	5.17e+06
SLO-5	$C_{27}H_{28}O_5$	14	7.251	1.16e+05
SLO-6	$C_{18}H_{18}O_6$	10	1.356	1.33e+06
SLO-7	$C_{18}H_{18}O_6$	10	5.534	1.33e+06
SLO-8	$C_{22}H_{26}O_8$	10	4.607	1.47e+05
SLO-9	$C_{21}H_{28}O_8$	8	6.644	7.04e+04
SLO-10	$C_{30}H_{38}O_9$	12	7.730	-

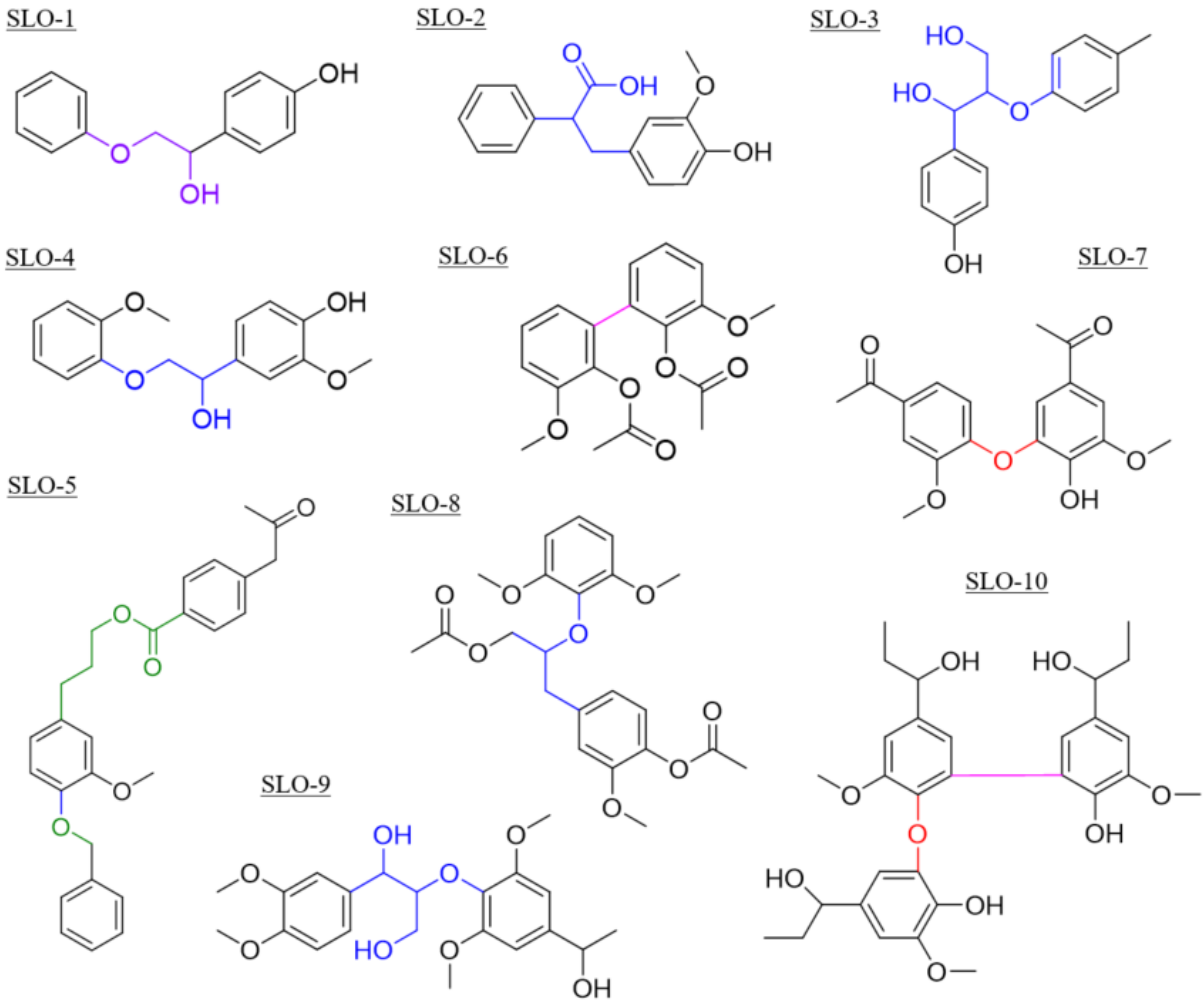


Figure 21. Synthetic lignin oligomers (SLO)

Validation of the optimized surrogate mixture against the experimental ^{1}H-NMR shifts found in in the spectrum showed excellent agreement (Figure 22). Using the optimized composition, 227 resonances were compared, yielding a MAE of 0.029 ppm (0.663%) and RMSD of 0.041 ppm (1.023%), values that fall well below the ~5% threshold typically considered acceptable for chemical-shift reproduction in complex mixtures. This outcome demonstrates that the optimized heavy-fraction surrogates, combined with the GC-MS-identified volatiles, provide a chemically consistent description of the bio-oil sample. Remaining deviations again are minor and arise mainly from solvent effects, subtle matrix-induced shifts, or structural motifs not captured within the current surrogate pool. Overall, this validation confirms that incorporating experimental data from synthetic lignin oligomers into surrogate-based modeling enables a reliable theoretical reconstruction of the bio-oil composition, helping to elucidate the contribution of the GC-MS-invisible heavy fraction.

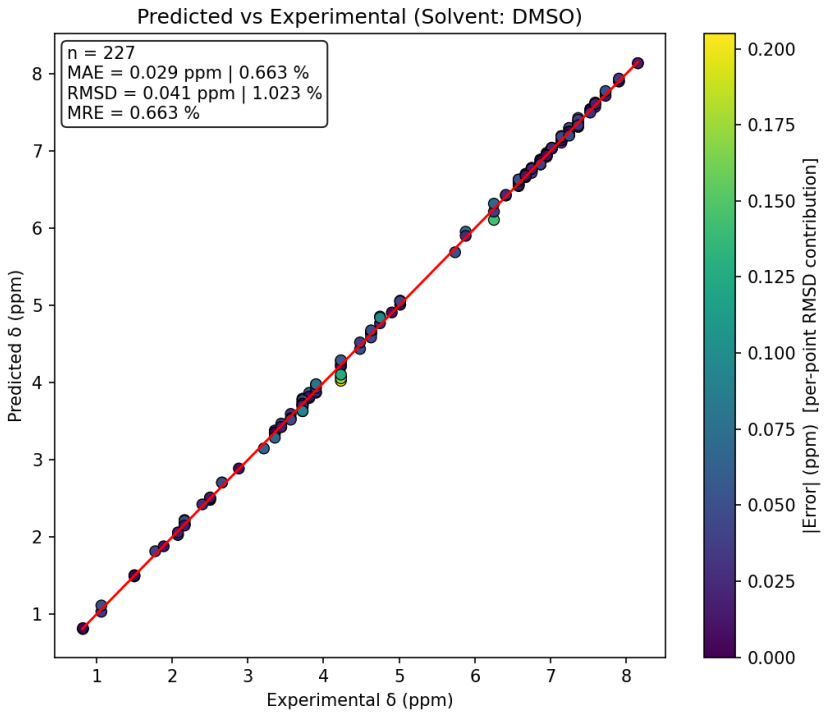


Figure 22. Experimental and simulated shifts for alkali lignin bio-oil.

3.4 Conclusion

This study employed a set of analytical techniques, including FTIR, GC-MS, 2D HSQC NMR, and UHRMS, to comprehensively characterize the composition and structural features of two lignin-derived bio-oils. These methods provided a multidimensional analysis, enabling a thorough characterization of the bio-oils and providing valuable insights into their oligomeric structures. Building on these results, modeling of the ^{1}H-NMR spectrum with surrogate mixtures composed of identified volatiles and synthetic lignin oligomers allowed us to reconstruct the spectra and evaluate the contribution of the heavy fraction, effectively linking experimental characterization with predictive representation.

Characterization results showed that the alkali lignin bio-oil, obtained from Miscanthus biomass, exhibited a broader chemical diversity than the Kraft lignin bio-oil. Functional group analysis revealed higher contents of hydroxyl, carbonyl, and aliphatic species, while the aromatic region of the NMR spectra confirmed the presence of guaiacyl, syringyl, and p-hydroxyphenyl units. GC-MS further identified a wide distribution of compound classes, including phenols, ketones, aliphatics, and nitrogenated species, together with small oligomers such as biphenyls, stilbenes, and 4-O-5 ethers. UHRMS corroborated these findings, showing a predominance of oxygenated dimers and trimers with oxygen numbers up to O_6, reflecting the oxygenated character of this bio-oil. In contrast, the Kraft lignin bio-oil was more enriched

in condensed aromatics and oxygenated phenols. GC-MS revealed the presence of polycyclic aromatic hydrocarbons (PAH), while UHRMS indicated more highly unsaturated oligomers and detected oligomer species up to tetramers, highlighting the higher degree of structural condensation in the Kraft lignin bio-oil compared to the alkali sample.

Building on these insights, modeling of the ^{1}H-NMR spectrum of the alkali lignin bio-oil demonstrated the essential role of the heavy fraction (i.e. dimers and trimers) in reproducing the experimental spectrum. Simulations based only on the volatile pool captured the spectral features poorly, and the introduction of synthetic lignin oligomers as surrogates to represent the heavy fraction substantially improved the agreement. Optimization of the surrogate composition resulted in a close match with the experimental spectrum (MAE = 15.64%, RMSD = 20.95%), while validation against experimental chemical shifts further confirmed the robustness of the model (MAE = 0.029 ppm, RMSD = 0.041 ppm). It is interesting to observe that exclusively dimers and trimers contribute to this surrogate mixture and seem to be sufficient in representing the heavy compounds. These results highlight the value of combining analytical characterization with a surrogate mixture modeling approach based on synthetic lignin oligomers to bridge the gap between well-identified compounds and the unresolved heavy fraction of bio-oils.

Chapter 4. Assessing phase behavior of pyrolytic lignin in fast pyrolysis bio-oil via advanced distillation curve

This chapter is based on the following publication:

Assessing phase behavior of pyrolytic lignin in fast pyrolysis bio-oil via advanced distillation curve

Ana C. C. Araujo, Frederico G. Fonseca, Nicolaus Dahmen, Axel Funke

Biomass and Bioenergy, 2026, 209, 109009. DOI: 10.1016/j.biombioe.2026.109009

4.1 Introduction

The complex composition of fast pyrolysis bio-oils (FPBOs) includes a large number of partially identified and unresolved compounds, particularly within the high-molecular-weight fraction derived from lignin. The pyrolytic lignin (PL) fraction, in particular, strongly affects properties such as volatility, viscosity, and stability, thereby influencing both the recovery and upgrading of FPBOs. A deeper understanding of the vapor–liquid equilibrium (VLE) behavior of this fraction is therefore essential for improving the accuracy of thermodynamic models and supporting process design.

In this chapter, the representation of pyrolytic lignin in FPBO models is investigated to find the most suitable approach for improving VLE predictions. To achieve this, advanced distillation curve (ADC) experiments are integrated with VLE simulations to obtain a comprehensive understanding of FPBO's phase behavior, with a particular emphasis on the influence of pyrolytic lignin. To isolate and elucidate the thermodynamic role of the pyrolytic lignin, FPBO derived specifically from lignin was selected, allowing for a focused assessment of this high molecular weight fraction without the interference of carbohydrate derived compounds. The ADC experiments yield detailed temperature and composition profiles for each distilled fraction of FPBO, which are subsequently modeled using the IDEAL model, the UNIFAC-Dortmund (DMD) activity coefficient model, and the Peng–Robinson Boston–Mathias (PR-BM) equation of state, using different surrogate molecules to represent the pyrolytic lignin fraction.

4.2 Materials and Methods

4.2.1 Lignin fast pyrolysis bio-oil

The lignin bio-oil used in this study was produced via fast pyrolysis of a *Miscanthus* derived lignin (Miscancell) and characterized via various analysis, as detailed in the previous chapters 2 and 3. The bio-oil was also characterized for its elemental composition, density, chemical composition via GC-MS/FID, and water content via Karl Fischer titration, with key results summarized in Table 15.

Table 15. Properties of the lignin FPBO.

Property	Lignin FPBO
General properties	
Water content (wt.%)	12.5
Density at 60 °C	1.11
Elemental composition (wt.%)	
C	60.4
H	6.1
O	30.7
Chemical composition (wt.%)	
Acids	8.2
Ketones	1.2
Furans	0.4
Benzenes	0.6
Phenols	3.9
Guaiacols	6.1
Syringols	3.6
Sugars	2.1

4.2.2 Advanced distillation curve experiments

Vacuum advanced distillation curve (ADC) experiments using the lignin bio-oil were conducted following the approach reported by Krutof and Hawboldt [90], a scheme of the setup can be found in Figure 23. At the beginning of each experiment, the system was evacuated to an absolute pressure of 15 kPa, after which heating was initiated. The sample was placed in a round-bottom flask enclosed by a heating jacket programmed to follow a three-stage heating ramp from room temperature to 285 °C. Vapor condensation was achieved using a counter-flow condenser initially cooled to 0 °C.

The first measurement was recorded when the first droplet of condensate was visually observed. Subsequently, kettle and head temperature were recorded at every 10 mL interval, as observed on the calibrated receiving flask. At the same points, 10 µL samples were taken using a syringe for GC-MS/FID analysis. Distillation continued until no further condensation was observed. All experiments were carried out in triplicate to ensure reproducibility.

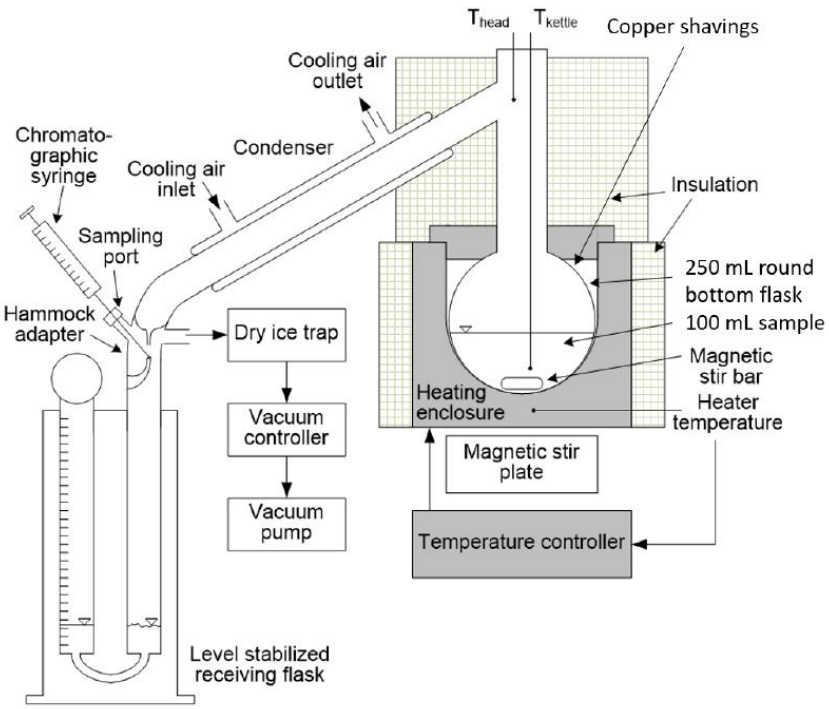


Figure 23. Scheme of ADC system adapted from [90]

A few adaptations were made to the setup described by Krutof and Hawboldt [90]. To accommodate for a smaller sample size used in the study (100 mL vs, 200 mL), a 250 mL round-bottom flask was used instead of a 500 mL one, and the gap between the flask and the heating enclosure was filled with copper shavings to improve heat transfer. In addition, the lower pressure of 5 kPa used in the original work could not be achieved due to equipment limitations.

4.2.3 Modeling

Modeling of the ADC experiments was carried out using Aspen Plus V14®. The thermodynamic property method used was either IDEAL, UNIFAC-DMD, or PR-BM for comparison purposes; property methods were not modified from their default configurations [129]. All molecules were characterized based on their structure, and in the case of the UNIFAC-DMD model, functional group information was also required. Properties that were not present in the Aspen Properties™ database were estimated using default options.

For the VLE calculations, a scheme of the flowsheet model can be found in Figure 24. The process was simulated at an absolute pressure of 15 kPa, with the lignin FPBO input as a liquid at 25 °C and a flow rate of 100 L/h. Each distillation stage was modeled as a *HEATER* + *SEP2* pair, with the *HEATER* configured to achieve a specified vapor fraction and the *SEP2* unit serving as an adiabatic separator. Design specifications were applied at each stage to evaporate

10 L of the liquid input by adjusting the vapor fraction in the *HEATER*, corresponding to the volume intervals of the ADC experiments. For the very first drop, however, no design specification was applied, and a fixed vapor fraction of 1×10^{-6} was used instead. The vapor stream produced in each stage was condensed using a second *HEATER*, enabling evaluation of the recovered condensate fractions. This approach allowed for the prediction of the temperature profile along the distillation and the component distribution across the collected fractions.

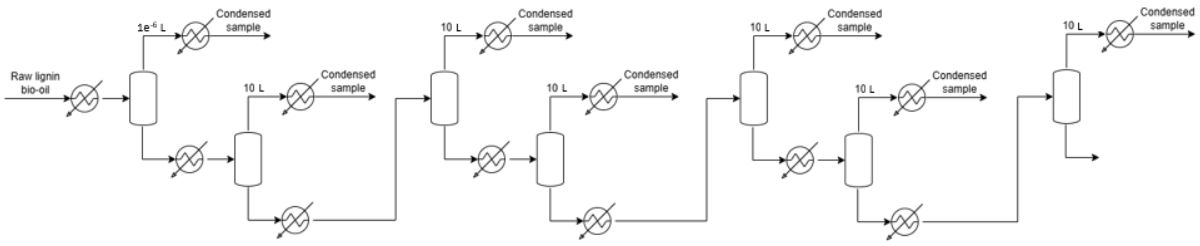


Figure 24. Scheme of the model implemented in Aspen Plus™, based on the flowsheet view of the software environment.

The lignin FPBO was modeled using a surrogate mixture, as detailed in Table 16.

Table 16. Surrogate mixture of the lignin FPBO.

Component	Mass Fraction	Formula	CAS
Acetic Acid	3.31%	$C_2H_4O_2$	64-19-7
Propionic acid	4.97%	$C_3H_6O_2$	79-09-4
Cyclopentenone	0.56%	C_5H_6O	930-30-3
Cyclopentenone, 2-methyl	1.32%	C_6H_8O	1120-73-6
2(H)-Furanone	0.42%	$C_5H_4O_2$	98-01-1
Benzofuran	0.59%	C_8H_6O	271-89-6
Phenol	1.32%	C_6H_6O	108-95-2
P-cresol	1.32%	C_7H_8O	106-44-5
Phenol, 4-ethyl	1.32%	$C_8H_{10}O$	123-07-9
Guaiacol	1.84%	$C_7H_8O_2$	90-05-1
Guaiacol, 4-vinyl	1.23%	$C_9H_{10}O_2$	7786-61-0
Isoeugenol	1.84%	$C_{10}H_{11}O_2$	5932-68-3
Acetoguaiacone	1.23%	$C_9H_{10}O_3$	498-02-2
Syringol	1.64%	$C_8H_{10}O_3$	91-10-1
Syringol, 4-vinyl	1.64%	$C_{10}H_{12}O_3$	28343-22-8
Acetosyringone	0.37%	$C_{10}H_{12}O_4$	2478-38-8
Levoglucosan	2.11%	$C_6H_{10}O_5$	498-07-7
Water	12.25%	H_2O	7732-18-5
Pyrolytic Lignin Surrogate	60.93%	-	-

The selected components and their weight fractions were based on data from the previous bio-oil characterization as described in Table 15; the remaining unidentified fraction was represented as pyrolytic lignin.

The pyrolytic lignin (PL) surrogate was modeled using various molecules proposed in the literature to evaluate which structure best represents the bio-oil properties and VLE behavior. The molecules were grouped according to shared structural characteristics, such as number of aromatic rings and linkage type, as summarized in Table 17. To illustrate these differences, Figure 25 shows a representative molecule from each group, highlighting the base structure and typical size. Full information on the structure, molecular formula, and functional groups of all PL molecules are provided in Table S8 and Table S9 the supporting information (SI).

From this point on, these molecules will be referred to as PL surrogates. In modeling, all molecules were evaluated individually, meaning no mixture of PL surrogates was considered. Additionally, a surrogate mixture excluding the PL surrogate, referred to as "No-PL," was evaluated to assess the importance of including pyrolytic lignin in the overall representation. This No-PL mixture contained only the components identified during the characterization, with their weight fractions normalized to 100%.

Table 17. Classification of PL surrogates by structure and linkage type.

Group	Description	Molecules	Reference
D-A	Dimer – β-O-4	D-A1, D-A2, D-A3, D-A4	[91,92]
D-B	Dimer – Biphenyl	D-B1, D-B2, D-B3, D-B4	[86,130]
D-C	Dimer – Stilbene	D-C1, D-C2	[89,130]
D-D	Dimer – Phenylcoumaran	D-D1, D-D2	[130]
D-E	Dimer – Resinol	D-E1	[130]
D-F	Dimer – Bridging double bond	D-F1	[131]
TR	Trimer – various links	TR1, TR2, TR3, TR4, TR5, TR6	[89,92,126]
TE	Tetramer – various links	TE1, TE2	[56,89]

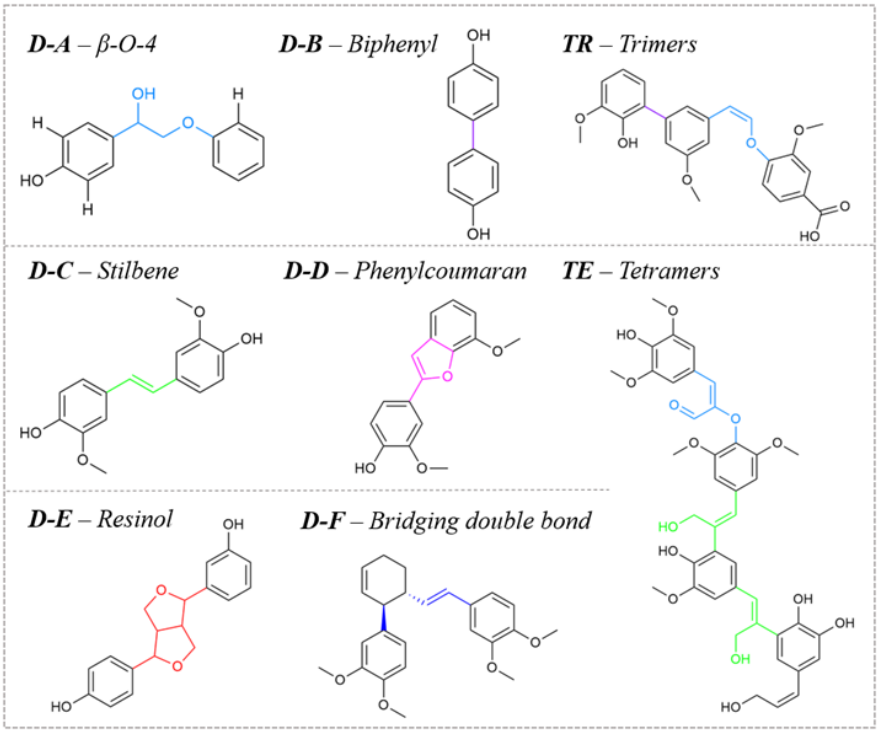


Figure 25. Representative PL surrogate from each group

For the calculation of bio-oil properties, the elemental composition was determined by performing a simple component balance over the surrogate mixtures, using the different PL surrogate molecules. Density predictions were obtained by simulating the bio-oil surrogate mixtures with each PL surrogate in Aspen at 60 °C, matching the reference temperature of the experimental data.

4.2.4 Data treatment

All simulated property values were processed and analyzed against the experimental data. Experimental reference values for elemental composition and density were taken from the characterization of the lignin FPBO reported on Table 15. For VLE evaluation, the experimental data of the ADC temperature profile and water content across the distillation range were used.

The relative error (RE) was used for properties with a single experimental value, such as density:

$$RE = \left| \frac{x_{sim} - x_{exp}}{x_{exp}} \right| \times 100 \quad (6)$$

For properties with multiple data points, such as elemental composition and ADC profiles, the mean absolute relative error (MARE) was used:

$$MARE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_{sim,i} - x_{exp,i}}{x_{exp.i}}\right| \times 100 \quad (7)$$

In addition, for evaluating the different groups of molecules, the mean MARE of each group's molecules was used, following the grouping logic defined in Table 17.

4.3 Results and Discussion

4.3.1 ADC experiments

The kettle temperature results obtained from the ADC experiments of the lignin FPBO can be seen in Table 18. At the distillation starting point (57.6 °C), temperature values were consistent across replicates, with a standard deviation of 0.5 °C. Following this, the temperature increased to an average of 96.9 °C at 10% distilled volume. An even sharper increase was observed at 20%, with temperature reaching 189.1 °C. This point also exhibited the highest deviation across replicates (22.6 °C), indicating a broader range of volatiles present in this fraction. From that point on, temperature gains became less pronounced, stabilizing toward the final measured values and showing lower deviations.

Table 18. Kettle temperatures from ADC of lignin fast pyrolysis bio-oil at 15 kPa.

Distilled volume %	Experimental kettle temperature (T_{kettle}, °C)				
	Run 1	Run 2	Run 3	Average	St. deviation
0	56.9	57.9	58	57.6	0.5
10	93.9	81.3	115.5	96.9	14.1
20	162.9	186.4	218	189.1	22.6
30	219	227	247	231	11.8
40	244	251	262	252.3	7.4

The temperature results presented in Table 18 are compared with the data reported by Krutof and Hawboldt [90], for a softwood bio-oil, in Figure 26. Given the limited availability of ADC data for bio-oils, particularly lignin-derived ones, the softwood bio-oil results are used here as a comparative reference. Comparing the curves, the two distillation profiles show significant differences. Krutof and Hawboldt [90] reported an earlier distillation onset (33.4 °C) and a convex profile with a lower incline. In contrast, this work presents a sharper incline and a more concave profile, with distillation starting at higher temperatures. Furthermore, the distillation of softwood bio-oil proceeds up to ~65 vol.% distilled, whereas in this work it was limited to

~40 vol.% Interestingly, despite these differences, both curves reach similar final temperatures near 260 °C.

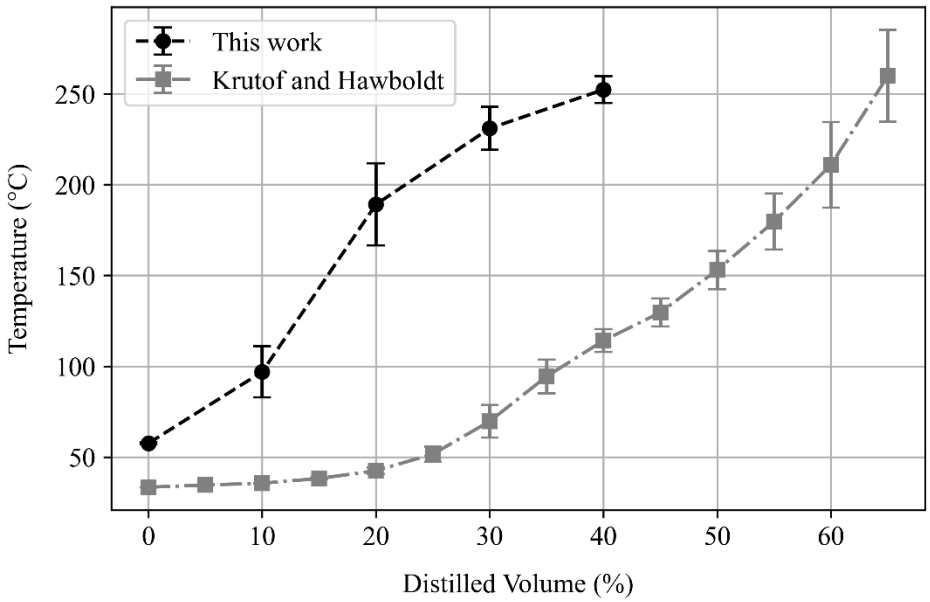


Figure 26. ADC temperature profiles of lignin FPBO at 15 kPa (this work) and softwood fast pyrolysis bio-oil at 5 kPa (Krutof and Hawboldt [90]).

These variations can be attributed to both the distinct nature of the bio-oils and the experimental conditions. The lignin-derived bio-oil used here contains a higher proportion of heavy phenolic oligomers and pyrolytic lignin, whereas softwood-derived oil likely includes more water and light volatiles. Additionally, the lower vacuum pressure employed by the reference (5 kPa vs 15 kPa in this study) would favor earlier and more extended evaporation of volatile fractions, further amplifying the observed differences.

In addition to the temperature data, the ADC experiment also provides insight into the distribution of compounds across the distilled fractions, as shown in Figure 27. Figure 27a shows how the water content varies throughout the distillation, starting as the dominant component at approximately 80 wt.% of the initial fraction and gradually decreasing as the distillation progresses. It is interesting to note that at 15 kPa, the boiling point of pure water is approximately 54.8 °C [132], closely aligning with the initial temperature observed in our experiments (57.6 °C). The slight elevation could be attributed to the presence of other solutes in the mixture. In contrast, at 5 kPa, water boils at around 32.9 °C [132], which corresponds well with the distillation onset reported by Krutof and Hawboldt (33.4 °C). This supports the conclusion that the difference in starting temperatures between the two curves shown in Figure 26 is primarily due to the difference in operating pressure.

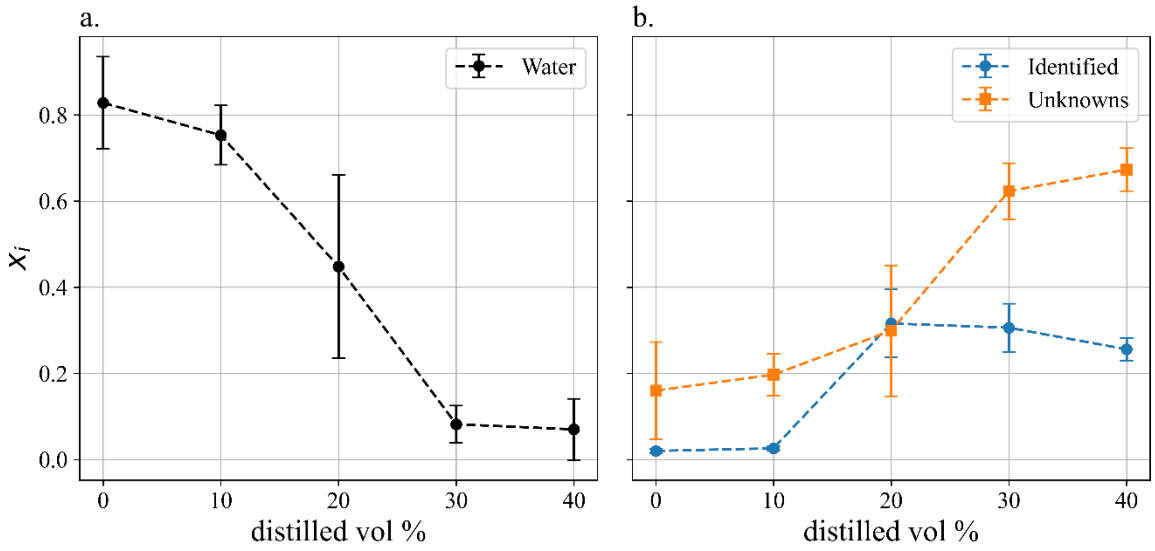


Figure 27. Mass fractions of water (a), and GC–MS/FID identified vs. unknown compounds (b) in the ADC distillates.

To further assess the phase behavior of water, its experimental activity coefficient (γ) in the initial distillation stage was estimated via modified Raoult's law. Based on the water content measured in the condensate at the starting point and the composition of the liquid surrogate mixture (Table 16), the mole fractions of water in each phase were determined. Two scenarios were considered to account for the uncertainty in the composition of the pyrolytic lignin (PL) surrogate: one assuming the smallest PL molecule in the surrogate set (D-B1) and another the largest (TE2). Calculations estimated activity coefficients of 1.76 and 1.42, respectively, which indicates a moderate positive deviation from ideal behavior, consistent with previous findings by Ille et al. [86]. Notably, the higher γ value was obtained for the smaller PL surrogate, which can be attributed to mole fraction balance, in which smaller molecules contribute a greater number of moles to the mixture, thereby lowering the mole fraction of water and increasing the calculated activity coefficient.

Figure 27b presents the mass fraction of compounds identified by GC–MS/FID and the remaining "unknown" fraction, calculated by subtracting the identified species and water content from the total. In the initial fractions, only a small portion of the distillate could be identified, and a notable fraction remained classified as unknown. The presence of these unidentified components at such early stages is somewhat unexpected and may arise from limitations in the GC–MS/FID analysis and/or an underestimation of the water content. At 20% distilled volume, both identified and unknown fractions reached amounts around 30 wt.% and exhibited the highest standard deviations. This correlates with the peak deviations in

temperature and water content, indicating an increased complexity in this fraction. In the final stages of the distillation (30-40% distilled volume), the amount of identified compounds gradually declined, and with water content at its lowest, the unknown fraction became increasingly dominant, probably composed of pyrolytic lignin oligomers that are beyond the detection range of the GC–MS/FID.

4.3.2 Modeling of FPBO properties

To evaluate how well different PL surrogate molecules represent the bulk characteristics of the lignin-derived FPBO, key physicochemical properties were analyzed. Figure 28 presents a Van Krevelen diagram (H/C vs. O/C) for the surrogate mixtures using different PL molecules, against the experimental value for the lignin bio-oil. Each point represents one surrogate, with the color scale indicating the MARE in elemental composition compared to the experimental result. Full results for each individual molecule can be found in Table S10 in the SI.

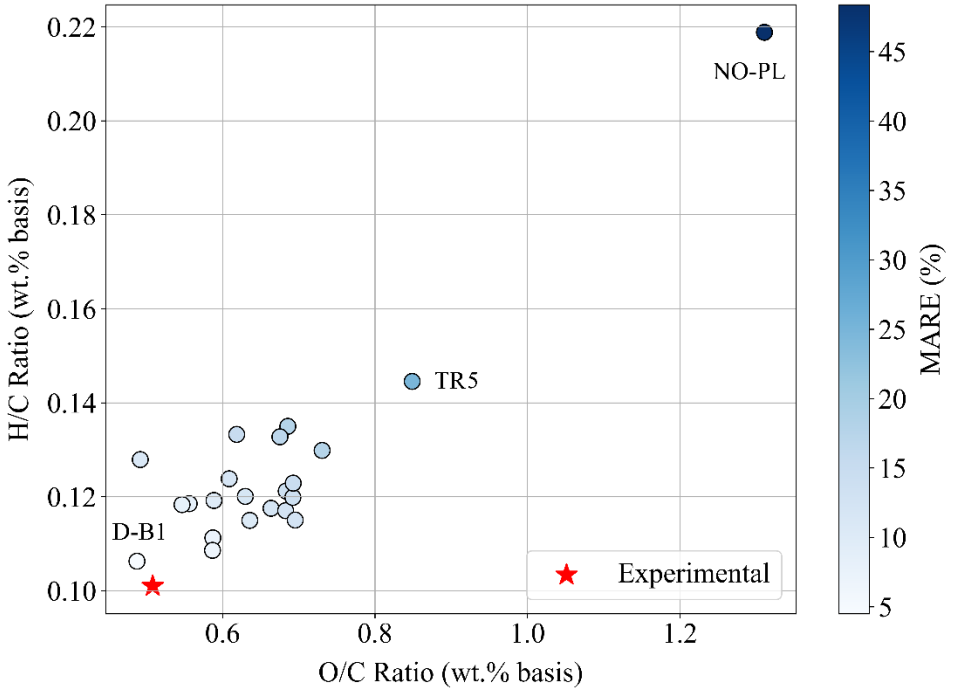


Figure 28. Van Krevelen diagram (H/C vs. O/C) of bio-oil composition predicted by the different surrogate molecules. Color scale shows the error (MARE) in reproducing the experimental values.

From the graph, it is clear that the introduction of a PL surrogate significantly improves the representation of the lignin bio-oil's elemental composition, with the No-PL case showing the largest error (45%), twice the error of worst performing surrogate. The best-performing PL surrogate, D-B1, is a small biphenyl-type dimer with just two oxygen atoms, yielding a MARE

of 4.5%. In contrast, TR5, a bulky trimer containing 15 oxygen atoms, produced the highest error at 24.8%. These results indicate that PL surrogates with lower oxygen and higher carbon content more effectively capture the elemental composition of pyrolytic lignin. This suggests that the actual PL structures present in the heavy fraction of the bio-oil may be less oxygenated than those often assumed in the literature.

Accurately predicting the density of the bio-oil is essential as it reflects how well the surrogates capture bulk physical properties. Figure 29 compares simulated density values (ρ) for surrogate molecules as a function of their molecular weight (MW) using two different methods: the Rackett equation, which is used by the IDEAL and UNIFAC-DMD models to estimate the liquid density, and PR-BM, which determines it from its own equation of state [129]. Full results are available in the SI (Table S10). For the Rackett estimations (Figure 29a), the best agreement was observed for dimers, particularly those in the 200–350 g/mol range, with most values falling close to the experimental target. As molecular weight increased, the predicted density generally declined, with most trimers and tetramers significantly underestimating the density, with the exception of TE1, which over predicted but still showed a considerably lower error.

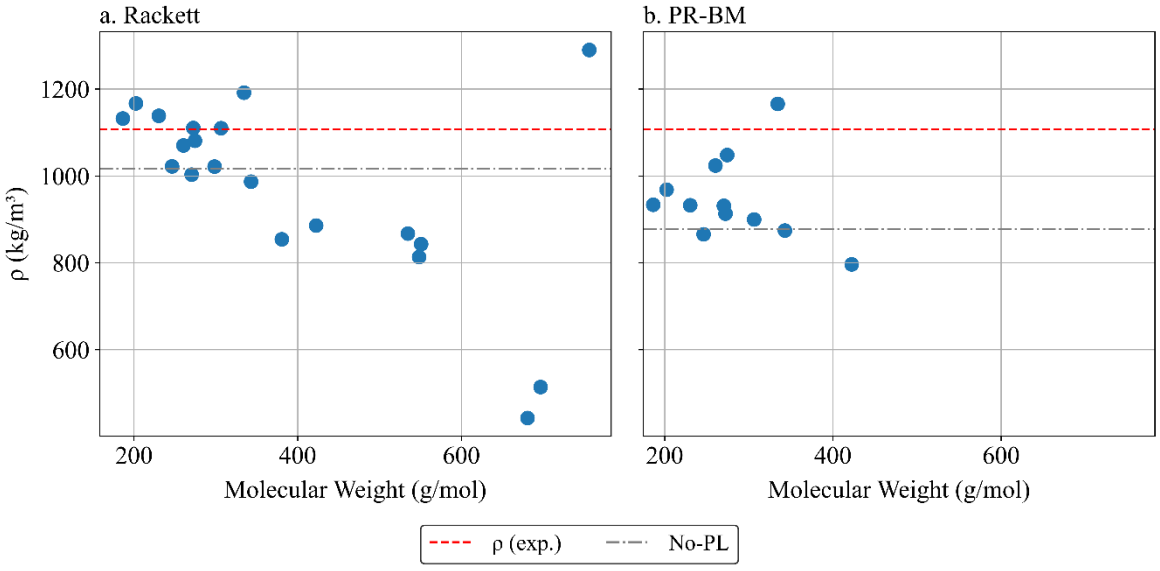


Figure 29. Liquid density (ρ) predictions of mixture featuring different PL surrogate molecules. (a) Using the Rackett method; (b) using the PR-BM model.

In contrast, the PR-BM method (Figure 29b) systematically predicted lower density values across all surrogates, showing larger deviations from the experimental target. While dimers still performed comparatively better than larger molecules, their predicted densities were generally lower than those obtained using the Rackett method. The discrepancy became more pronounced

for trimers and tetramers, which showed limited predictive reliability, with many of these molecules encountering simulation issues and yielding zero-value predictions that were no included in the graph. These findings reinforce the limitations of cubic equations of state in representing liquid densities of heavy aromatics, as noted in previous studies [84].

The No-PL case underestimated the experimental density with both methods (Figure 29), with the Rackett method performing better than PR-BM, which showed a larger underestimation. This shows that including a PL surrogate can improve predictive accuracy; however, this improvement depends strongly on the surrogate choice, as higher molecular weight candidates can often underestimate the density or lead to simulation issues.

The combined analysis of elemental composition and density showed that dimers yielded more reliable predictions, whereas high molecular weight surrogates, particularly those rich in oxygen, were less accurate and more prone to simulation errors. The D-B group (biphenyl-type dimer) consistently showed the best performance as surrogate for the pyrolytic lignin fraction, with other surrogates from the D-A (β-O-4), D-C (stilbene), and D-D (phenylcoumaran) groups also yielding strong results. These findings suggest that small, condensed aromatic dimers with low oxygen content are particularly well-suited to represent the physical properties of pyrolytic lignin in thermodynamic modelling.

4.3.3 Modeling of vapor-liquid equilibrium

To evaluate the ability of PL surrogates to reproduce VLE behavior, their performance was assessed using the experimental ADC data as benchmark. At first, the surrogates were evaluated by group, using the mean MARE for each group, with error bars indicating the standard deviation within the group. Some surrogates encountered simulation issues: TE2 failed to converge with the UNIFAC-DMD model and, while it ran with the IDEAL model, produced unrealistic results and it was removed from the analysis. In the PR-BM model, convergence errors were observed for TR5 and TE3. Figure 30 compares the mean MARE values for temperature and water content predictions, highlighting how the choice of surrogate group and thermodynamic model influences accuracy. Data associated with individual PL surrogates can be found in the SI (Table S11).

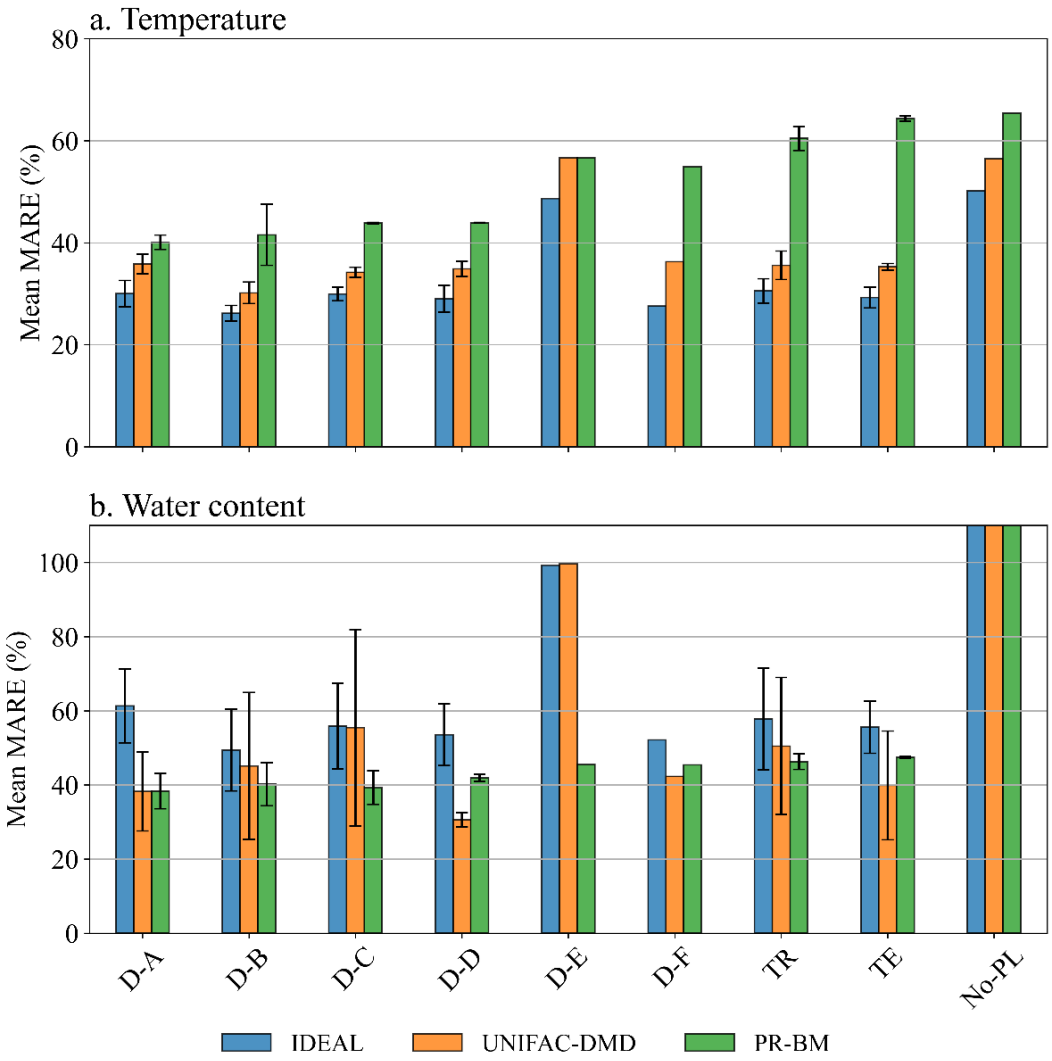


Figure 30. Mean MARE values for (a) temperature and (b) water content prediction across surrogate groups using IDEAL, UNIFAC-DMD, and PR-BM models.

For the temperature predictions, shown in Figure 30a, most molecular groups performed comparably similar, with the exception of group D-E that clearly showed the worst performance. Regarding the influence of thermodynamic models, IDEAL consistently achieved the best performance (with MARE values often below 30%), closely followed by UNIFAC-DMD which showed similar performance in several groups, while PR-BM consistently produced the highest deviations (>40%). This is surprising given that the IDEAL mixture model neglects intermolecular interactions, which are typically considered critical for representing PL behavior. The absence of a PL surrogate (No-PL) led to significantly higher deviations (>50%), underscoring the importance of including a representative for this fraction in the mixture. Variability within groups was generally small, with the exception of group D-B for the PR-BM model, suggesting that functional group variations may have little effect on temperature predictions.

Figure 30b shows the deviations for the evolution of the water content predictions. In contrast to the temperature predictions, here the IDEAL model consistently produced the largest errors (>50%). PR-BM showed comparatively smaller errors compared to its temperature predictions, and was even able to achieve lower errors for D-E, where the other models deviated by almost 100%. The UNIFAC-DMD model showed a broader error range (30-60%) and the widest intragroup variation (D-C), yet it also provided the best-performing individual PL surrogates for all groups except the D-E group. These findings suggest that, in contrast to temperature predictions, functional group variations may be more relevant for water content predictions, as reported by Ille et al. [86]. The No-PL case once again produced the highest errors (>100%), reinforcing the need to include a representative surrogate for the pyrolytic lignin fraction.

4.3.4 Best pyrolytic lignin surrogate candidates

The best PL surrogate molecules for each property and model are compared in Table 19. While no single PL surrogate excelled across all evaluations, dimers dominated the rankings overall. The biphenyl group consistently ranked among the top performers, with D-B1 standing out as the best surrogate across most properties, achieving the lowest elemental composition error (RE = 4.50 %) and the best temperature prediction under the IDEAL model (MARE = 24.09 %). D-B2 and D-B3 also showed strong performance, reaching error values of 6.70 % and 25.67 %, respectively, for elemental composition and water content predictions. It is worth noting that D-B2 (3,4,4-biphenyltriol) was also selected in previous studies by Ille et al. [86] and Krutof and Hawboldt [90] as a representative compound for modeling the pyrolytic lignin in VLE, further supporting its relevance. Other dimer groups, including D-A (β-O-4), D-C (stilbene), and D-D (phenylcoumaran), also performed well across several properties. For example, D-A1 achieved a MARE of 28.47 % for water content under UNIFAC-DMD, and D-D2 showed one of the lowest density deviations (RE = 0.25 %). In contrast, trimers and tetramers only appear in the ranking for water content predictions, with surrogates TR2 (MARE = 27.59 %) and TE3 (MARE = 29.53 %) under the UNIFAC-DMD model. These results reinforce the superior suitability of dimers as PL surrogates, as trimers and tetramers not only ranked lower in most categories but also led to simulation convergence issues in some cases.

Table 19. Top five performing surrogates across all categories, ranked by lowest error values.

	Input bio-oil properties					
	C-H-O (wt. %)		Density – Rackett		Density – PR-BM	
	Surrogate	**MARE (%)**	**Surrogate**	**RE (%)**	**Surrogate**	**RE (%)**
1	D-B1	4.50	D-B4	0.25	D-A2	5.27
2	D-B2	6.70	D-D2	0.25	D-A4	5.35
3	D-A1	7.39	D-C1	0.25	D-B2	12.5
4	D-E1	8.01	D-B1	2.26	D-B1	15.7
5	D-D2	8.44	D-A4	2.37	D-A1	15.8
	Temperature profile					
	IDEAL		**UNIFAC-DMD**		**PR-BM**	
	Surrogate	**MARE (%)**	**Surrogate**	**MARE (%)**	**Surrogate**	**MARE (%)**
1	D-B1	24.09	D-B1	28.01	D-B2	36.37
2	D-B3	26.40	D-B2	29.14	D-B1	36.45
3	D-B2	26.47	D-B3	30.70	D-A4	38.67
4	D-D1	27.17	D-B4	32.86	D-A3	39.55
5	D-A4	27.26	D-C2	33.47	D-A1	40.17
	Water content profile					
	IDEAL		**UNIFAC-DMD**		**PR-BM**	
	Surrogate	**MARE (%)**	**Surrogate**	**MARE (%)**	**Surrogate**	**MARE (%)**
1	D-B3	35.80	D-B3	25.67	D-A4	32.40
2	TR2	39.03	TR2	27.59	D-B2	33.35
3	D-B1	46.34	D-A1	28.47	D-C2	36.12
4	D-A1	47.53	D-D2	29.30	D-A3	36.80
5	D-D2	47.66	TE3	29.53	D-B1	38.59

The results also show that the choice of thermodynamic model and PL surrogate are closely linked, with rankings varying across the different models. While for IDEAL and UNIFAC-DMD, the best performing surrogates were similar, with the D-B group dominating across all properties, for the PR-BM model, surrogates from the D-A group (particularly D-A2 and D-A4) also ranked among the top. This highlights the importance of evaluating surrogate performance in conjunction with the thermodynamic model used, as different models prioritize different molecular features.

In addition to the evaluation of errors, the ability to model distillation profiles must also be considered. Figure 31 compares the top two surrogates for each model and the No-PL case against the experimental ADC data for temperature and water content profiles. The selected

profiles highlight how individual surrogates capture the VLE behavior of the pyrolytic lignin fraction, allowing for direct comparison across different thermodynamic models. For clarity, only the top two molecules from each model are shown in the main figures, while the complete top five rankings are provided in Figure S2 the SI.

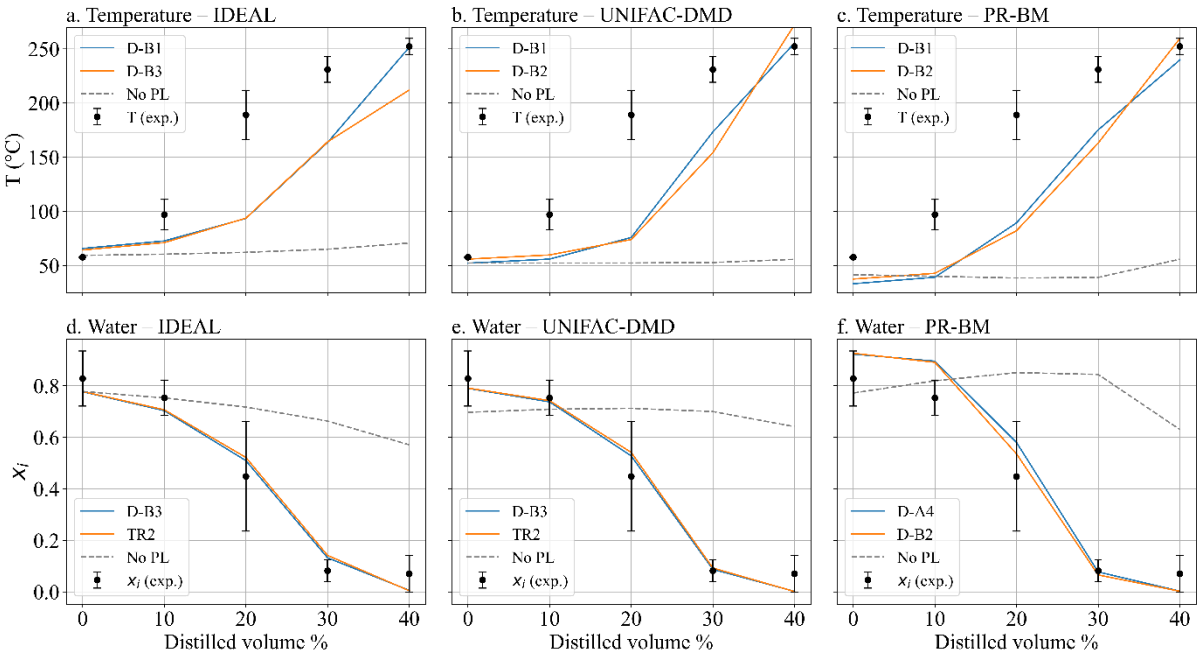


Figure 31. Simulated ADC temperature and water content profiles for top-performing surrogate molecules and No-PL case across thermodynamic models: (a) & (d) IDEAL, (b) & (e) UNIFAC-DMD and (c) & (f) PR-BM.

Regarding the temperature, all PL surrogates tended to underestimate the experimental curve, resulting in a convex shaped simulation profile, similar to the one observed by Krutof and Hawboldt [90], and more accurately captured in their modeling. For the UNIFAC-DMD model (Figure 31b), D-B2 closely matched the initial distillation temperature, while D-B1 showed a slight underestimation. The IDEAL model (Figure 31a) overestimated this point, whereas PR-BM (Figure 31c) strongly underestimated it. All simulated curves fail to capture the sharp temperature rise observed between 10-20 vol.%. Instead, they consistently underestimate the boiling temperature throughout the distillation and show the steepest incline to the 20-30-40 vol.% range, where the experimental curve actually begins to flatten. The last distillation point is well captured by D-B1 in both the IDEAL and UNIFAC-DMD models and by D-B2 in the PR-BM model (Figure 31a-c). While the surrogates still captured key features of the temperature curve, the No-PL case consistently failed to do so, producing a flat and inaccurate profile across all models. This again reaffirms the need to include a pyrolytic lignin surrogate to achieve more accurate VLE predictions.

Figure 31d, e, and f show the water content profiles obtained with the different thermodynamic models. All models are able to capture the general trend of the experimental data across the distillation range, with the UNIFAC-DMD model (Figure 31e) showing the best overall agreement and the lowest errors. Despite its comparatively larger deviations, the IDEAL model (Figure 31d) follows a curve shape similar to that of UNIFAC-DMD. Higher errors for this model may stem primarily from mismatches at the final distillation points. In contrast, the PR-BM model (Figure 31f) was the only one to overestimate the water content in the initial stages of the distillation, followed by a sharper decline. Notably, the overall shape of the water content curve remained consistent across different PL surrogates, with the choice of molecule mainly influencing the magnitude of the errors rather than the general trend of the predictions. As with the temperature predictions, the No-PL case failed to capture the experimental trend, producing an inaccurate water content profile across all models.

Evaluating the simulated activity coefficient of water at the first distillation point for molecule D-B1 reveals important differences. For the UNIFAC-DMD model, the predicted γ_{water} was 1.87, which aligns well with the experimentally estimated value of 1.76. This agreement, together with its accurate water content predictions, highlights the model's suitability for capturing non-ideal behavior in fast pyrolysis bio-oils. In contrast, PR-BM predicted a much higher γ_{water} of 6.3 for the same surrogate, correlating with its poor performance in the initial distillation region and suggesting limited applicability for these systems.

It is also important to note that deviations between modeled and experimental curves may also arise from uncertainties in both the experimental procedure and analytical methods. In particular, offsets in temperature and distilled volume readings can contribute to the observed deviations. Additionally, uncertainties in quantification of the fast pyrolysis bio-oil and especially in the small distillate samples can also affect the results.

4.3.5 Assessment of unknowns

The assessment of the unknown fraction was treated separately from the main VLE error analysis due to its inherently higher uncertainty. The analysis focused on evaluating trends in the predicted profiles, providing qualitative insights rather than quantitative error assessments. Figure 32 compares the simulated profiles of selected PL surrogates against the experimental distribution of the unknown fraction using the IDEAL, UNIFAC-DMD, and PR-BM models.

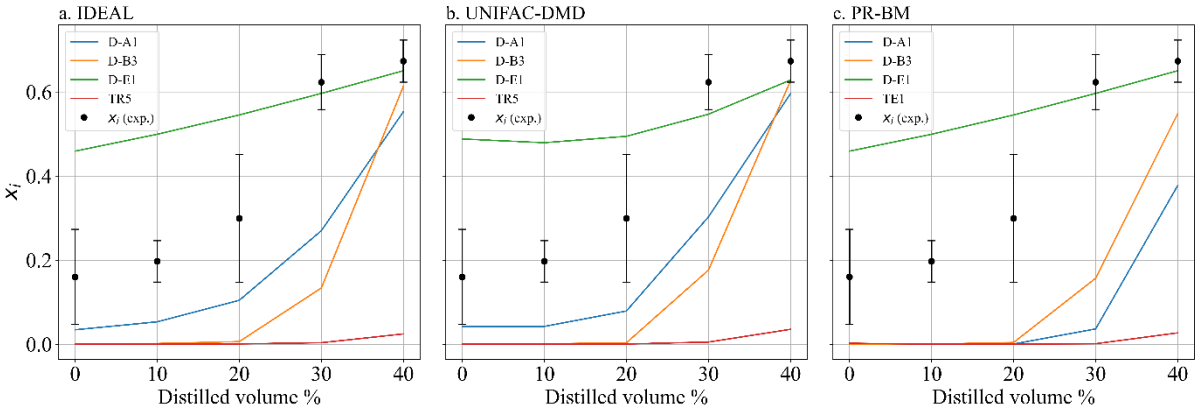


Figure 32. Simulated ADC unknowns profile for selected PL surrogates across the thermodynamic models (a) IDEAL, (b) UNIFAC-DMD and (c) PR-BM.

Overall, most dimers started to distil around 20 vol.%, which aligns well with the experimental profile if we consider that the apparently high concentration of unknowns in the early fractions is likely caused by an underestimation of water content and/or other volatiles. A few dimers, however, showed earlier onset of distillation, particularly D-E1 for all models and D-A1, D-C1 and D-D2 for the IDEAL and UNIFAC-DMD model. Among these, D-E1 exhibited an especially atypical behavior, predicting disproportionately high concentrations of unknowns in the distillate from the very beginning of the distillation. This unusual trend is consistent with D-E1's poorer performance in simulating other properties, suggesting that it may not be a suitable standalone surrogate for representing the PL fraction. For the trimers and tetramers, distillation typically began later, around 30-40 vol.%, and in some cases even beyond this range. Among all tested candidates, the surrogate that best captured the overall trend of the unknown fraction was D-A1, followed by D-B3. Additional profiles are provided in the SI (Figure S3).

These findings provide insights into the distillation behavior of the lignin-derived fraction of the bio-oil. While dimers dominate the onset of unknowns release, their contributions are not uniform, with some species exhibiting earlier volatility and others aligning more closely with the mid-distillation range. In contrast, bulkier oligomers contribute primarily at later stages, reflecting their higher molecular weights. Importantly, even surrogates such as D-E1, despite their unrealistic concentration patterns, could still play a useful role when designing surrogate mixtures with multiple representatives for the PL fractions, as they can contribute to capturing early-distilling behavior. Understanding the distinct contributions of each surrogate enables for a more informed design of surrogate mixtures for the pyrolytic lignin, supporting a more accurate phase equilibrium modeling.

4.4 Conclusion

In this study, we investigated the phase equilibrium of fast pyrolysis bio-oil (FPBO) by combining advanced distillation curve (ADC) experiments with vapor-liquid equilibrium (VLE) simulations. The influence of various pyrolytic lignin (PL) surrogates was assessed across different thermodynamic models to evaluate their ability to reproduce experimental distillation behavior. These findings contribute to a deeper understanding of FPBO phase behavior and help guide the design of surrogate mixtures that can better predict VLE calculations of FPBO, ultimately improving the predictive accuracy of process simulations.

The results highlight the critical role of surrogate selection in accurately representing the PL fraction. In comparison to the No-PL case, using a PL surrogate generally improved model accuracy, particularly for elemental composition and VLE behavior. For the specific case of density predictions, a more cautious selection is recommended, as some surrogates yielded less accurate results than when no surrogate was used.

Among all tested molecules, dimers consistently outperformed larger oligomers in capturing key FPBO properties and distillation behavior. In particular, the biphenyl-based surrogates (D-B group) ranked among the best performers across most evaluations, while other dimers such as D-A (β-O-4), D-C (stilbene), and D-D (phenylcoumaran) also showed strong performance. Among larger surrogates, such as trimers and tetramers, a few molecules (e.g., TR2 and TE3) performed reasonably well for water content predictions using the UNIFAC-DMD model. However, overall, these higher molecular weight surrogates were less suitable as standalone representations of the PL fraction, as some also led to simulation errors, especially when using the PR-BM model.

The evaluation of thermodynamic methods showed that the IDEAL model, despite not accounting for molecular interactions, delivered reasonably good results, including the lowest deviations in the temperature curve. This performance may be partly attributed to a favorable offset in the initial boiling point, which coincidentally overlaps the experimental data. Overall, its predictions were comparable to those of the UNIFAC-DMD model, which was the most accurate at predicting the water content and also showed consistency with the calculated experimental activity coefficient of water, indicating its suitability for modeling non-ideal behavior in fast pyrolysis bio-oils. In contrast, the PR-BM model consistently showed higher deviations in temperature predictions and frequently delivered inferior performance overall. Furthermore, its strong overestimation of the water activity coefficient suggests that it may be less appropriate for accurately capturing the phase behavior of such systems.

The results also revealed a strong dependence between the optimal surrogate choice and the thermodynamic model employed. While IDEAL and UNIFAC-DMD generally favored similar top-performing surrogates, with the D-B group dominating, the PR-BM model showed a preference for group D-A in several cases.

Analysis of the unknown fraction provided further insights into the distillation characteristics of lignin-derived compounds. While some dimers initiate distillation earlier than expected, most remain stable and started distilling around the 20 vol.% mark. In contrast, larger oligomers only started distilling at higher distillation volumes (30-40 vol.%). Given these differences, mixtures of surrogates with complementary characteristics could help improve how the PL fraction is represented in phase equilibrium simulations in future calculations.

Chapter 5. Liquid-liquid equilibrium prediction in fast pyrolysis bio-oil systems: A framework for incorporating bio-oil complexity

This chapter is based on the following publication:

Liquid–Liquid Equilibrium Prediction in Fast Pyrolysis Bio-Oil Systems: A Framework for Incorporating Bio-Oil Complexity

Myriam Rojas Salas, Frederico Gomes Fonseca, Ana Cristina Corrêa de Araujo, Manuel García-Perez, Axel Funke

Energy & Fuels, 2024, 38, 19, 18769–18780. DOI: 10.1021/acs.energyfuels.4c03387

5.1 Introduction

Building upon the findings of the previous chapter, which addressed the vapor–liquid equilibrium (VLE) behavior of fast pyrolysis bio-oils (FPBOs), this chapter focuses on the complementary study of liquid–liquid equilibria (LLE), which is equally important for understanding and designing separation processes. The coexistence of compounds with widely differing polarities often leads FPBOs to form separate liquid phases that can play a decisive role in product recovery and upgrading. However, accurately predicting these equilibria remains difficult due to the compositional complexity of FPBOs and the presence of unresolved high molecular weight fraction, mainly derived from lignin.

In this chapter, we investigate the influence of representative surrogate molecules for describing the phase behavior of pyrolytic lignin (PL) oligomers in LLE modeling of FPBOs. The set of surrogate molecules applied here is similar to the one used in the previous chapter for the VLE study, allowing for a consistent assessment of their predictive capability across different phase equilibrium conditions. Experimental data available in the literature are used as reference for validation, and the results highlight key trends in the phase separation behavior of relevant model compounds.

5.2 Methodology

This study is based on experimental results for FPBO extraction (from BTG, The Netherlands), as reported by Han et al [133]. To reduce the uncertainties by including a large number of compounds with unknown thermophysical properties, the number of different species in the surrogate mixture was lowered according to the heuristics proposed by Fonseca and Funke [84]. The original mixture and the two case surrogate mixtures employed are presented in Table 20. The characterization of the original mixture was performed using GC-MS analysis to identify and quantify volatile components; in addition, the pyrolignin (PL) was quantified via cold water precipitation. Therefore, the resulting unknowns from GC-MS analysis most likely include primarily substituted sugars. The compounds used in the surrogate mixture are numbered in the first column; all compounds without numbers are added to the previous surrogate to determine its fraction (e.g., butyrolactone and 2(5H)-furanone are added to furfural). Case 1 refers to a surrogate mixture with a direct normalization of all surrogates to meet 100%. Case 2 refers to the scenario in which the unknown fraction of the original mixture (~21 %) is added to the PL fraction prior to normalization. Null case (∅), was simulated to identify the effect of representing the PL with lignin oligomer structures in the system.

Table 20. FPBO composition provided by Han et al. [133] and definition of relevant cases (null, 1, 2)

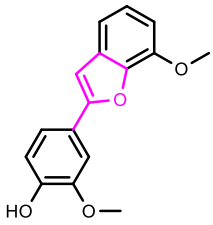
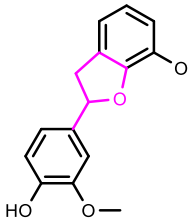
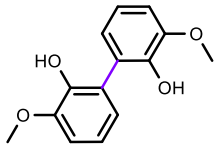
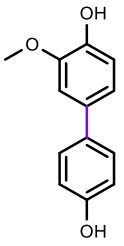
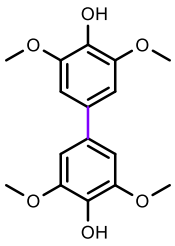
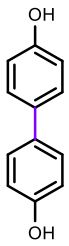
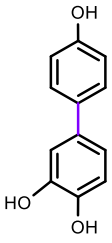
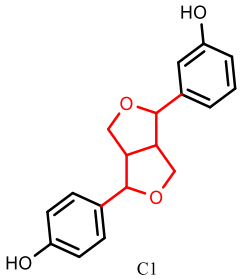
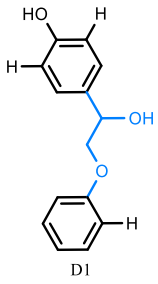
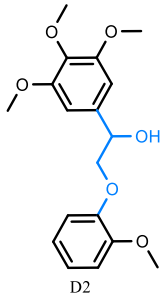
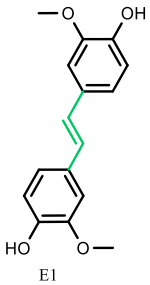
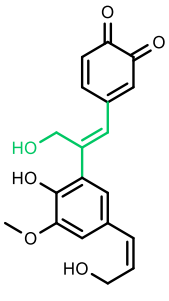
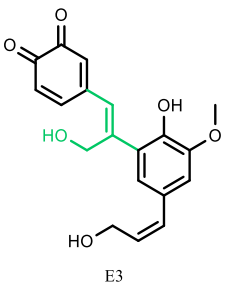
Number	Compound	Base FPBO (wt.%)	Null Case (Ø) (wt.%)	Case 1 (wt.%)	Case 2 (wt.%)
1	Water (by KF titration)	32.4	56.92	41.07	35.64
Carboxylic Acids, Ketones, and Furans:					
2	Glycolaldehyde	0.68	1.46	1.05	0.91
3	Acetic acid	1.4	3.00	2.16	1.88
4	Acetol	1.21	2.59	1.87	1.62
	Butanedial	0.07			
5	2-cyclopenten-1-one	0.19	0.41	0.29	0.25
6	Furfural	0.31	1.20	0.87	0.75
	Butyrolactone	0.06			
	2(5H)-furanone	0.19			
7	Corylon	0.6	1.29	0.93	0.80
8	Furfural. 5-hydroxymethyl-	0.33	0.71	0.51	0.44
	Other carboxylic acids/ketones/furans	1.09			
Phenolics:					
9	Phenol	0.13	0.34	0.24	0.21
10	Phenol. 2-methoxy-	0.76	1.97	1.42	1.23
	Phenol. 2-methyl-	0.1			
	Phenol. 4-methyl-	0.24			
	Phenol. 3-methyl-	0.05			
11	Phenol. 2-methoxy-4-methyl-	0.86	2.22	1.60	1.39
	Phenol. 3.4-dimethyl-	0.12			
	Phenol. 4-ethyl-	0.05			
12	Phenol. 4-ethyl-2-methoxy-	0.46	1.19	0.86	0.75
	Phenol. 2-methoxy-4-vinyl-	0.07			
13	Phenol. 2-methoxy-5-(2-propenyl)-	0.34	2.25	1.62	1.41
	1.2-benzenediol	0.12			
	Phenol. 2.6-dimethoxy-	0.05			
	Phenol. 2-methoxy-4-(1-propenyl)-. (E)-	0.27			
	Phenol. 2-methoxy-4-(1-propenyl)-. (Z)-	0.26			
	Phenol. 4-methoxy-3-(methoxymethyl)-	0.05			
14	Vanillin	0.41	1.06	0.77	0.66
	Phenol. 2-methoxy-4-propyl-	0.24			
15	4-acetyl-guaiacol	0.3	1.22	0.88	0.76
	Guaiacyl-acetone	0.17			
	Other phenols	0.78			
Sugars:					
16	Levoglucosan	9.57	22.14	15.97	13.86
	Other anhydrosugars	3.03			
Others:					
17	Pyrolignin (PL)	22		27.89	37.41
	Total	**78.96**	**100**	**100**	**100**

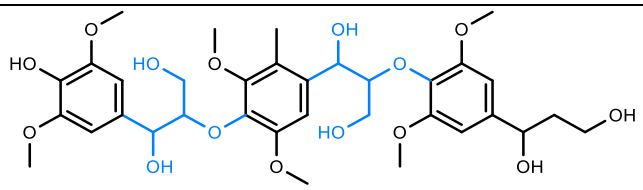
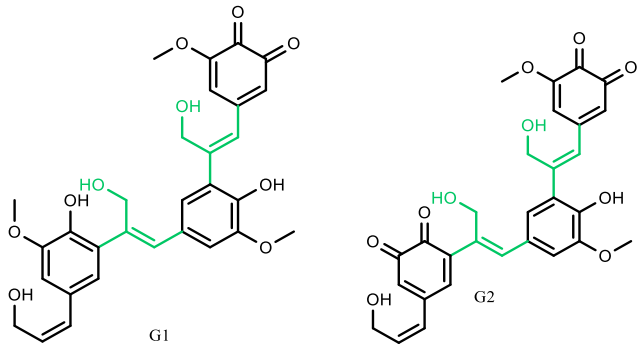
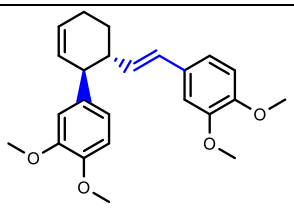
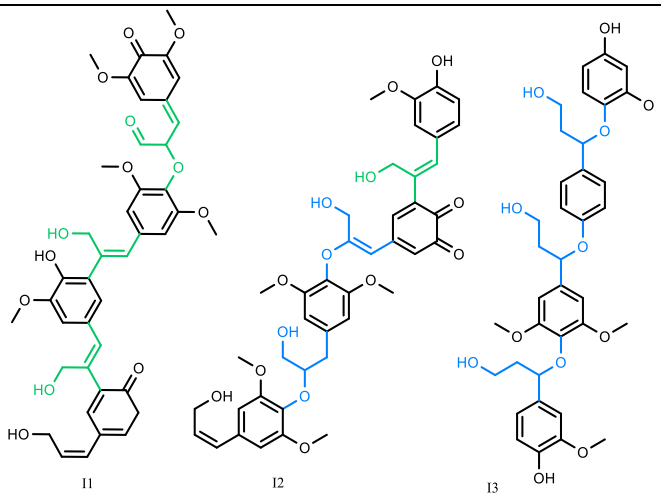
Twenty possible identities were considered to represent the PL fraction to study the influence on Liquid-Liquid Equilibrium (LLE) prediction, which are presented in Table 21. As shown in Table 21, the lignin oligomer structures were grouped based on the similarity of their bonds and organized by letters from A to H (dimers and trimers). The tetramers in group I consist of two homogeneous structures with either β-5 or β-O-4 bonds and one heterogeneous structure containing both types of bonds. SM is a mixture of structures proposed by Fonts et al. [88] to represent PL, including low molecular mass lignin (LMM) (molecular mass of 400 amu and polydispersity of 1.7) and high molecular mass lignin (HMM) (molecular mass of 1050 amu and polydispersity of 2.3).

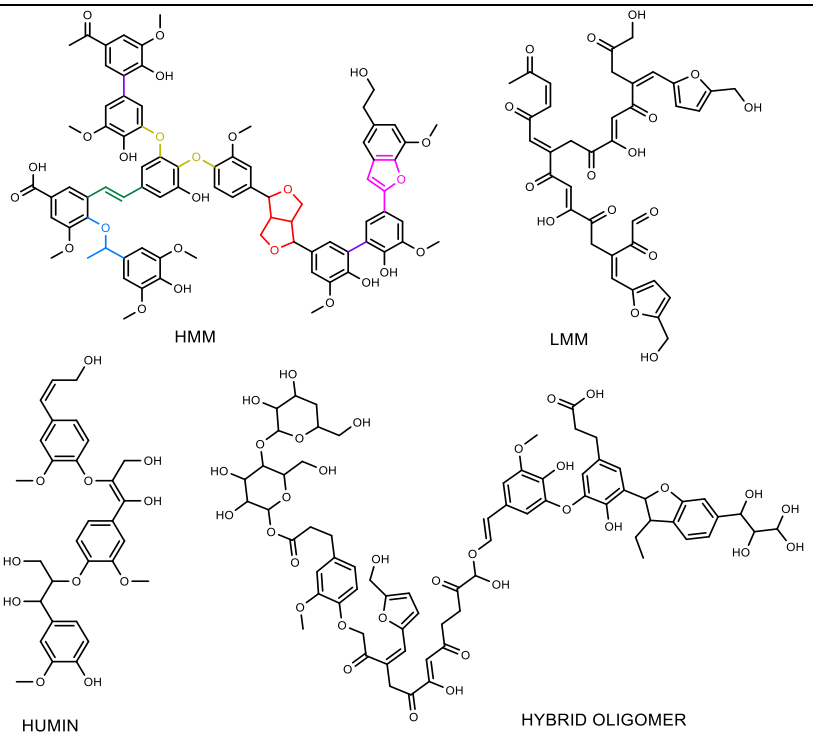
The system was modeled using Aspen Plus™ V14 (see scheme in Figure S4), where the bio-oil (FPBO) is modeled as the mixture presented in Table 20 (case 1 or case 2). Pyrolignin refers to one of the proposed oligomer structures summarized in Table 21, or to a case where no pyrolignin is present. All streams (bio-oil, water, solvent) are modeled as liquid streams at operation temperature (293 K), according to the methodology employed in the experiments that serve as reference datasets for this study [133]. The liquid streams are mixed in a mixer unit block, and the product is channeled to a decanter unit block. The decanter operates at 293 K and water is defined as the key component of the 2nd liquid, branding it as an aqueous phase. Phase split is modeled by equalizing component fugacities in the two liquids, and liquid-liquid coefficients are estimated from the property method. The property method employed is UNIFAC-Dortmund (DMD), chosen due to its ability to model molecules which little experimental data is available, such as many of the molecules present in the bio-oil surrogate used in this work. The software is able to automatically generate the group distribution based on the inputted molecular structure for different molecules, in order to estimate thermophysical properties such as normal boiling point, critical properties, etc... The division of LLE UNIFAC groups of the structures used in this work is detailed in Table S12 in the SI.

The Aspen Plus™ model was connected to Microsoft Excel™ via the Aspen Simulation Workbook™ to facilitate data collection. Two sets of results are considered based on varying proportions of bio-oil, water, and butanol, according to the experimental setup used for validation [133]. Within the BE system (extraction with butanol), water was set at 30 g and bio-oil at 15 g and kept constant, while the solvent (n-butanol) input was varied six times (BE_1: 6 g, BE_2:10 g, BE_3:15 g, BE_4:22.5 g, BE_5:30 g, and BE_6:60 g). In the WE system (extraction with water), the solvent was set at 22.5 g and the bio-oil at 15 g and kept constant; the water input was varied six times (WE_1: 7.5, WE_2:15, WE_3:22.5, WE_4:30, WE_5:45, and WE_6:60 g).

Table 21. Chemical structures of the pyrolignin model molecules.

Group	Structure	Relevant link types	Reference
	Dimers		
A	 A1  A2	Phenylcoumaran β-5	[130]
B	 B1  B2  B3  B4  B5	Biphenyl 5–5 linkage	[130,134,135]
C	 C1	Resinol β-β	[130]
D	 D1  D2	Alkyl-aryl ether β-O-4	[136,137]
E	 E1  E2  E3	Stilbene β-5	[89]

Trimers			
F	 F1	Alkyl-aryl ether β-O-4	[138,139]
G	 G1 G2	Stilbene β-5	[89]
H	 H1	Bridging double bond	[131]
Tetramers			
I	 I1 I2 I3	Mixed links: stilbene β-5 (sometimes enolized), alkyl-aryl ether β-O-4	[89,138,139]

SM [a]		Custom mixture of structures, including polysaccharides and hybrid solutions.	[88]

a: HMM (28.08%), LMM (25.96%), HUMIN (17.87%), and Hybrid oligomer (28.08%). Structures and proportions adapted from [88].

Han et al. [133] defined parameters for evaluating their experimental results. The mass ratio between the organic and aqueous phase (Ratio O/A) was estimated according to Eq. 8. The partitional coefficient of compound or component *i* that distributes in both phases ($K_{OW,i}$) is estimated according to Eq. 9.

$$Ratio\ O/A = \frac{m_{org}\ [g]}{m_{aq}\ [g]} \quad 8$$

$$K_{OW,i} = \frac{m_{i,org}\ [g]/m_{org}\ [g]}{m_{i,aq}\ [g]/m_{aq}\ [g]} \quad 9$$

According to Han et al [133], the chemical composition of the phases, including the amount of n-butanol, was determined after separation by GC/MS. The water content was measured by Karl-Fisher titration, and the bio-oil content was determined by the difference for each phase in mass percentage as Eq. 10.

$$m_{bio-oil} = 100 - m_{butanol} - m_{water} \quad 10$$

Data filtering and processing, trend analysis, error calculations, and figure creation were performed using Matlab 2024a®. Relative deviations to the experimental data of Han et al. [133]

were estimated for every species present, as well as a compound entity named the *Bio-oil*, comprised of all species except water and solvent. For each case (1, 2) and system (BE, WE), deviations between calculations and experimental data are reported as mean relative error (MRE) as Eq. 11, and as relative errors (Re) for the individual points of each system obtained by Eq. 12 and Eq. 13.

$$MRE = \frac{1}{n}\sum_{i=1}^{n}|Re_i| \quad (11)$$

$$Abs\ Error = |x_m - x_e| \quad (12)$$

$$Re\ (\%) = \frac{Abs\ Error}{x_e} * 100 \quad (13)$$

Where x_e is each data predicted for the model and x_e is the experimental value.

The results presented in figures and data tables in the manuscript represent a choice of different trends and predictive capability. The complete curves and data are included in the supplementary information (SI). The main components evaluated were water, Bio-oil (considered all the compounds in Table 20, except water), and solvent (n-butanol).

5.3 Results and Discussion

5.3.1 Prediction of phase separation of the system water, n-butanol, and bio-oil

In Figure 33, one can see the evolution of the Ratio O/A with the increasing input of butanol (a, b, e, f) or water (c, d, g, h), for the best performant model molecules identified with letters and numbers as Table 21. Most of the model molecules present a very similar behavior, which can be attributed to their hydrophobic center and possible steric issues or shortcomings of using UNIFAC-DMD for the prediction of LLE. A clear outlier is the null case (∅), which reinforces the importance of employing a representation for the pyrolignin phase when modeling LLE phenomena, reinforcing the findings of previous researchers [84,140]. The worst MREs are found for the null case (∅), followed by the SM case (details for all curves are in Figure S5 in the SI). As SM is a mixture comprised of 4 structures of high molecular weight [88], the bad performance may reflect shortcomings when LLE modeling using UNIFAC-DMD.

The WE system (WE_2-WE_6 as in Figure 33 c, d, g, h; WE_1 consistently lays outside of the value range predicted by the model using any pyrolignin representative) presents better

predictions for the Ratio O/A (Figure 33) for either Case than the BE system (BE_1-BE_5 as in Figure 33 a, b, e, f; BE_6 consistently lays outside of the value range predicted by the model using any pyrolignin representative). While water is a common extraction medium for FPBO fractioning [80,82], the nature of the pyrolignin representatives (Table 21) would indicate, at a first glance, a larger tendency to migration to the organic phase.

Despite the worse predictions (Figure 33), the BE system is well predicted up to BE_3, and higher fractions of n-butanol cause the model to present increasingly higher degrees of underestimation of the Ratio O/A, indicating the model predicts a higher migration to the aqueous phase than was is observed experimentally. This is interesting due to the relatively high fractions of phenolics and pyrolignin present in the surrogate mixture (Table 20).

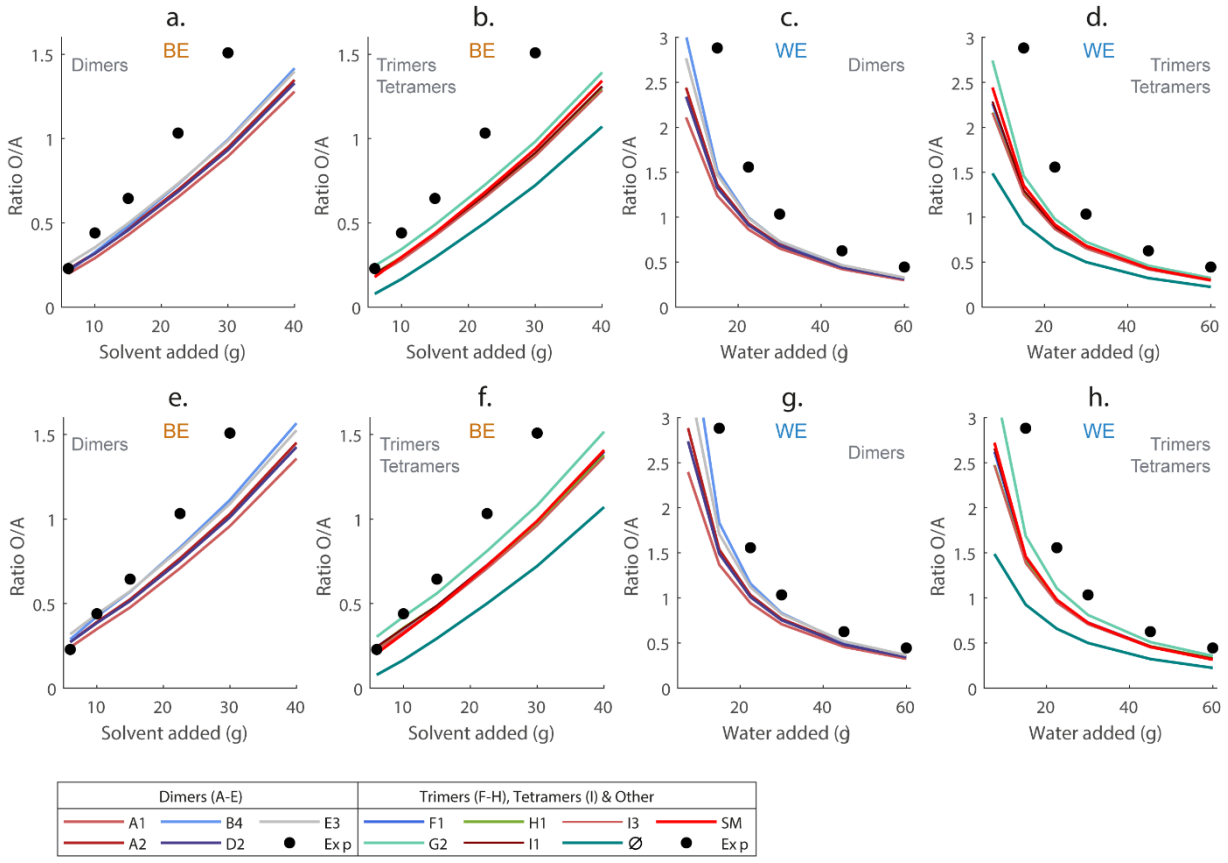


Figure 33. The variation of solvent (BE) and water (WE) and its impact on the MRE of the Ratio O/A for Case 1: a-d), and Case 2: e-h.

Figure 34 presents the deviations in the prediction of the partition coefficient (K_{OW}) for the main components of the system WE (*Bio-oil*, water, solvent). Case 1, which excluded water-soluble molecules from the pyrolignin fraction, demonstrated superior performance in predicting the distribution of *Bio-oil* compared to Case 2 (max MRE: 47% for Case 1 vs min MRE: >66% for Case 2). Similarly, a high MRE was found for the null case (ØMRE: 67%).

These results show that including water-soluble molecules in the pyrolignin fraction affects model performance and highlights the importance of accurately accounting for pyrolignin during bio-oil modeling. The system BE was not considered due to a lack of experimental data. For Case 1, the best results were obtained with G2 (MRE: 15 %), followed by E2, E3, C1, I1, and SM (MRE: <18 %), while the worse results were found for the dimers B5, B4, and B1 (MRE: 47%, 32%, and 25%, respectively).

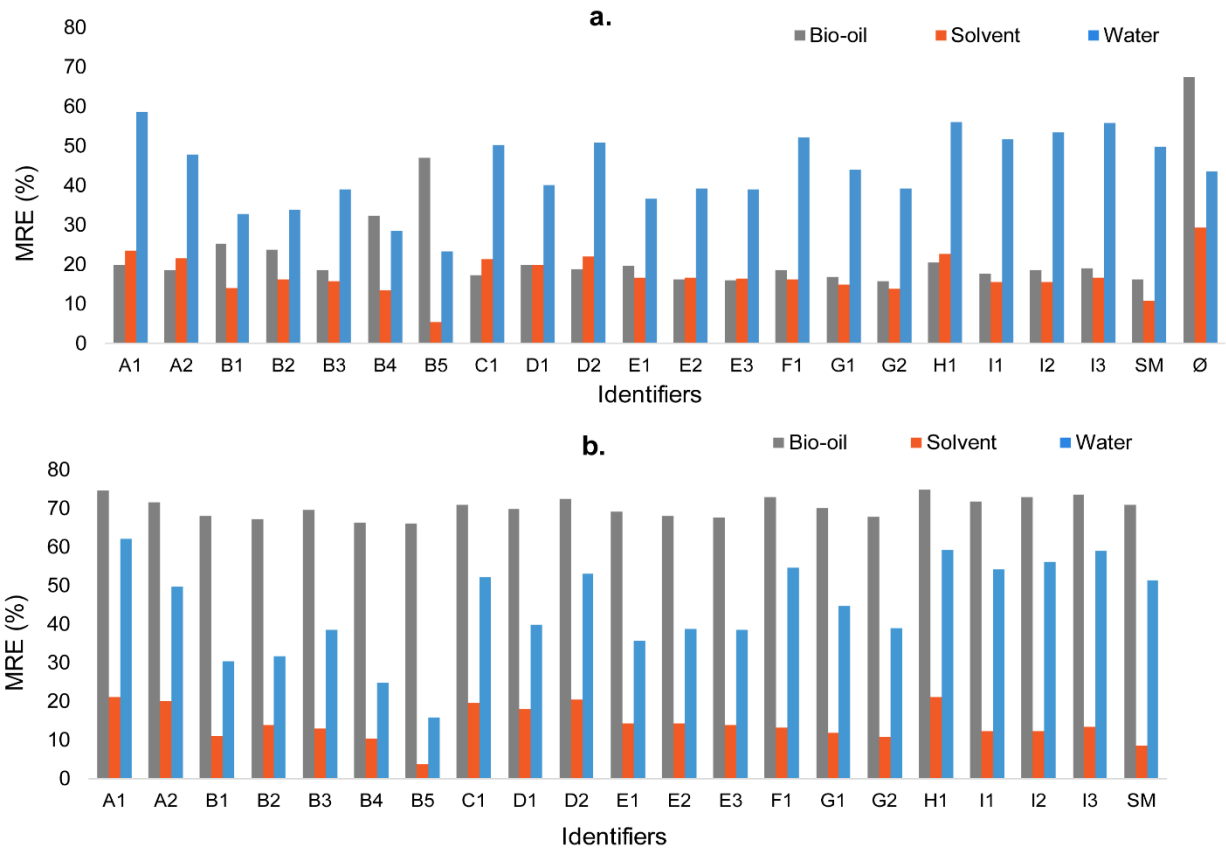


Figure 34. MRE values for the prediction of the K_{OW} for the main factions within the system WE: Bio-oil, solvent, and water for a) Case 1 and b) Case 2.

It is important to keep in mind that the fraction of pyrolignin in the bio-oil was determined by cold water precipitation [133]. This means the unknowns present in the base reports of FPBO composition correspond to a water-soluble fraction (sugars, polar aromatics, oxygenates of various natures) that is either not present in the database or not volatile at GC/MS operation ranges [84,141] and therefore could not be identified and quantified in the analysis reported by Han et al. [133] This means that Case 2 (Figure 34 b) may incur an implicit error by including heavy and water-soluble molecules in the pyrolignin fraction, which may help explain the more pronounced errors observed when using Case 2.

The K_{OW} of the *Bio-oil* fraction is increasing when adding water and decreasing when solvent is added (Figure S6 in the SI). This can be explained by the polar nature of the solvent (n-butanol), which is expected to partially distribute into the aqueous phase.

While Case 1 outperformed Case 2 for bio-oil prediction, both methods yielded comparable results for the water fraction (Figure 34). Additionally, the model consistently struggled to accurately predict the K_{OW} of the water fraction (Figure S7), regardless of the chosen case (MRE: 23-58% for Case 1, 15-65% for Case 2). The authors suggest that limitations in the chosen model molecules for sugar oligomers might contribute to this discrepancy. Overall, the best results were found for B5 (MRE: 23%) and the worst for A1 (MRE: 58%), while the null case (∅) results in an MRE of 43%.

When discussing the K_{OW} of the solvent (n-butanol), unlike for the bio-oil, the B5 structure showed the best approximation to the experimental data (MRE: 5.4%). The null case (∅) also performed the worst in this case (MRE: 29.1%). Interestingly, for all model molecules within Case 2 (Figure 34 b), better predictions were found for the solvent fraction when compared to Case 1 (Figure 34 a). The resulting trend shows that the K_{OW} of n-butanol initially increases and tends to become constant as more water or solvent is added (Figure S8).

Regarding the modeled pyrolignin fraction (Figure S9–Figure S11), the K_{ow} varied according to each evaluated structure. All structures showed an affinity for the organic phase as expected, with K_{ow} values ranging from 2 to approximately 966,000,000. This wide range indicates the diversity of the structures, with some being entirely transferred to the organic phase (K_{ow} >> 10,000) such as A1, H1, G1, F1, I1 and I3. In the range of 100 to 1,000, showing slightly less refinement, are A2, E2, E3, and D2. The range of 10 to 100 includes B1, B2, E1, and D1. The molecules with the lowest affinity for the organic phase, with K_{ow} < 10, are B3, B4, and B5. Although these latter molecules prefer the organic phase, their polarity allows for some low concentration transfer to the aqueous phase.

5.3.2 Evaluation of relevant individual species

Figure 35 presents a comparison between the predicted K_{OW} and the experimental value for three relevant species in FPBO: 2-cyclopenten-1-one, guaiacol, and acetic acid. Experimentally, the former presented K_{OW} values between 1.5 and 2.6, lowering for increasing amounts of solvent and water, revealing a moderate affinity to the organic phase (Figure 35 a-d). Case 1 using the WE system (Figure 35 b) presents the best predictions overall, with clear outliers when using E3 and G2. For the remaining cases, the predicted trends plateau and do

not accompany the trend of the experimental data. Overall, best results are found for D2, F1, H1, I2 and I3 (MRE: 6.6%–7.4%) while the worse are found for B4, E2 and E3 (MRE: ~ 35%).

The variation with the experimental K_{OW} values for guaiacol do not follow clear trends (Figure 35 e-h), indicating possible issues with the experimental data, but are overall increasingly positive (K_{OW} 7.5-22.5) for increasing additions of solvent and water, for both cases. In terms of prediction, the K_{OW} of this molecule is consistently underestimated for all cases except Case 1 $WE_{1,2}$. Overall, the best results are found for D2, H1, A2 and A1 MRE: 23.9% - 25.8%) while the worse are found for G2 and B5 (MRE: 28% and 33 %, respectively).

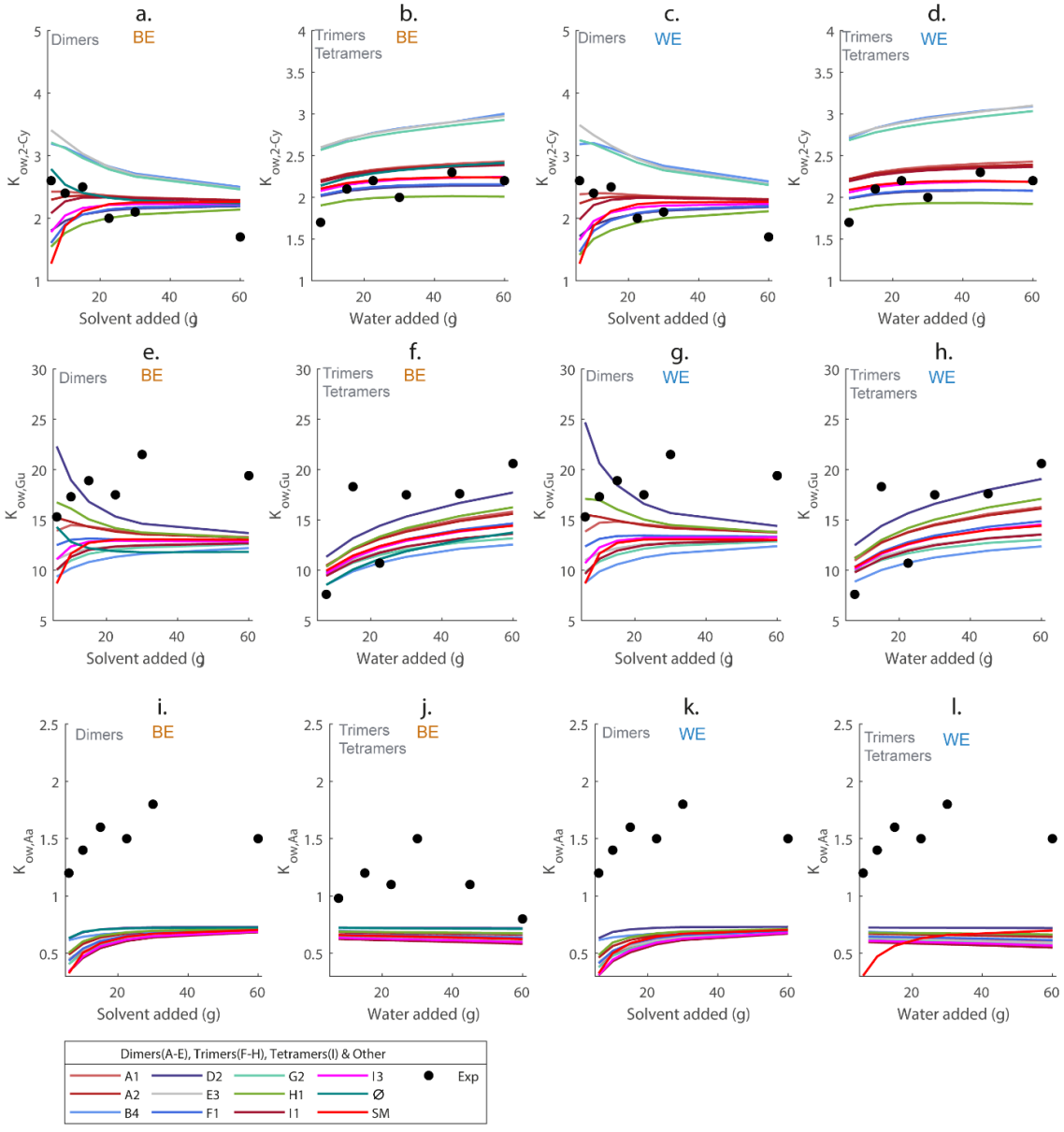


Figure 35. Variation of the predicted K_{OW} for 2-cyclopenten-1-one: (a - b) case 1, (c - d) case 2; guaiacol: (e-f) case1, (g-h) case 2; and acetic acid: (i- j) case 1, and (k-l) case 2.

Regarding acetic acid, the experimental K_{OW} values reflect a better distribution between the organic and aqueous phases (K_{OW} 0.75-1.75), with an even less clear trend than the one seen for guaiacol, due to the high affinity to water at high water additions (Figure 35 j, WE_6). The predictions overwhelmingly underestimate the K_{OW} of acetic acid, estimating an unrealistic concentration in the aqueous phase, which may be attributed to shortcomings of the model (simple decanter) or the UNIFAC-DMD model. Overall, the best results are found for D2 and B1 (MRE: 34.2 % and 34.9%, respectively) while the worse are found for C1 and I1 (MRE: ~ 43.5%).

Figure 36 mirrors the analysis presented in Figure 35, but focusing on levoglucosan. Experimental results indicate that this sugar distributes preferentially to the aqueous phase (K_{OW} ~ 0.2). Interestingly, the K_{OW} tends to further favor the aqueous phase for increasingly higher additions of water and n-butanol. This trend seems to be well predicted in the model, expect for some clear outliers, namely B4 and G2 (MRE: 75.2% and *62*.6% respectively), while the best performants were I3 and I2 (MRE: 8.7% and 10.5%, respectively). The MREs were slightly larger for Case 2. Curves for all molecules can be found in the SI (Figure S12).

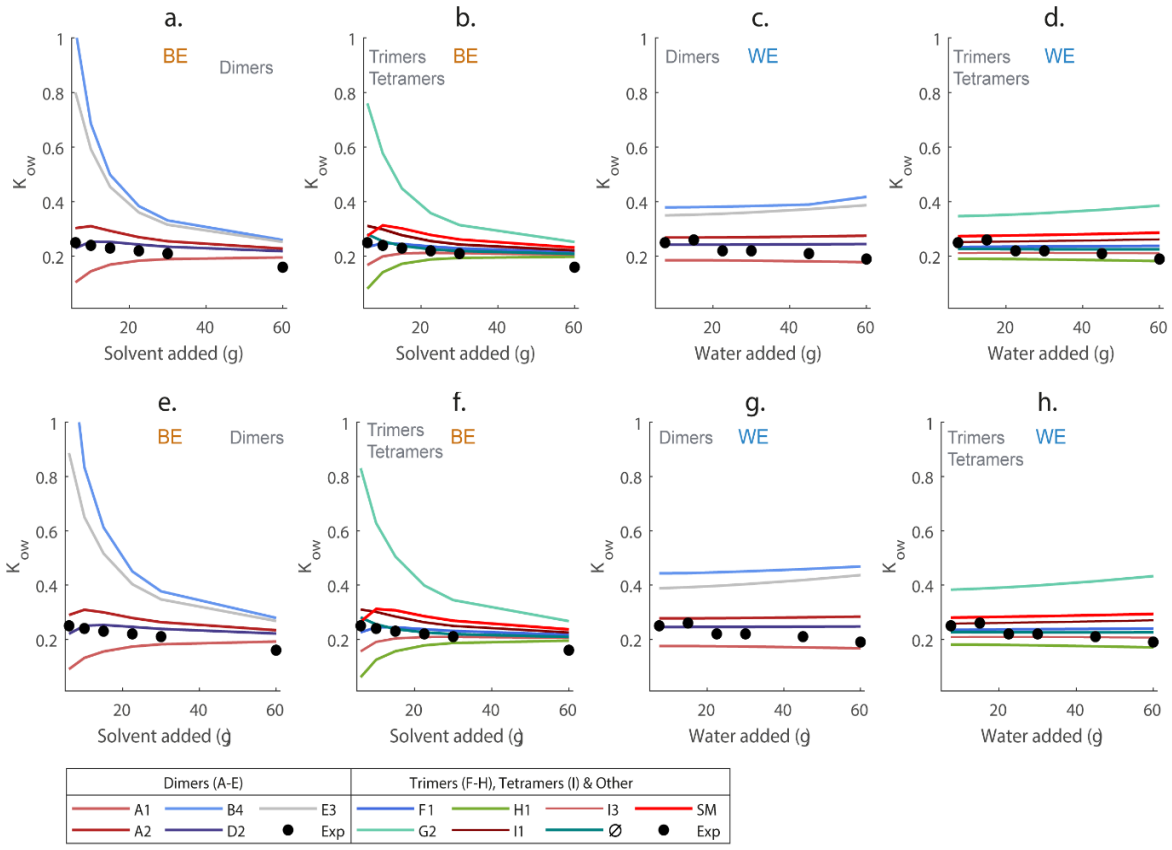


Figure 36. Evolution of the experimental and predicted partition coefficients for levoglucosan with added solvent (BE) and water (WE). Case 1: a-d, and Case 2: e-h.

In Figure 37, one can see the trends presents in Figure 36 in bar graph form, allowing for a more in depth analysis of the performance of each pyrolignin representative. Greatly contrasting with the trends observed in Figure 33, Figure 34, and Figure 35, the best results can be found in WE system for the null case (∅) and the larger representatives SM, I2, and I3 (MRE <10%), as well as some dimers (D2) and trimers (F1) (MRE: 11.2% - 12.6%). The MRE values were slightly higher for the BE system. This may be explained by their complexity and number of different interaction groups present, which better model potential interactions between levoglucosan and pyrolignin in solution.

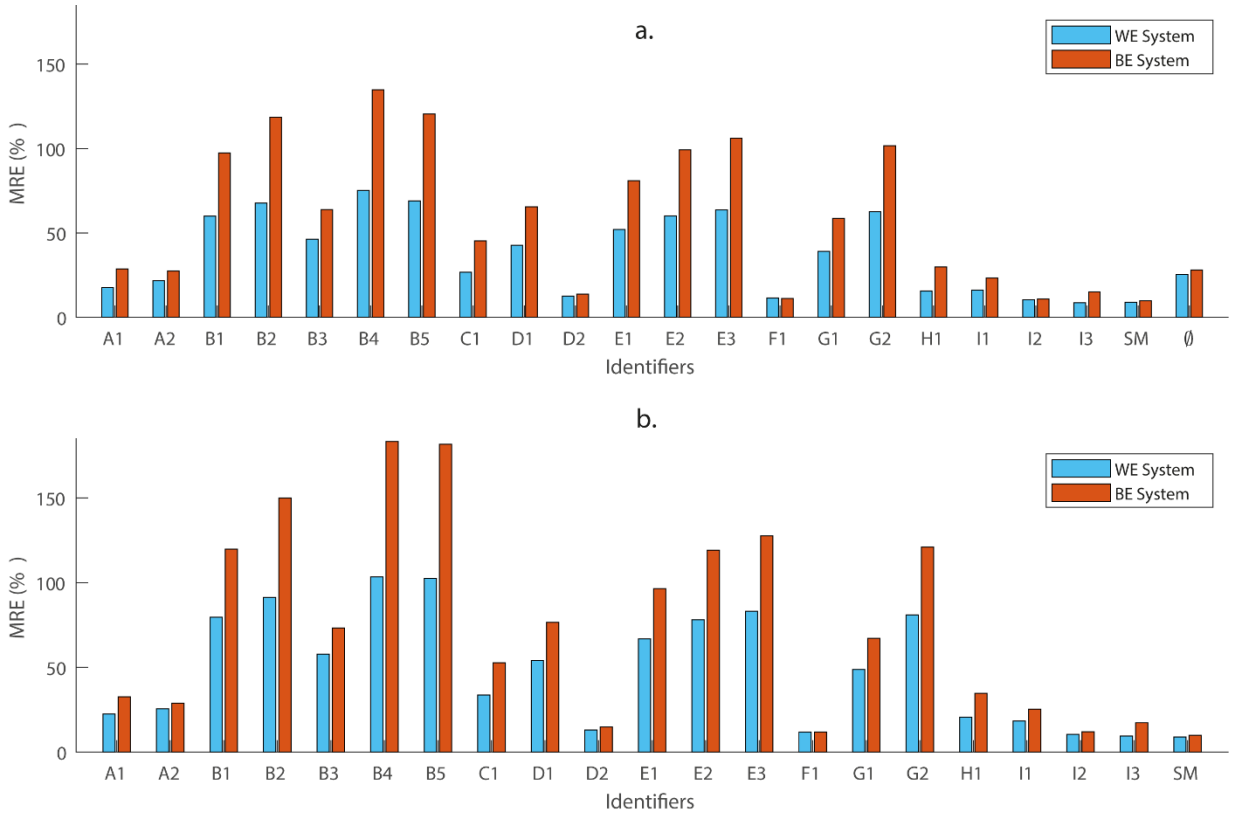


Figure 37. Bar diagrams of MRE values for K_{OW} prediction for levoglucosan for a) Case 1 and b) Case 2.

It is important to highlight that most of the surrogates evaluated in this study successfully captured the trend of both systems (WE and BE), demonstrating a good approximation to the real system and paving the way for a better understanding of the chemical composition of FPBO and the potential of extraction. Similar results, capturing the trends in levoglucosan phase distribution, have been found by Parku et al. [75] using conventional water extraction of this compound from already condensed FPBOs, who also reported considerable errors when comparing to modeling results also using UNIFAC-DMD (RMSD ≈ 22%). However, it is crucial to consider that the present results have been compared only with a single set of

experimental data. This precision could be significantly improved with new experiments and measurements, as well as by considering other sugar oligomer structures to better represent the polar fraction of the system. Candidate sugar oligomers for evaluation include those proposed by Denson et al. [141] which are from monomers until tetramers.

The pyrolignin representatives from group A (dimers, Table 21) present good results (MRE < 21%), with slight variations between the two group members. The benzofuran ring in A1 contributes to hydrophobic interactions due to its higher affinity, leading to phase separation when mixed with polar solvents such as water, and preferring the organic phase ($K_{OW,A1}$ > 1000). Partial saturation of the dihydrobenzofuran ring in A2 increases the polarity ($126 \geq K_{OW,A2} \leq 328$), which may reduce the tendency for phase separation in polar solvents compared to A1. In comparison, group B representatives (methoxy-substituted dimers, Table 21) performed markedly worse (MRE > 45%), but captured the trends observed in Figure 36.

Group D is composed of two molecules with similar bonds (β-O-4 dimers, Table 21 but which resulted in very different behaviors in the simulation. The MREs (≥ 50%) observed for D1 may be attributed to due to conformational configurations of these molecules that indubitably affect solubility in either phase, especially in an acid environment such as FPBO for which the supplier (BTG) reports a general low pH-value of around 2.5-3. This factor is most likely not well reflected when using the UNIFAC-DMD model. On the other hand, D2 is one of the best performing representatives for the prediction of the partition of levoglucosan (MRE <14%), which may be due to the presence of four methoxy groups and resonance effects between the aryl groups, and possible enol conformations in the hydroxy of the β-O-4 bond. Similar resonance effects are also heavily featured in the structure of F1 (β-O-4 trimer, Table 21), that performs even better than D1.

The difference in performance between G1 and G2 (MRE: <59% and >60%, respectively) within group G (stilbene β-5 trimers, Table 21) further clarify shortcomings in the methodology, namely the inability of the software and the thermodynamic model to account for resonance effects, especially when the difference between the molecules in the presence of a methoxy-hydroxy substituted ring compared to a quinone ring. H1 (saturated intermediate ring, double-ring bond, Table 21) performs comparably with group A, I1 and SM, and may be explained by the saturation of the intermediate ring and the presence of methoxy substituents in the aromatic rings.

Group I (Table 21) features tetramers with different terminals and interbonding, but most rings feature methoxy substituents. The best performance between these is found for I2, which

may be attributed to only employing stilbene β-5 bonds, which are desaturated, compared to other bonds employed by I1 and I3.

Upon reviewing probable trends or correlations between certain characteristics of the evaluated molecules, a considerable correlation was found between the presence of methoxy groups (-OCH_3) in the molecular structure and the accuracy of $K_{OW.LVG}$ prediction with R^2 values ranging between 0.61 - 0.68. The size of the molecules showed only a slight correlation (R^2: 0.35 - 0.37), while the number and variety of functional groups present did not exhibit significant trends (Figure S13). Therefore, it can be hypothesized that structures with greater diversity and quantity of functional groups, especially methoxy groups, would be appropriate for modeling the liquid-liquid phase equilibrium when seeking to extract levoglucosan. This may also be favorable for other sugars of interest.

It is important to highlight that most of the surrogates evaluated in this study successfully captured the trend of both systems (WE and BE), demonstrating a good approximation to the real system and paving the way for a better understanding of the chemical composition of FPBO and the potential of extraction. It is important to acknowledge the inherent limitations of modeling techniques and potential for errors (~30%), which were significantly improved with several structures, resulting in errors of less than 10%. Similar results, capturing the trends in levoglucosan phase distribution, have been found by Parku et al. [75] using conventional water extraction from already condensed FPBOs, who also reported considerable errors when comparing to modeling results also using UNIFAC-DMD (RMSD ≈ 22%). However, it is crucial to consider that both the results presented in this work and the work by Parku et al. [75] have been compared only with a single set of experimental data. This precision could be significantly improved with new experiments and measurements, as well as by considering other sugar oligomer structures to better represent the polar fraction of the system. Candidate sugar oligomers for evaluation include those proposed by Denson et al. [141] (monomers to tetramers).

5.3.3 Evaluation of a pyrolignin representative mixture

As can be observed in Figure 36, the pyrolignin representatives A1 and G2 presented contrasting trends for BE system. It was proposed that mixtures of these two molecules may give room for optimization as representative of pyrolignin to more accurately predict the K_{OW} of levoglucosan. Mixtures named M1-9 have been tested, corresponding to increasing mass fractions of A1 versus G2 (M1:10%/90%, M2:20%/80%, M3: 30%/70%, M4:40%/60%, M5:50%/50%, M6:60%/40%, M7:70%/30%, M8:80%/20%, and M9:90%/10%). The results

indicate only minor improvements of the MRE of the ratio O/A (MRE: BE 28.4% to 34.6%, WE 43.2% to 48.0%) when compared to the studies performed with either A1 (MRE: BE 35.1%, WE 48.6%) or G2 (MRE: BE 28.1%, WE 42.4%).

The distribution of the *bio-oil*, *solvent*, *and water* fractions was improved by introducing mixtures M1-M9. Significant enhancement was achieved in the phase separation Ratio OA in the WE system, reducing the MRE from 42% and 48% (for A1 and G2 individually) to 28% with M8 and M9. However, no improvements were observed in the BE system. It was also possible to slightly improve the K_{ow} prediction of the bio-oil and the solvent (less than a 3% decrease), with the water fraction showing better results, achieving a 7% decrease.

Figure 38 presents the evolution of the K_{OW} of levoglucosan when employing the M1-9 as pyrolignin representatives. Despite very low MREs (< 10%) when using these mixtures, several of them overestimate the $BE_{1\text{-}5}$ points, while all of them overestimate BE_6, with parallel plateaus. The trends for WE are better, also with parallel plateaus. The best approximations were found for the mixture M2 with MRE of 8.8%, while increasingly worse results are observed for higher fractions of G2.

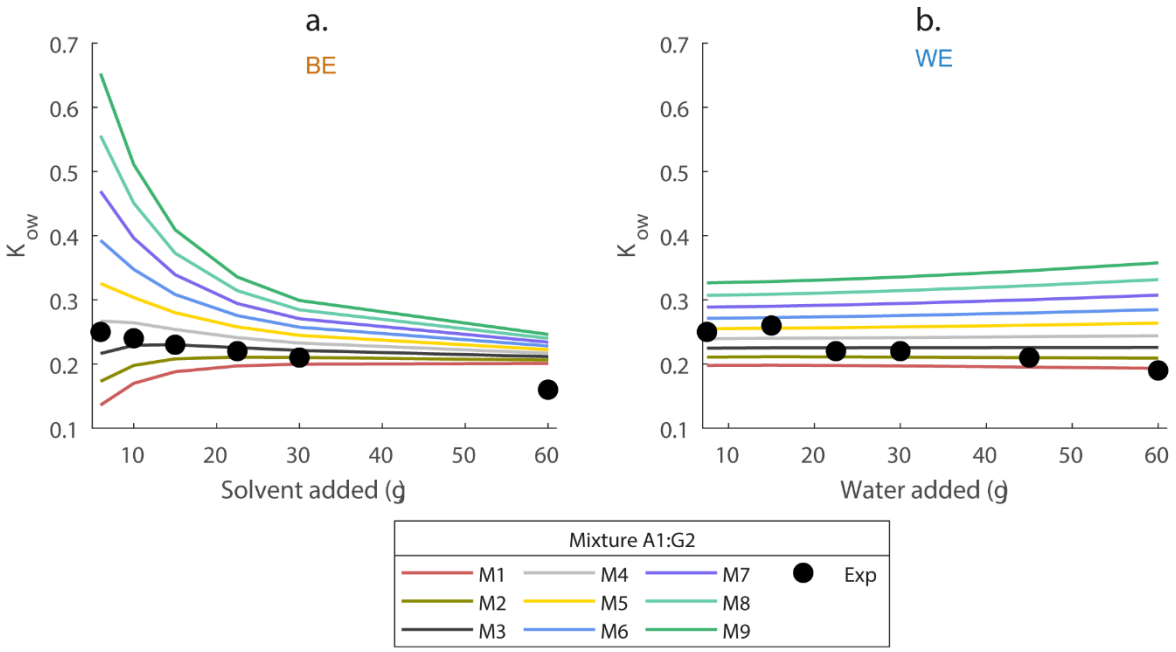


Figure 38. Evolution of the experimental and estimated value for the K_{OW} of levoglucosan for the mixtures M1-M9. a. BE system, b. WE system.

This study suggests that mixing two molecules (dimer and trimer) to represent the pyrolignin fraction improves the results because it captures a broader range of molecular interactions and physicochemical properties present in the pyrolignin fraction of FPBO. Different molecules

interact with solvents and other components in unique ways, and combining two molecules with opposite tendencies (in K_{ow} levoglucosan prediction) allows the model to simulate these complex interactions more accurately. Complementary properties of different molecules, such as different solubility in the organic and aqueous phases, help achieve a balanced representation.

However, this mixture did not produce significant improvements for other compounds; for instance, for 2-cyclopenten-1-one, the best mixture was M1 (MRE: 12.8%), while A1 individually had an MRE of 12.9% and G2 had 30% (for the BE system). Similarly, no significant improvements were observed for guaiacol, acetic acid, and other compounds. This outcome is likely because the selection of structures for mixing was based on the trends of K_{ow} for levoglucosan rather than for other compounds.

While the model represents an important first step in describing liquid-liquid phase equilibria for FPBO, further details need to be clarified in order to achieve a model that is suitable from an engineering perspective. This primarily involves the availability of thermophysical data of individual compounds and other relevant parameters to run phase equilibria models since this limits the definition of a surrogate mixture and consequently the quality of prediction. From this study it becomes also clear that experimental data for phase equilibria, specifically in the domain of LLE, is very scarce and further efforts fur high quality data are desperately needed to enable detailed optimization of phase equilibrium calculations.

Future work must focus on applying such tests, both experimentally and modeling-wise, to broadly different bio-oils obtained from different feedstocks and process conditions, in order to estimate the level of error expected from this method. It is also recommended that the modeling refines the range of possible representatives for the pyrolytic lignin fraction, but also employs representatives for sugar oligomers (e.g., humins), to further improve the accuracy of phase separation predictions.

5.4 Conclusions

This work presents a first approach to predict liquid-liquid phase equilibria during extraction of FPBO with n-butanol and water using the UNIFAC-Dortmund model. With a carefully chosen surrogate mixture for the FPBO, lumped parameters like the mass ratio of organic to aqueous fraction (O/A) and partition coefficients of individual species are predicted with a reasonable error. Representing the largely unknown oligomers in FPBO with a lignin-derived representative compound (‘pyrolignin’) achieves an MRE of around 27 % for the ratio O/A and is an important improvement as compared to modeling phase equilibria without a pyrolignin

representation (MRE > 57%). The K_{OW} of the solvent n-butanol was generally predicted better than that of water. The choice of the pyrolignin representative affects the quality of prediction for the different partition coefficients in opposing ways (i.e., improving one while worsening others), giving little general preference for one representative over the other. Interestingly, 3,4,4'-biphenyltriol, which has also been shown to be a promising representative for pyrolignin for vapour-liquid phase equilibria calculations, is among the slightly better performing representatives. This indicates that a unified surrogate mixture might become reasonable to model both liquid-liquid as well as vapor liquid phase equilibria (VLE).

Similar to VLE prediction, the deviation from experimental results can be very different for individual compounds in the bio-oil and is also strongly affected by the choice of pyrolignin representative. For example, the general trend of the partition coefficient for levoglucosan is captured very well, which decreases with the addition of n-butanol and increases with water addition. For this specific case, the surrogate mixture without pyrolignin surprisingly provided a very good prediction (MRE <10%) which was also achieved with some pyrolignin representatives (an alkyl-aryl ether β-O-4 trimer and mixed tetramers, all with MRE values between 8 to 12%). This is an indication that sugar oligomers might play an important role during LLE and that the model could be improved by considering representatives for them.

Different to modelling VLE, some pyrolignin representatives show opposing trends when compared to experimental data, allowing for optimization with a mixture of representatives. E.g., combining a dimer (phenylcoumaran β-5) and a trimer (stilbene β-5) significantly improved the model's prediction for levoglucosan extraction with an MRE <10% for the K_{OW} of levoglucosan in both the n-butanol and water extraction. This indicates that carefully curated mixtures of surrogates can enhance the accuracy of phase separation predictions.

Chapter 6. Conclusions and Outlook

This chapter summarizes the main findings obtained throughout the work and provides a critical assessment of the achieved outcomes. The discussion is divided into three parts corresponding to the main research themes: (i) fast pyrolysis of lignin, (ii) analytical characterization of lignin-derived FPBOs, and (iii) thermodynamic modeling of phase equilibria. Finally, an outlook highlights perspectives for future research and potential developments arising from the present results.

Conclusions

Lignin Pyrolysis

The pilot-scale fast pyrolysis experiments demonstrated the technical feasibility of processing different lignin types in a twin-screw reactor. Two lignins with distinct origins and properties were investigated: Indulin AT, a Kraft lignin, and Miscancell, a *Miscanthus*-derived alkali lignin. The experiments revealed that lignin type has a pronounced influence on both operability and product composition.

For Indulin AT, the original reactor configuration caused severe agglomeration within the heated section, preventing continuous operation. The reversal of the inlet configuration, where the heat carrier was introduced before the feed, successfully eliminated blockages inside the reactor, confirming that the strong mechanical forces of the twin-screw reactor effectively disperse molten lignin. However, under these new conditions, Indulin AT showed persistent deposition and clogging in the feeding drop pipe, shifting the operational limitation to the feeding section. In contrast, the Miscancell lignin could be processed smoothly, yielding a stable operation without any issues. Its lower tendency to melt and form agglomerates highlights the importance of feedstock properties for ensuring continuous processing.

Product analysis showed the expected bio-oil composition, with FPBOs rich in phenols and oxygenated phenols. The Kraft lignin bio-oil was dominated by guaiacyl-type compounds, whereas the Miscancell lignin bio-oil exhibited a broader composition including syringyl-type species, reflecting the nature of the original biomasses. In addition, the Miscancell lignin FPBO also showed the presence of sugar species, likely related to residual cellulose or hemicellulose impurities. The good agreement between pilot-scale and micro-pyrolysis data confirmed the reliability of the reactor and experimental approach. Overall, the study demonstrated that twin-screw reactor configuration tested is suitable for lignin fast pyrolysis, provided that the feeding section is adapted to the melting behavior of the feedstock.

Lignin Bio-Oil Characterization

Comprehensive characterization of the lignin-derived FPBOs was achieved through the combination of FTIR, GC–MS, 2D-HSQC NMR, and UHRMS, offering complementary perspectives on functional groups, structural motifs, and molecular distributions. The alkali lignin bio-oil from *Miscanthus* showed higher chemical diversity, with more hydroxyl and carbonyl functionalities and greater aliphatic content, while the Kraft lignin bio-oil exhibited higher aromatic condensation and oxygenated phenols.

UHRMS analysis provided crucial molecular-level insights into the compositional complexity of the bio-oils. Thousands of individual molecular formulas were assigned, revealing similar oxygenated class distributions (mainly O_2–O_6) for both samples, but with clear differences in molecular weight and structural patterns. The alkali lignin bio-oil contained a higher proportion of lighter, oxygen-rich dimers and trimers, whereas the Kraft lignin bio-oil showed more condensed oligomer species extending up to tetramers. These results confirmed the presence of heavy oxygenated compounds in both oils and reflected the higher structural condensation of the Kraft bio-oil sample.

Building upon these findings, ^{1}H-NMR spectral modeling was performed to connect the analytical characterization with a molecular representation of the bio-oil. A surrogate mixture to represent the bio-oil was composed of GC-detectable volatiles and synthetic lignin oligomers (SLOs) to describe the unknown heavy fraction. The selected SLOs, for which NMR data were available from a literature database, were used to account for the spectral contribution of the heavy GC-invisible compounds. The optimized surrogate composition, in which the SLOs consisted exclusively of dimers and trimers, successfully reproduced the experimental spectrum of the alkali lignin bio-oil, indicating that similar oligomer structures are likely present in the FPBO.

Taken together, the analytical results advanced the understanding of lignin derived FPBOs by linking measurable molecular features with the unresolved heavy fraction, and showing that SLOs can help uncover the structures hidden within it.

Thermodynamic Modeling

The thermodynamic modeling work comprised two complementary studies addressing the vapor–liquid (VLE) and liquid–liquid (LLE) equilibrium behavior of fast pyrolysis bio-oils. In both cases, the complex FPBO matrix was represented by a surrogate mixture composed of identifiable and quantifiable volatile compounds combined with representative structures for the unknown heavy fraction. Different lignin oligomers were systematically evaluated as

potential surrogates in order to assess their influence on phase-equilibria predictions and to identify suitable molecular candidates for modeling the pyrolytic lignin fraction.

For the VLE study, advanced distillation curve (ADC) experiments were performed using the lignin derived FPBO to evaluate the VLE behavior and better assess the contribution of the heavy fraction during distillation. Simulation of the experimental results showed that including a pyrolytic lignin (PL) surrogate in the surrogate mixture significantly improved predictive accuracy compared to calculations without a heavy component. Among the tested candidates, biphenyl-type dimers (D-B group) provided the best overall agreement with experimental distillation data, followed by β-O-4-type dimers, particularly for the PR-BM model. Dimers in general outperformed larger oligomers such as trimers and tetramers, which often caused instability or higher deviations in the simulations. Despite its simplicity, the IDEAL model performed surprisingly well, yielding results comparable to those of the UNIFAC-DMD model, which gave the most accurate composition predictions and an activity coefficient for water consistent with experimental data. In contrast, the PR-BM model showed the largest deviations and overestimated the water activity coefficient, indicating its limited suitability for modeling the phase behavior of these complex systems.

For the LLE study, the surrogate approach was applied to describe the liquid–liquid extraction of FPBO with n-butanol and water, using literature data as a reference. The calculations were performed with the UNIFAC-DMD model, which was selected for its ability to handle non-ideal systems. As in the VLE study, introducing a representative lignin oligomer surrogate for the heavy fraction improved the accuracy of the predicted phase separation, reducing the mean relative error of the organic-to-aqueous mass ratio from over 57 % to 27 %. The biphenyl-type dimers again showed good performance for describing the overall phase split, in line with the VLE results. However, when considering the partitioning of individual compounds, particularly levoglucosan, better agreement with experimental data was achieved using β-O-4-type dimers, a trimer, or a tetramer.

Overall, the results from both studies showed that biphenyl-type dimers are the most reliable representatives for modeling phase equilibria of FPBOs, particularly for predicting bulk properties. At the same time, when targeting compound specific interactions or extraction behavior, other molecular structures may provide additional accuracy. In addition, the UNIFAC-DMD model consistently provided the accurate predictions, confirming its suitability for describing the non-ideal behavior of these systems. These results indicate that optimal surrogate selection depends on the targeted property, and that combining molecules with complementary characteristics could further improve model performance.

Outlook

This work established a framework combining experimental, analytical, and modeling approaches to better understand and predict the behavior of lignin oligomers in FPBOs. Building upon these findings, several directions for future work can be identified to further improve its research.

For future lignin fast pyrolysis experiments, efforts should focus on improving the feeding section for lignins with severe melting behavior, such as Kraft lignin. While the modified reactor configuration successfully prevented agglomeration within the reaction zone, the issue was transferred to the feed inlet. Developing reliable feeding strategies is therefore essential to enable continuous operation of the pilot scale twin-screw reactor unit at IKFT. Approaches such as cooled screw feeders and slug injectors could be explored and adapted to the unit for the processing of different lignins.

For bio-oil characterization, future work could extend the current approach to include ^{13}C-NMR data, since these spectra are already available in existing databases and can complement the ^{1}H-NMR modeling. Expanding these databases with additional synthetic lignin oligomers and corresponding reference spectra would further improve coverage and reliability. These molecules could also be analyzed using complementary techniques such as UHRMS, FTIR, and GC-MS to generate cross-method reference data. Finally, developing methods to isolate similar oligomer structures directly from FPBO would allow their experimental validation through comparison with analytical and modeled results.

For thermodynamic modeling, future work should explore a broader range of surrogate molecules to represent the heavy fraction and evaluate their performance across different phase-equilibrium conditions. Testing mixtures of surrogates rather than single representatives could also provide a more realistic description of FPBO composition. Also, a deeper understanding of the physicochemical properties of these molecules is essential, as most current parameters are derived from estimation methods. Determining or calculating key properties such as vapor pressure, heat capacity, and binary interaction parameters would significantly improve model reliability. Different estimation and group-contribution methods could be compared for this purpose, while ab initio calculations may help elucidate molecular level interactions that define these properties. In addition, synthetic lignin oligomers (SLOs) could serve as model compounds for experimental determination of such parameters, providing valuable reference data to validate and refine the thermodynamic models.

References

[1] J.L. Holechek, H.M.E. Geli, M.N. Sawalhah, R. Valdez, A Global Assessment: Can Renewable Energy Replace Fossil Fuels by 2050?, Sustainability 14 (2022) 4792. https://doi.org/10.3390/su14084792.

[2] M. Filonchyk, M.P. Peterson, L. Zhang, V. Hurynovich, Y. He, Greenhouse gases emissions and global climate change: Examining the influence of CO2, CH4, and N2O, Science of The Total Environment 935 (2024) 173359. https://doi.org/10.1016/j.scitotenv.2024.173359.

[3] J. Delbeke, A. Runge-Metzger, Y. Slingenberg, J. Werksman, The Paris Agreement, (2019). https://doi.org/10.4324/9789276082569-2.

[4] European Commission. Directorate General for Communication., The European Green Deal: delivering the EU's 2030 climate targets., Publications Office, LU, 2023. https://data.europa.eu/doi/10.2775/783179 (accessed September 30, 2025).

[5] G. Centi, G. Iaquaniello, S. Perathoner, Chemical engineering role in the use of renewable energy and alternative carbon sources in chemical production, BMC Chem Eng 1 (2019) 5. https://doi.org/10.1186/s42480-019-0006-8.

[6] M.-A. Perea-Moreno, E. Samerón-Manzano, A.-J. Perea-Moreno, Biomass as Renewable Energy: Worldwide Research Trends, Sustainability 11 (2019) 863. https://doi.org/10.3390/su11030863.

[7] R.T. Alarcon, K.J. Lamb, G. Bannach, M. North, Opportunities for the Use of Brazilian Biomass to Produce Renewable Chemicals and Materials, ChemSusChem 14 (2021) 169–188. https://doi.org/10.1002/cssc.202001726.

[8] R.S. Abolore, S. Jaiswal, A.K. Jaiswal, Green and sustainable pretreatment methods for cellulose extraction from lignocellulosic biomass and its applications: A review, Carbohydrate Polymer Technologies and Applications 7 (2024) 100396. https://doi.org/10.1016/j.carpta.2023.100396.

[9] V. Sharma, M.-L. Tsai, P. Nargotra, C.-W. Chen, P.-P. Sun, R.R. Singhania, A.K. Patel, C.-D. Dong, Journey of lignin from a roadblock to bridge for lignocellulose biorefineries: A comprehensive review, Science of The Total Environment 861 (2023) 160560. https://doi.org/10.1016/j.scitotenv.2022.160560.

[10] C.C. Zandonadi Nunes, H.B. De Paula, I.F. Demuner, M.O. De Paula, L.G. Pedroti, A.M.M.L. Carvalho, Kraft lignin biorefinery: from pulping side streams to concrete plasticizers, Eur. J. Wood Prod. 82 (2024) 849–860. https://doi.org/10.1007/s00107-024-02044-8.

[11] S. Suresh, V. Viswanathan, M. Angamuthu, G.P. Dhakshinamoorthy, K.P. Gopinath, A. Bhatnagar, Lignin waste processing into solid, liquid, and gaseous fuels: a comprehensive review, Biomass Conv. Bioref. 13 (2023) 4515–4553. https://doi.org/10.1007/s13399-021-01497-8.

[12] M. Ullah, P. Liu, S. Xie, S. Sun, Recent Advancements and Challenges in Lignin Valorization: Green Routes towards Sustainable Bioproducts, Molecules 27 (2022) 6055. https://doi.org/10.3390/molecules27186055.

[13] A. Prakash, R. Singh, B. Balagurumurthy, T. Bhaskar, A.K. Arora, S.K. Puri, Thermochemical Valorization of Lignin, in: Recent Advances in Thermo-Chemical Conversion of Biomass, Elsevier, 2015: pp. 455–478. https://doi.org/10.1016/B978-0-444-63289-0.00016-8.

[14] R. Rinaldi, R. Jastrzebski, M.T. Clough, J. Ralph, M. Kennema, P.C.A. Bruijnincx, B.M. Weckhuysen, Paving the Way for Lignin Valorisation: Recent Advances in Bioengineering, Biorefining and Catalysis, Angew Chem Int Ed 55 (2016) 8164–8215. https://doi.org/10.1002/anie.201510351.

[15] B. Joffres, D. Laurenti, N. Charon, A. Daudin, A. Quignard, C. Geantet, Thermochemical Conversion of Lignin for Fuels and Chemicals: A Review, Oil Gas Sci. Technol. – Rev. IFP Energies Nouvelles 68 (2013) 753–763. https://doi.org/10.2516/ogst/2013132.
[16] A. Ekielski, P.K. Mishra, Lignin for Bioeconomy: The Present and Future Role of Technical Lignin, IJMS 22 (2020) 63. https://doi.org/10.3390/ijms22010063.
[17] Z. Strassberger, S. Tanase, G. Rothenberg, The pros and cons of lignin valorisation in an integrated biorefinery, RSC Adv. 4 (2014) 25310–25318. https://doi.org/10.1039/C4RA04747H.
[18] P. Azadi, O.R. Inderwildi, R. Farnood, D.A. King, Liquid fuels, hydrogen and chemicals from lignin: A critical review, Renewable and Sustainable Energy Reviews 21 (2013) 506–523. https://doi.org/10.1016/j.rser.2012.12.022.
[19] V.K. Thakur, M.K. Thakur, P. Raghavan, M.R. Kessler, Progress in Green Polymer Composites from Lignin for Multifunctional Applications: A Review, ACS Sustainable Chem. Eng. 2 (2014) 1072–1092. https://doi.org/10.1021/sc500087z.
[20] G. Gellerstedt, A. Majtnerova, L. Zhang, Towards a new concept of lignin condensation in kraft pulping. Initial results, Comptes Rendus. Biologies 327 (2004) 817–826. https://doi.org/10.1016/j.crvi.2004.03.011.
[21] L.A. Zevallos Torres, A. Lorenci Woiciechowski, V.O. De Andrade Tanobe, S.G. Karp, L.C. Guimarães Lorenci, C. Faulds, C.R. Soccol, Lignin as a potential source of high-added value compounds: A review, Journal of Cleaner Production 263 (2020) 121499. https://doi.org/10.1016/j.jclepro.2020.121499.
[22] M.J. De La Torre, A. Moral, M.D. Hernández, E. Cabeza, A. Tijero, Organosolv lignin for biofuel, Industrial Crops and Products 45 (2013) 58–63. https://doi.org/10.1016/j.indcrop.2012.12.002.
[23] F.F. Menezes, V.M. Nascimento, G.R. Gomes, George.J.M. Rocha, M. Strauss, T.L. Junqueira, C. Driemeier, Depolymerization of enzymatic hydrolysis lignin: Review of technologies and opportunities for research, Fuel 342 (2023) 127796. https://doi.org/10.1016/j.fuel.2023.127796.
[24] S. Kang, X. Li, J. Fan, J. Chang, Hydrothermal conversion of lignin: A review, Renewable and Sustainable Energy Reviews 27 (2013) 546–558. https://doi.org/10.1016/j.rser.2013.07.013.
[25] A.V. Bridgwater, Review of fast pyrolysis of biomass and product upgrading, Biomass and Bioenergy 38 (2012) 68–94. https://doi.org/10.1016/j.biombioe.2011.01.048.
[26] A.A. Shah, K. Sharma, M.S. Haider, S.S. Toor, L.A. Rosendahl, T.H. Pedersen, D. Castello, The Role of Catalysts in Biomass Hydrothermal Liquefaction and Biocrude Upgrading, Processes 10 (2022) 207. https://doi.org/10.3390/pr10020207.
[27] C. Chio, M. Sain, W. Qin, Lignin utilization: A review of lignin depolymerization from various aspects, Renewable and Sustainable Energy Reviews 107 (2019) 232–249. https://doi.org/10.1016/j.rser.2019.03.008.
[28] I. Graça, J.M. Lopes, H.S. Cerqueira, M.F. Ribeiro, Bio-oils Upgrading for Second Generation Biofuels, Ind. Eng. Chem. Res. 52 (2013) 275–287. https://doi.org/10.1021/ie301714x.
[29] H. Wang, Y. Pu, A. Ragauskas, B. Yang, From lignin to valuable products–strategies, challenges, and prospects, Bioresource Technology 271 (2019) 449–461. https://doi.org/10.1016/j.biortech.2018.09.072.
[30] K. Ye, Y. Liu, S. Wu, J. Zhuang, A review for lignin valorization: Challenges and perspectives in catalytic hydrogenolysis, Industrial Crops and Products 172 (2021) 114008. https://doi.org/10.1016/j.indcrop.2021.114008.
[31] D.D.S. Argyropoulos, C. Crestini, C. Dahlstrand, E. Furusjö, C. Gioia, K. Jedvert, G. Henriksson, C. Hulteberg, M. Lawoko, C. Pierrou, J.S.M. Samec, E. Subbotina, H.

Wallmo, M. Wimby, Kraft Lignin: A Valuable, Sustainable Resource, Opportunities and Challenges, ChemSusChem 16 (2023) e202300492. https://doi.org/10.1002/cssc.202300492.
[32] A. Aho, N. Kumar, K. Eränen, B. Holmbom, M. Hupa, T. Salmi, D.Yu. Murzin, Pyrolysis of Softwood Carbohydrates in a Fluidized Bed Reactor, IJMS 9 (2008) 1665–1675. https://doi.org/10.3390/ijms9091665.
[33] H. Kawamoto, Lignin pyrolysis reactions, J Wood Sci 63 (2017) 117–132. https://doi.org/10.1007/s10086-016-1606-z.
[34] D.J. Nowakowski, A.V. Bridgwater, D.C. Elliott, D. Meier, P. de Wild, Lignin fast pyrolysis: Results from an international collaboration, Journal of Analytical and Applied Pyrolysis 88 (2010) 53–72. https://doi.org/10.1016/j.jaap.2010.02.009.
[35] B. Shrestha, Y. Le Brech, T. Ghislain, S. Leclerc, V. Carré, F. Aubriet, S. Hoppe, P. Marchal, S. Pontvianne, N. Brosse, A. Dufour, A Multitechnique Characterization of Lignin Softening and Pyrolysis, ACS Sustainable Chem. Eng. 5 (2017) 6940–6949. https://doi.org/10.1021/acssuschemeng.7b01130.
[36] B. Shrestha, Y. Le Brech, T. Ghislain, S. Leclerc, V. Carré, F. Aubriet, S. Hoppe, P. Marchal, S. Pontvianne, N. Brosse, A. Dufour, A Multitechnique Characterization of Lignin Softening and Pyrolysis, ACS Sustainable Chem. Eng. 5 (2017) 6940–6949. https://doi.org/10.1021/acssuschemeng.7b01130.
[37] S. Zhou, R.C. Brown, X. Bai, The use of calcium hydroxide pretreatment to overcome agglomeration of technical lignin during fast pyrolysis, Green Chem. 17 (2015) 4748–4759. https://doi.org/10.1039/C5GC01611H.
[38] S. Ghysels, B. Dubuisson, M. Pala, L. Rohrbach, J. Van Den Bulcke, H.J. Heeres, F. Ronsse, Improving fast pyrolysis of lignin using three additives with different modes of action, Green Chem. 22 (2020) 6471–6488. https://doi.org/10.1039/D0GC02417A.
[39] D.J. Nowakowski, A.V. Bridgwater, D.C. Elliott, D. Meier, P. de Wild, Lignin fast pyrolysis: Results from an international collaboration, Journal of Analytical and Applied Pyrolysis 88 (2010) 53–72. https://doi.org/10.1016/j.jaap.2010.02.009.
[40] P.J. de Wild, W.J.J. Huijgen, H.J. Heeres, Pyrolysis of wheat straw-derived organosolv lignin, Journal of Analytical and Applied Pyrolysis 93 (2012) 95–103. https://doi.org/10.1016/j.jaap.2011.10.002.
[41] A. Tumbalam Gooty, D. Li, F. Berruti, C. Briens, Kraft-lignin pyrolysis and fractional condensation of its bio-oil vapors, Journal of Analytical and Applied Pyrolysis 106 (2014) 33–40. https://doi.org/10.1016/j.jaap.2013.12.006.
[42] D. Li, C. Briens, F. Berruti, Improved lignin pyrolysis for phenolics production in a bubbling bed reactor – Effect of bed materials, Bioresource Technology 189 (2015) 7–14. https://doi.org/10.1016/j.biortech.2015.04.004.
[43] A. Singh-Morgan, A. Puente-Urbina, J.A. van Bokhoven, Technology Overview of Fast Pyrolysis of Lignin: Current State and Potential for Scale-Up, ChemSusChem 15 (2022) e202200343. https://doi.org/10.1002/cssc.202200343.
[44] E. Pienihäkkinen, C. Lindfors, T. Ohra-aho, J. Lehtonen, T. Granström, M. Yamamoto, A. Oasmaa, Fast Pyrolysis of Hydrolysis Lignin in Fluidized Bed Reactors, Energy Fuels 35 (2021) 14758–14769. https://doi.org/10.1021/acs.energyfuels.1c01719.
[45] H. Shafaghat, P.S. Rezaei, D. Ro, J. Jae, B.-S. Kim, S.-C. Jung, B.H. Sung, Y.-K. Park, In-situ catalytic pyrolysis of lignin in a bench-scale fixed bed pyrolyzer, Journal of Industrial and Engineering Chemistry 54 (2017) 447–453. https://doi.org/10.1016/j.jiec.2017.06.026.
[46] I. Charisteidis, P. Lazaridis, A. Fotopoulos, E. Pachatouridou, L. Matsakas, U. Rova, P. Christakopoulos, K. Triantafyllidis, Catalytic Fast Pyrolysis of Lignin Isolated by Hybrid Organosolv—Steam Explosion Pretreatment of Hardwood and Softwood

Biomass for the Production of Phenolics and Aromatics, Catalysts 9 (2019) 935. https://doi.org/10.3390/catal9110935.
[47] T.N. Trinh, P.A. Jensen, Z. Sárossy, K. Dam-Johansen, N.O. Knudsen, H.R. Sørensen, H. Egsgaard, Fast Pyrolysis of Lignin Using a Pyrolysis Centrifuge Reactor, Energy Fuels 27 (2013) 3802–3810. https://doi.org/10.1021/ef400527k.
[48] S. Mukkamala, M.C. Wheeler, A.R.P. Van Heiningen, W.J. DeSisto, Formate-Assisted Fast Pyrolysis of Lignin, Energy Fuels 26 (2012) 1380–1384. https://doi.org/10.1021/ef201756a.
[49] F. Campuzano, R.C. Brown, J.D. Martínez, Auger reactors for pyrolysis of biomass and wastes, Renewable and Sustainable Energy Reviews 102 (2019) 372–409. https://doi.org/10.1016/j.rser.2018.12.014.
[50] A. Funke, M. Tomasi Morgano, N. Dahmen, H. Leibold, Experimental comparison of two bench scale units for fast and intermediate pyrolysis, Journal of Analytical and Applied Pyrolysis 124 (2017) 504–514. https://doi.org/10.1016/j.jaap.2016.12.033.
[51] S.-S. Liaw, Z. Wang, P. Ndegwa, C. Frear, S. Ha, C.-Z. Li, M. Garcia-Perez, Effect of pyrolysis temperature on the yield and properties of bio-oils obtained from the auger pyrolysis of Douglas Fir wood, Journal of Analytical and Applied Pyrolysis 93 (2012) 52–62. https://doi.org/10.1016/j.jaap.2011.09.011.
[52] A. Niebel, A. Funke, C. Pfitzer, N. Dahmen, N. Weih, D. Richter, B. Zimmerlin, Fast Pyrolysis of Wheat Straw—Improvements of Operational Stability in 10 Years of Bioliq Pilot Plant Operation, Energy Fuels 35 (2021) 11333–11345. https://doi.org/10.1021/acs.energyfuels.1c00851.
[53] N. Weih, A. Niebel, C. Pfitzer, A. Funke, G.K. Parku, N. Dahmen, Operational experience with miscanthus feedstock at the bioliq® fast pyrolysis plant, Journal of Analytical and Applied Pyrolysis 177 (2024) 106338. https://doi.org/10.1016/j.jaap.2023.106338.
[54] A. Funke, D. Richter, A. Niebel, N. Dahmen, J. Sauer, Fast Pyrolysis of Biomass Residues in a Twin-screw Mixing Reactor, JoVE (2016) 54395. https://doi.org/10.3791/54395.
[55] A. Oasmaa, Y. Solantausta, V. Arpiainen, E. Kuoppala, K. Sipilä, Fast Pyrolysis Bio-Oils from Wood and Agricultural Residues, Energy Fuels 24 (2010) 1380–1388. https://doi.org/10.1021/ef901107f.
[56] B. Scholze, C. Hanser, D. Meier, Characterization of the water-insoluble fraction from fast pyrolysis liquids (pyrolytic lignin), Journal of Analytical and Applied Pyrolysis 58–59 (2001) 387–400. https://doi.org/10.1016/S0165-2370(00)00173-X.
[57] A. Moutsoglou, B. Lawburgh, J. Lawburgh, Fractional condensation and aging of pyrolysis oil from softwood and organosolv lignin, Journal of Analytical and Applied Pyrolysis 135 (2018) 350–360. https://doi.org/10.1016/j.jaap.2018.08.016.
[58] L. Zhang, R. Liu, R. Yin, Y. Mei, Upgrading of bio-oil from biomass fast pyrolysis in China: A review, Renewable and Sustainable Energy Reviews 24 (2013) 66–72. https://doi.org/10.1016/j.rser.2013.03.027.
[59] M. Staš, D. Kubička, J. Chudoba, M. Pospíšil, Overview of Analytical Methods Used for Chemical Characterization of Pyrolysis Bio-oil, Energy Fuels 28 (2014) 385–402. https://doi.org/10.1021/ef402047y.
[60] W. Mu, H. Ben, A. Ragauskas, Y. Deng, Lignin Pyrolysis Components and Upgrading—Technology Review, Bioenerg. Res. 6 (2013) 1183–1204. https://doi.org/10.1007/s12155-013-9314-7.
[61] B. Scholze, D. Meier, Characterization of the water-insoluble fraction from pyrolysis oil (pyrolytic lignin). Part I. PY–GC/MS, FTIR, and functional groups, Journal of Analytical and Applied Pyrolysis 60 (2001) 41–54. https://doi.org/10.1016/S0165-2370(00)00110-8.

[62] T.M. Santos, W.R.D. Silva, J.D.C. Carregosa, A. Wisniewski, Comprehensive characterization of cattle manure bio-oil for scale-up assessment comparing non-equivalent reactor designs, Journal of Analytical and Applied Pyrolysis 162 (2022) 105465. https://doi.org/10.1016/j.jaap.2022.105465.
[63] W.R. Silva, T.M. Santos, J.C. Carregosa, C.C. Schmitt, K. Raffelt, N. Dahmen, A. Wisniewski, Bio-oil as a source of renewable chemicals: the chemistry of pyrolytic lignin, Biomass Conv. Bioref. (2024). https://doi.org/10.1007/s13399-024-05622-1.
[64] E. Terrell, V. Carré, A. Dufour, F. Aubriet, Y. Le Brech, M. Garcia-Pérez, Contributions to Lignomics: Stochastic Generation of Oligomeric Lignin Structures for Interpretation of MALDI–FT-ICR-MS Results, ChemSusChem 13 (2020) 4428–4445. https://doi.org/10.1002/cssc.202000239.
[65] Z. Zhou, L. Zhu, C. Cui, H. Liu, Y. Shen, W. Yuan, F. Qi, Pyrolysis of lignocellulosic biomass: Molecular-level insights with online ultrahigh-resolution mass spectrometry, Fuel Processing Technology 236 (2022) 107439. https://doi.org/10.1016/j.fuproc.2022.107439.
[66] R. Bayerbach, D. Meier, Characterization of the water-insoluble fraction from fast pyrolysis liquids (pyrolytic lignin). Part IV: Structure elucidation of oligomeric molecules, Journal of Analytical and Applied Pyrolysis 85 (2009) 98–107. https://doi.org/10.1016/j.jaap.2008.10.021.
[67] I.I. Pikovskoi, D.S. Kosyakov, Kendrick mass defect analysis — a tool for high-resolution Orbitrap mass spectrometry of native lignin, Anal Bioanal Chem 415 (2023) 3525–3534. https://doi.org/10.1007/s00216-023-04742-3.
[68] A.P. Pinheiro Pires, J. Arauzo, I. Fonts, M.E. Domine, A. Fernández Arroyo, M.E. Garcia-Perez, J. Montoya, F. Chejne, P. Pfromm, M. Garcia-Perez, Challenges and Opportunities for Bio-oil Refining: A Review, Energy Fuels 33 (2019) 4683–4720. https://doi.org/10.1021/acs.energyfuels.9b00039.
[69] A. Oasmaa, J. Lehto, Y. Solantausta, S. Kallio, Historical Review on VTT Fast Pyrolysis Bio-oil Production and Upgrading, Energy Fuels 35 (2021) 5683–5695. https://doi.org/10.1021/acs.energyfuels.1c00177.
[70] S. Papari, K. Hawboldt, A review on condensing system for biomass pyrolysis process, Fuel Processing Technology 180 (2018) 1–13. https://doi.org/10.1016/j.fuproc.2018.08.001.
[71] A.-C. Johansson, K. Iisa, L. Sandström, H. Ben, H. Pilath, S. Deutch, H. Wiinikka, O.G.W. Öhrman, Fractional condensation of pyrolysis vapors produced from Nordic feedstocks in cyclone pyrolysis, Journal of Analytical and Applied Pyrolysis 123 (2017) 244–254. https://doi.org/10.1016/j.jaap.2016.11.020.
[72] R.J.M. Westerhof, N.J.M. Kuipers, S.R.A. Kersten, W.P.M. van Swaaij, Controlling the Water Content of Biomass Fast Pyrolysis Oil, Ind. Eng. Chem. Res. 46 (2007) 9238–9247. https://doi.org/10.1021/ie070684k.
[73] R.J.M. Westerhof, D.W.F. Brilman, M. Garcia-Perez, Z. Wang, S.R.G. Oudenhoven, W.P.M. van Swaaij, S.R.A. Kersten, Fractional Condensation of Biomass Pyrolysis Vapors, Energy Fuels 25 (2011) 1817–1829. https://doi.org/10.1021/ef2000322.
[74] N.T. Machado, S.A.P.D. Mota, R.A.C. Leão, R.O.M.A.D. Souza, S. Duvoisin Junior, L.E.P. Borges, A.D.A.M.D. Mota, Upgrading/Deacidification of Bio-Oils by Liquid–Liquid Extraction Using Aqueous Methanol as a Solvent, Energies 17 (2024) 2713. https://doi.org/10.3390/en17112713.
[75] G.K. Parku, S.R. Pulicanti, A. Funke, N. Dahmen, Phase Equilibria Aided Optimization of Levoglucosan Extraction during Condensation of Fast Pyrolysis Bio-Oils, Energy Fuels 38 (2024) 14343–14350. https://doi.org/10.1021/acs.energyfuels.4c01097.

[76] M.R. Rover, P.A. Johnston, T. Jin, R.G. Smith, R.C. Brown, L. Jarboe, Production of Clean Pyrolytic Sugars for Fermentation, ChemSusChem 7 (2014) 1662–1668. https://doi.org/10.1002/cssc.201301259.
[77] M. Basafa, K. Hawboldt, A review on sources and extraction of phenolic compounds as precursors for bio-based phenolic resins, Biomass Conv. Bioref. (2021). https://doi.org/10.1007/s13399-021-01408-x.
[78] A. Oasmaa, E. Kuoppala, Y. Solantausta, Fast Pyrolysis of Forestry Residue. 2. Physicochemical Composition of Product Liquid, Energy Fuels 17 (2003) 433–443. https://doi.org/10.1021/ef020206g.
[79] A. Oasmaa, T. Sundqvist, E. Kuoppala, M. Garcia-Perez, Y. Solantausta, C. Lindfors, V. Paasikallio, Controlling the Phase Stability of Biomass Fast Pyrolysis Bio-oils, Energy Fuels 29 (2015) 4373–4381. https://doi.org/10.1021/acs.energyfuels.5b00607.
[80] H. Jeon, J.-Y. Park, J. Lee, C.-H. Oh, J.-K. Kim, J. Yoon, Fractional Composition Analysis for Upgrading of Fast Pyrolysis Bio-Oil Produced from Sawdust, Energies 15 (2022) 2054. https://doi.org/10.3390/en15062054.
[81] E.B. Hassan, H. Abou-Yousef, P. Steele, Increasing the efficiency of fast pyrolysis process through sugar yield maximization and separation from aqueous fraction bio-oil, Fuel Processing Technology 110 (2013) 65–72. https://doi.org/10.1016/j.fuproc.2012.11.003.
[82] A. Oasmaa, I. Fonts, M.R. Pelaez-Samaniego, M.E. Garcia-Perez, M. Garcia-Perez, Pyrolysis Oil Multiphase Behavior and Phase Stability: A Review, Energy Fuels 30 (2016) 6179–6200. https://doi.org/10.1021/acs.energyfuels.6b01287.
[83] P. Kostetskyy, L.J. Broadbelt, Progress in Modeling of Biomass Fast Pyrolysis: A Review, Energy Fuels 34 (2020) 15195–15216. https://doi.org/10.1021/acs.energyfuels.0c02295.
[84] F.G. Fonseca, A. Funke, Modeling of Liquid–Vapor Phase Equilibria of Pyrolysis Bio-oils: A Review, Ind. Eng. Chem. Res. 63 (2024) 13401–13420. https://doi.org/10.1021/acs.iecr.4c00775.
[85] M. Li, M. Zhang, Y. Yu, H. Wu, Ternary System of Pyrolytic Lignin, Mixed Solvent, and Water: Phase Diagram and Implications, Energy Fuels 32 (2018) 465–474. https://doi.org/10.1021/acs.energyfuels.7b02943.
[86] Y. Ille, F. Kröhl, A. Velez, A. Funke, S. Pereda, K. Schaber, N. Dahmen, Activity of water in pyrolysis oil—Experiments and modelling, Journal of Analytical and Applied Pyrolysis 135 (2018) 260–270. https://doi.org/10.1016/j.jaap.2018.08.027.
[87] M.R. Salas, F.G. Fonseca, A.C.C. De Araujo, M. García-Perez, A. Funke, Liquid–Liquid Equilibrium Prediction in Fast Pyrolysis Bio-Oil Systems: A Framework for Incorporating Bio-Oil Complexity, Energy Fuels 38 (2024) 18769–18780. https://doi.org/10.1021/acs.energyfuels.4c03387.
[88] I. Fonts, M. Atienza-Martínez, H.-H. Carstensen, M. Benés, A.P. Pinheiro Pires, M. Garcia-Perez, R. Bilbao, Thermodynamic and Physical Property Estimation of Compounds Derived from the Fast Pyrolysis of Lignocellulosic Materials, Energy Fuels 35 (2021) 17114–17137. https://doi.org/10.1021/acs.energyfuels.1c01709.
[89] R. Manrique, E. Terrell, P. Kostetskyy, F. Chejne, M. Olarte, L. Broadbelt, M. García-Pérez, Elucidating Biomass-Derived Pyrolytic Lignin Structures from Demethylation Reactions through Density Functional Theory Calculations, Energy Fuels 37 (2023) 5189–5205. https://doi.org/10.1021/acs.energyfuels.2c04292.
[90] A. Krutof, K.A. Hawboldt, Thermodynamic model of fast pyrolysis bio-oil advanced distillation curves, Fuel 261 (2020) 116446. https://doi.org/10.1016/j.fuel.2019.116446.
[91] C.W. Lahive, P.C.J. Kamer, C.S. Lancefield, P.J. Deuss, An Introduction to Model Compounds of Lignin Linking Motifs; Synthesis and Selection Considerations for

Reactivity Studies, ChemSusChem 13 (2020) 4238–4265. https://doi.org/10.1002/cssc.202000989.
[92] M. Rojas, F.G. Fonseca, U. Hornung, A. Funke, N. Dahmen, Synthetic Lignin Oligomers: Analytical Techniques, Challenges, and Opportunities, ChemSusChem 18 (2025). https://doi.org/10.1002/cssc.202402334.
[93] E. Henrich, N. Dahmen, F. Weirich, R. Reimert, C. Kornmayer, Fast pyrolysis of lignocellulosics in a twin screw mixer reactor, Fuel Processing Technology 143 (2016) 151–161. https://doi.org/10.1016/j.fuproc.2015.11.003.
[94] A. Funke, D. Richter, A. Niebel, N. Dahmen, J. Sauer, Fast Pyrolysis of Biomass Residues in a Twin-screw Mixing Reactor, JoVE (2016) 54395. https://doi.org/10.3791/54395.
[95] A. Demirbaş, Calculation of higher heating values of biomass fuels, Fuel 76 (1997) 431–434. https://doi.org/10.1016/S0016-2361(97)85520-2.
[96] M. Windt, D. Meier, J.H. Marsman, H.J. Heeres, S. De Koning, Micro-pyrolysis of technical lignins in a new modular rig and product analysis by GC–MS/FID and GC×GC–TOFMS/FID, Journal of Analytical and Applied Pyrolysis 85 (2009) 38–46. https://doi.org/10.1016/j.jaap.2008.11.011.
[97] B. Dubis, K. Bułkowska, M. Lewandowska, W. Szempliński, K.J. Jankowski, J. Idźkowski, N. Kordala, K. Szymańska, Effect of different nitrogen fertilizer treatments on the conversion of Miscanthus×giganteus to ethanol, Bioresource Technology 243 (2017) 731–737. https://doi.org/10.1016/j.biortech.2017.07.005.
[98] N. Brosse, A. Dufour, X. Meng, Q. Sun, A. Ragauskas, *Miscanthus* : a fast-growing crop for biofuels and chemicals production, Biofuels Bioprod Bioref 6 (2012) 580–598. https://doi.org/10.1002/bbb.1353.
[99] K.P. Woli, M.B. David, J. Tsai, T.B. Voigt, R.G. Darmody, C.A. Mitchell, Evaluating silicon concentrations in biofuel feedstock crops Miscanthus and switchgrass, Biomass and Bioenergy 35 (2011) 2807–2813. https://doi.org/10.1016/j.biombioe.2011.03.007.
[100] T. Han, N. Sophonrat, A. Tagami, O. Sevastyanova, P. Mellin, W. Yang, Characterization of lignin at pre-pyrolysis temperature to investigate its melting problem, Fuel 235 (2019) 1061–1069. https://doi.org/10.1016/j.fuel.2018.08.120.
[101] M.E. Moustaqim, A.E. Kaihal, M.E. Marouani, S. Men-La-Yakhaf, M. Taibi, S. Sebbahi, S.E. Hajjaji, F. Kifani-Sahban, Thermal and thermomechanical analyses of lignin, Sustainable Chemistry and Pharmacy 9 (2018) 63–68. https://doi.org/10.1016/j.scp.2018.06.002.
[102] D.R. Naron, F.-X. Collard, L. Tyhoda, J.F. Görgens, Production of phenols from pyrolysis of sugarcane bagasse lignin: Catalyst screening using thermogravimetric analysis – Thermal desorption – Gas chromatography – Mass spectroscopy, Journal of Analytical and Applied Pyrolysis 138 (2019) 120–131. https://doi.org/10.1016/j.jaap.2018.12.015.
[103] E. Jakab, Analytical Techniques as a Tool to Understand the Reaction Mechanism, in: Recent Advances in Thermo-Chemical Conversion of Biomass, Elsevier, 2015: pp. 75–108. https://doi.org/10.1016/B978-0-444-63289-0.00003-X.
[104] P.A. Lazaridis, A.P. Fotopoulos, S.A. Karakoulia, K.S. Triantafyllidis, Catalytic Fast Pyrolysis of Kraft Lignin With Conventional, Mesoporous and Nanosized ZSM-5 Zeolite for the Production of Alkyl-Phenols and Aromatics, Front. Chem. 6 (2018) 295. https://doi.org/10.3389/fchem.2018.00295.
[105] A. Aho, N. Kumar, K. Eränen, B. Holmbom, M. Hupa, T. Salmi, D.Yu. Murzin, Pyrolysis of Softwood Carbohydrates in a Fluidized Bed Reactor, IJMS 9 (2008) 1665–1675. https://doi.org/10.3390/ijms9091665.
[106] M. Bergs, G. Völkering, T. Kraska, R. Pude, X.T. Do, P. Kusch, Y. Monakhova, C. Konow, M. Schulze, Miscanthus x giganteus Stem Versus Leaf-Derived Lignins

Differing in Monolignol Ratio and Linkage, IJMS 20 (2019) 1200. https://doi.org/10.3390/ijms20051200.
[107] M. Bergs, X.T. Do, J. Rumpf, P. Kusch, Y. Monakhova, C. Konow, G. Völkering, R. Pude, M. Schulze, Comparing chemical composition and lignin structure of *Miscanthus x giganteus* and *Miscanthus nagara* harvested in autumn and spring and separated into stems and leaves, RSC Adv. 10 (2020) 10740–10751. https://doi.org/10.1039/C9RA10576J.
[108] P.S. Marathe, R.J.M. Westerhof, S.R.A. Kersten, Fast pyrolysis of lignins with different molecular weight: Experiments and modelling, Applied Energy 236 (2019) 1125–1137. https://doi.org/10.1016/j.apenergy.2018.12.058.
[109] F.G. Fonseca, A. Anca-Couce, A. Funke, N. Dahmen, Challenges in Kinetic Parameter Determination for Wheat Straw Pyrolysis, Energies 15 (2022) 7240. https://doi.org/10.3390/en15197240.
[110] A. Moutsoglou, B. Lawburgh, J. Lawburgh, Fractional condensation and aging of pyrolysis oil from softwood and organosolv lignin, Journal of Analytical and Applied Pyrolysis 135 (2018) 350–360. https://doi.org/10.1016/j.jaap.2018.08.016.
[111] T.N. Trinh, P.A. Jensen, K. Dam-Johansen, N.O. Knudsen, H.R. Sørensen, S. Hvilsted, Comparison of Lignin, Macroalgae, Wood, and Straw Fast Pyrolysis, Energy Fuels 27 (2013) 1399–1409. https://doi.org/10.1021/ef301927y.
[112] A.V. Bridgwater, Review of fast pyrolysis of biomass and product upgrading, Biomass and Bioenergy 38 (2012) 68–94. https://doi.org/10.1016/j.biombioe.2011.01.048.
[113] M. Borella, A.A. Casazza, G. Garbarino, P. Riani, G. Busca, A Study of the Pyrolysis Products of Kraft Lignin, Energies 15 (2022) 991. https://doi.org/10.3390/en15030991.
[114] J. Gracia-Vitoria, S.C. Gándara, E. Feghali, P. Ortiz, W. Eevers, K.S. Triantafyllidis, K. Vanbroekhoven, The chemical and physical properties of lignin bio-oils, facts and needs, Current Opinion in Green and Sustainable Chemistry 40 (2023) 100781. https://doi.org/10.1016/j.cogsc.2023.100781.
[115] W. Mu, H. Ben, A. Ragauskas, Y. Deng, Lignin Pyrolysis Components and Upgrading—Technology Review, Bioenerg. Res. 6 (2013) 1183–1204. https://doi.org/10.1007/s12155-013-9314-7.
[116] B. Scholze, D. Meier, Characterization of the water-insoluble fraction from pyrolysis oil (pyrolytic lignin). Part I. PY–GC/MS, FTIR, and functional groups, Journal of Analytical and Applied Pyrolysis 60 (2001) 41–54. https://doi.org/10.1016/S0165-2370(00)00110-8.
[117] J.O. Ajikashile, M.-J. Alhnidi, G.K. Parku, A. Funke, A. Kruse, A study on the fast pyrolysis of millet and sorghum straws sourced from arid and semi-arid regions of Nigeria in a twin-screw mixing reactor, Materials Science for Energy Technologies 6 (2023) 388–398. https://doi.org/10.1016/j.mset.2023.03.007.
[118] G.R. Gomes, E.G. De Jesus, J.C.C. Jacintho, D.L.G. García, B.R.A. Alencar, F.P. Gabetto, J.J. Gomes, J.L.N. Carvalho, M. Strauss, C. Driemeier, Peculiarities of bio-oil and biochar obtained from the lignin-rich residue of the enzymatic hydrolysis of sugarcane bagasse, Renewable Energy 241 (2025) 122282. https://doi.org/10.1016/j.renene.2024.122282.
[119] S.H. Beis, S. Mukkamala, N. Hill, J. Joseph, C. Baker, B. Jensen, E.A. Stemmler, C. Wheeler, B.G. Frederick, A. Van Heiningen, A.G. Berg, W.J. DeSisto, Fast pyrolysis of lignins, BioRes 5 (2010) 1408–1424. https://doi.org/10.15376/biores.5.3.1408-1424.
[120] A.G. Margellou, P.A. Lazaridis, I.D. Charisteidis, C.K. Nitsos, C.P. Pappa, A.P. Fotopoulos, S. Van Den Bosch, B.F. Sels, K.S. Triantafyllidis, Catalytic fast pyrolysis of beech wood lignin isolated by different biomass (pre)treatment processes: Organosolv, hydrothermal and enzymatic hydrolysis, Applied Catalysis A: General 623 (2021) 118298. https://doi.org/10.1016/j.apcata.2021.118298.

[121] T.M. Santos, W.R.D. Silva, J.D.C. Carregosa, C.C. Schmitt, R. Moreira, K. Raffelt, N. Dahmen, A. Wisniewski, Thermal Conversion of Sugarcane Bagasse Coupled with Vapor Phase Hydrotreatment over Nickel-Based Catalysts: A Comprehensive Characterization of Upgraded Products, Catalysts 12 (2022) 355. https://doi.org/10.3390/catal12040355.
[122] S. Ralph, J. Ralph, F. Lu, NMR Database of Lignin and Cell Wall Model Compounds., (2024). https://doi.org/10.11578/2409191.
[123] C. Crestini, H. Lange, M. Sette, D.S. Argyropoulos, On the structure of softwood kraft lignin, Green Chem. 19 (2017) 4104–4121. https://doi.org/10.1039/C7GC01812F.
[124] M.P. Pandey, C.S. Kim, Lignin Depolymerization and Conversion: A Review of Thermochemical Methods, Chem Eng & Technol 34 (2011) 29–41. https://doi.org/10.1002/ceat.201000270.
[125] D.S. Zijlstra, C.W. Lahive, C.A. Analbers, M.B. Figueirêdo, Z. Wang, C.S. Lancefield, P.J. Deuss, Mild Organosolv Lignin Extraction with Alcohols: The Importance of Benzylic Alkoxylation, ACS Sustainable Chem. Eng. 8 (2020) 5119–5131. https://doi.org/10.1021/acssuschemeng.9b07222.
[126] J. Prothmann, P. Spégel, M. Sandahl, C. Turner, Identification of lignin oligomers in Kraft lignin using ultra-high-performance liquid chromatography/high-resolution multiple-stage tandem mass spectrometry (UHPLC/HRMSn), Anal Bioanal Chem 410 (2018) 7803–7814. https://doi.org/10.1007/s00216-018-1400-4.
[127] S.O. Asare, F. Huang, B.C. Lynn, Characterization and sequencing of lithium cationized β-O-4 lignin oligomers using higher-energy collisional dissociation mass spectrometry, Analytica Chimica Acta 1047 (2019) 104–114. https://doi.org/10.1016/j.aca.2018.09.068.
[128] Aspen Technology, Inc., Aspen Plus® V14 – Physical Property Methods and Models, (2024).
[129] R. Hempfling, H.-R. Schulten, Chemical characterization of the organic matter in forest soils by Curie point pyrolysis-GC/MS and pyrolysis-field ionization mass spectrometry, Organic Geochemistry 15 (1990) 131–145. https://doi.org/10.1016/0146-6380(90)90078-E.
[130] E. Ranzi, P.E.A. Debiagi, A. Frassoldati, Mathematical Modeling of Fast Biomass Pyrolysis and Bio-Oil Formation. Note I: Kinetic Mechanism of Biomass Pyrolysis, ACS Sustainable Chem. Eng. 5 (2017) 2867–2881. https://doi.org/10.1021/acssuschemeng.6b03096.
[131] O.C. Bridgeman, E.W. Aldrich, Vapor Pressure Tables for Water, Journal of Heat Transfer 86 (1964) 279–286. https://doi.org/10.1115/1.3687121.
[132] Y. Han, A. Paiva Pinheiro Pires, M. Denson, A.G. McDonald, M. Garcia-Perez, Ternary Phase Diagram of Water/Bio-Oil/Organic Solvent for Bio-Oil Fractionation, Energy Fuels 34 (2020) 16250–16264. https://doi.org/10.1021/acs.energyfuels.0c03100.
[133] F.G. Fonseca, A. Funke, N. Dahmen, Aspen plusTM modeling of fractional condensation schemes for production of fast pyrolysis bio-oil, in: 27th European Biomass Conference and Exhibition, Lisbon, Portugal, 2019: pp. 1227–1233.
[134] Q. Li, D.G. Vlachos, 5–5 Lignin Linkage Cleavage over Ru: A Density Functional Theory Study, ACS Sustainable Chem. Eng. 9 (2021) 16143–16152. https://doi.org/10.1021/acssuschemeng.1c04838.
[135] E. Adler, E. Eriksoo, E. Boss, A. Caglieris, Guaiacylglycerol and its beta-Guaiacyl Ether., Acta Chem. Scand. 9 (1955) 341–342. https://doi.org/10.3891/acta.chem.scand.09-0341.
[136] F. Nakatsubo, K. Sato, T. Higuchi, Synthesis of Guaiacylglycerol-β-guaiacyl Ether, Holzforschung 29 (1975) 165–168. https://doi.org/10.1515/hfsg.1975.29.5.165.

[137] T. Kishimoto, Y. Uraki, M. Ubukata, Synthesis of Bromoacetophenone Derivatives as Starting Monomers for β-O-4 Type Artificial Lignin Polymers, J. of Wood Chem. & Tech. 28 (2008) 97–105. https://doi.org/10.1080/02773810802124894.
[138] T. Kishimoto, Y. Uraki, M. Ubukata, Synthesis of β-O-4-type artificial lignin polymers and their analysis by NMR spectroscopy, Org. Biomol. Chem. 6 (2008) 2982. https://doi.org/10.1039/b805460f.
[139] Y. Ille, F.A. Sánchez, N. Dahmen, S. Pereda, Multiphase Equilibria Modeling of Fast Pyrolysis Bio-Oils. Group Contribution Associating Equation of State Extension to Lignin Monomers and Derivatives, Ind. Eng. Chem. Res. 58 (2019) 7318–7331. https://doi.org/10.1021/acs.iecr.9b00227.
[140] M.D. Denson, E. Terrell, P. Kostetskyy, M. Olarte, L. Broadbelt, M. Garcia-Perez, Elucidation of Structure and Physical Properties of Pyrolytic Sugar Oligomers Derived from Cellulose Depolymerization/Dehydration Reactions: A Density Functional Theory Study, Energy Fuels 37 (2023) 7834–7847. https://doi.org/10.1021/acs.energyfuels.3c00641.

Supporting Information

Supporting Information – Chapter 2

Table S1. Detailed composition of organic rich condensate (ORC) from Kraft lignin quantified by GC-MS/FID.

CAS No.	Compound	wt.% wet	
	NONAROMATIC COMPOUNDS		
	Acids	**0.000**	
	Nonaromatic Esters	**0.000**	
	Nonaromatic Alcohols	**0.000**	
	Nonaromatic Aldehydes	**0.000**	
	Nonaromatic Ketones	**0.704**	
123-42-2	2-Pentanone, 4-hydroxy-4-methyl- = Diacetone alcohol = impurity of Acetone	0.248	c
1121-05-7	Cyclopenten-1-one, 2,3-dimethyl-2-	0.074	c
1120-73-6	Cyclopenten-1-one, 2-methyl-2-	0.056	c
2758-18-1	Cyclopenten-1-one, 3-methyl-2-	0.123	c
	4-Penten-2-one, 4-methyl- (NIST MQ 92)	0.071	#
	3-Penten-2-one, 4-methyl- (NIST MQ 92)	0.103	#
	2-Cyclopenten-1-one, x,y-dimethyl-	0.028	#
	Aliphatic Hydrocarbons	**0.000**	
	HETEROCYCLIC COMPOUNDS		
	Furans	**0.091**	
96-48-0	Butyrolactone, γ-	0.091	c
	Pyrans	**0.000**	
	AROMATIC COMPOUNDS		
	Benzenes	**0.405**	
108-88-3	Toluene	0.011	c
494-99-5	Toluene, 3,4-dimethoxy-	0.163	c
	Benzene, 1-methoxy-2,3-dimethyl- (NIST MQ 88)	0.013	#
	Naphthalene, 1-phenyl- (impurity in IS = Fluoranthene)	0.036	#
	C18H18: Retene = Phenanthrene, 7-isopropyl-1-methyl- (NIST MQ 89)	0.182	#
	Catechols	**n.q.**	
120-80-9	Catechol (Benzene, 1,2-dihydroxy-)	n.q.	n.q.

452-86-8	Catechol, 4-methyl-	n.q.	n.q.
	Benzenediol, methyl-	n.q.	n.q.
	Catechol, 4-ethyl- (NIST MQ 92)	n.q.	n.q.
	Aromatic Alcohols	**0.000**	
	Aromatic Aldehydes	**0.000**	
	Aromatic Ketones	**0.053**	
1131-62-0	Acetophenone, 3,4-dimethoxy-	0.053	c
	Aromatic Esters	**0.000**	
	Lignin derived Phenols	**5.784**	
108-95-2	Phenol	1.126	c
95-48-7	Cresol, o-	0.941	c
106-44-5	Cresol, p-	0.625	c
108-39-4	Cresol, m-	0.303	c
95-87-4	Phenol, 2,5-dimethyl-	0.863	c
576-26-1	Phenol, 2,6-dimethyl-	0.074	c
108-68-9	Phenol, 3,5-dimethyl-	0.047	c
527-60-6	Phenol, 2,4,6-trimethyl-	0.060	c
90-00-6	Phenol, 2-ethyl-	0.090	c
620-17-7	Phenol, 3-ethyl-	0.138	c
123-07-9	Phenol, 4-ethyl-	0.351	c
2628-17-3	Phenol, 4-vinyl-	0.182	#
85960-81-2	Phenol , cis 4-propenyl-	0.106	#
5932-68-3	Phenol, trans 4-propenyl-	0.168	#
99-93-4	Acetophenone, 4-hydroxy-	0.138	c
	Phenol, ethyl-methyl-	0.354	#
	Phenol, 4-methyl-2-(2-propenyl)- (NIST MQ 88)	0.217	#
	Guaiacols (Methoxy phenols)	**15.905**	
90-05-1	Guaiacol	4.095	c
18102-31-3	Guaiacol, 3-methyl-	0.271	#
93-51-6	Guaiacol, 4-methyl-	2.294	c
	Guaiacol, 3-ethyl-	0.158	#
2785-89-9	Guaiacol, 4-ethyl-	1.117	c
7786-61-0	Guaiacol, 4-vinyl-	0.799	#
97-53-0	Guaiacol, 4-allyl-; (Eugenol)	0.354	c
2785-87-7	Guaiacol, 4-propyl-	0.186	c
97-54-1	Guaiacol, 4-propenyl- cis (Isoeugenol)	0.758	c
5932-68-3	Guaiacol, 4-propenyl-(trans) (Isoeugenol)	2.168	c
121-33-5	Vanillin	0.592	c
2305-13-7	Dihydroconiferyl alcohol	1.420	#
498-02-2	Ethanone, 1-(4-hydroxy-3-methoxyphenyl)- (Acetoguaiacone)	1.139	c
1835-14-9	Propioguaiacone	0.214	#

2503-46-0	Guaiacyl acetone	0.341	c
	Syringols (Dimethoxy phenols)	**0.000**	
	CARBOHYDRATES		
	Sugars	**0.000**	
	OTHER ORGANIC COMPOUNDS		
	N-compounds	**0.968**	
	unknown N-Compound (no NIST spectrum found)	0.113	#
	unknown N-Compound (no NIST spectrum found)	0.219	#
	4-Piperidinone, 2,2,6,6-tetramethyl- (NIST MQ 87)	0.553	#
	unknown N-Compound (no NIST spectrum found)	0.083	#
	Acetates	**0.000**	
	Terpenes	**0.000**	
	unknown compounds	**0.000**	
	Miscellaneous	**0.000**	

c = calibrated compound
n.q. = not quantified
# = estimated response factor

Table S2. Detailed composition of aqueous condensate (AC) from Kraft lignin quantified by GC-MS/FID.

CAS No.	Compound	wt.% wet	
	NONAROMATIC COMPOUNDS		
	Acids	**0.251**	
64-19-7	Acetic acid	0.251	c
	Nonaromatic Esters	**0.000**	
	Nonaromatic Alcohols	**0.025**	
71-23-8	Propanol, 1-	0.006	#
	2-Propen-1-ol (NIST MQ 91)	0.019	#
	Nonaromatic Aldehydes	**0.000**	
	Nonaromatic Ketones	**1.433**	
123-42-2	2-Pentanone, 4-hydroxy-4-methyl- = Diacetone alcohol = impurity of Acetone	1.335	c
78-93-3	Butanone, 2-	0.007	c
	4-Penten-2-one, 4-methyl- (NIST MQ 92)	0.009	#
	3-Penten-2-one, 4-methyl- (NIST MQ 92)	0.082	#
	Aliphatic Hydrocarbons	**0.000**	
	HETEROCYCLIC COMPOUNDS		
	Furans	**0.000**	
	Pyrans	**0.000**	
	AROMATIC COMPOUNDS		
	Benzenes	**0.000**	
	Catechols	**0.000**	
	Aromatic Alcohols	**0.000**	
	Aromatic Aldehydes	**0.000**	
	Aromatic Ketones	**0.000**	
	Aromatic Esters	**0.000**	
	Lignin derived Phenols	**0.524**	
108-95-2	Phenol	0.309	c
95-48-7	Cresol, o-	0.103	c
106-44-5	Cresol, p-	0.047	c
108-39-4	Cresol, m-	0.025	c

95-87-4	Phenol, 2,5-dimethyl-	0.026	c
576-26-1	Phenol, 2,6-dimethyl-	0.004	c
90-00-6	Phenol, 2-ethyl-	0.003	c
123-07-9	Phenol, 4-ethyl-	0.006	c
	Phenol, ethyl-methyl-	0.002	#
	Guaiacols (Methoxy phenols)	**0.504**	
90-05-1	Guaiacol	0.435	c
93-51-6	Guaiacol, 4-methyl-	0.061	c
2785-89-9	Guaiacol, 4-ethyl-	0.008	c
	Syringols (Dimethoxy phenols)	**0.000**	
	CARBOHYDRATES		
	Sugars	**0.000**	
	OTHER ORGANIC COMPOUNDS		
	N-compounds	**1.843**	
	unknown N-Compound (no NIST spectrum found)	0.006	#
	poss. 2-Pentanone, 4-amino-4-methyl- (NIST MQ 77)	0.571	#
	Pyrrole, 1-methyl-3-(1,1-dimethylethyl)- (NIST MQ 84)	0.060	#
	unknown N-Compound (no NIST spectrum found)	0.011	#
	unknown N-Compound (no NIST spectrum found)	0.988	#
	unknown N-Compound (no NIST spectrum found)	0.007	#
	unknown N-Compound (no NIST spectrum found)	0.027	#
	4-Piperidinone, 2,2,6,6-tetramethyl- (NIST MQ 87)	0.173	#
	Acetates	**0.000**	
	Terpenes	**0.000**	
	unknown compounds	**0.000**	
	Miscellaneous	**0.000**	

c = calibrated compound
n.q. = not quantified
# = estimated response factor

Table S3. Detailed composition of organic rich condensate (ORC) from Miscancell lignin quantified by GC-MS/FID.

CAS No.	Compound	wt.% wet	
	NONAROMATIC COMPOUNDS		
	Acids	**8.144**	
64-19-7	Acetic acid	3.442	c
79-09-4	Propionic acid	4.703	c
	Nonaromatic Esters	**0.000**	
	Nonaromatic Alcohols	**0.000**	
	Nonaromatic Aldehydes	**0.000**	
	Nonaromatic Ketones	**1.481**	
123-42-2	2-Pentanone, 4-hydroxy-4-methyl- = Diacetone alcohol = impurity of Acetone	0.198	c
78-93-3	Butanone, 2-	0.040	c
592-20-1	Propan-2-one, 1-acetyloxy-	0.037	c
120-92-3	Cyclopentanone	0.077	c
930-30-3	Cyclopenten-1-one, 2-	0.099	c
1120-73-6	Cyclopenten-1-one, 2-methyl-2-	0.120	c
2758-18-1	Cyclopenten-1-one, 3-methyl-2-	0.235	c
5682-69-2	Cylopenten-1-one, 3-ethyl-2-	0.072	#
80-71-7	Cyclopenten-1-one, 2-hydroxy-3-methyl-2-	0.564	c
	3-Penten-2-one, 4-methyl- (NIST MQ 92)	0.040	#
	Aliphatic Hydrocarbons	**0.000**	
	HETEROCYCLIC COMPOUNDS		
	Furans	**0.462**	
497-23-4	Furanone, 2(5H)-	0.171	c
1192-62-7	Ethanone, 1-(2-furanyl)-	0.019	c
22122-36-7	Furan-2-one, 3-methyl-, (5H)-	0.042	c
96-48-0	Butyrolactone, γ-	0.229	c
	Pyrans	**0.000**	
	AROMATIC COMPOUNDS		
	Benzenes	**0.553**	
494-99-5	Toluene, 3,4-dimethoxy-	0.034	c
83-33-0	Inden-1-one, 2,3-dihydro-1H-	0.051	c
	Benzofuran, 2,3-dihydro- (NIST MQ 86)	0.430	#
	1H-Indenol (NIST MQ 82)	0.038	#
	Catechols	**n.q.**	

123-31-9	Hydroquinone (Benzene, 1,4-dihydroxy-)	n.q.	n.q.
108-46-3	Resorcinol (Benzene, 1,3-dihydroxy-)	n.q.	n.q.
	Benzenediol, methyl-	n.q.	n.q.
	Aromatic Alcohols	**0.000**	
	Aromatic Aldehydes	**0.000**	
	Aromatic Ketones	**0.000**	
	Aromatic Esters	**0.000**	
	Lignin derived Phenols	**3.871**	
108-95-2	Phenol	0.913	c
95-48-7	Cresol, o-	0.357	c
106-44-5	Cresol, p-	0.408	c
108-39-4	Cresol, m-	0.317	c
95-87-4	Phenol, 2,5-dimethyl-	0.122	c
105-67-9	Phenol, 2,4-dimethyl-	0.155	c
108-68-9	Phenol, 3,5-dimethyl-	0.043	c
90-00-6	Phenol, 2-ethyl-	0.038	c
620-17-7	Phenol, 3-ethyl-	0.171	c
123-07-9	Phenol, 4-ethyl-	0.886	c
85960-81-2	Phenol , cis 4-propenyl-	0.082	#
5932-68-3	Phenol, trans 4-propenyl-	0.115	#
150-19-6	Phenol, 3-methoxy-	0.059	c
	Phenol, ethyl-methyl-	0.135	#
	Phenol, 3-methoxy-5-methyl- (NIST MQ 88)	0.071	#
	Guaiacols (Methoxy phenols)	**6.672**	
90-05-1	Guaiacol	1.029	c
93-51-6	Guaiacol, 4-methyl-	0.544	c
	Guaiacol, 3-ethyl-	0.072	#
2785-89-9	Guaiacol, 4-ethyl-	0.533	c
7786-61-0	Guaiacol, 4-vinyl-	1.104	#
97-53-0	Guaiacol, 4-allyl-; (Eugenol)	0.182	c
2785-87-7	Guaiacol, 4-propyl-	0.077	c
97-54-1	Guaiacol, 4-propenyl- cis (Isoeugenol)	0.466	c
5932-68-3	Guaiacol, 4-propenyl-(trans) (Isoeugenol)	1.551	c
121-33-5	Vanillin	0.363	c
498-02-2	Ethanone, 1-(4-hydroxy-3-methoxyphenyl)- (Acetoguaiacone)	0.557	c
2503-46-0	Guaiacyl acetone	0.195	c
	Syringols (Dimethoxy phenols)	**3.971**	
91-10-1	Syringol	1.014	c
6638-05-7	Syringol, 4-methyl-	0.391	c
14059-92-8	Syringol, 4-ethyl-	0.269	c

28343-22-8	Syringol, 4-vinyl-	0.509	#
6627-88-9	Syringol, 4-allyl-	0.386	c
26624-13-5	Syringol, 4-(1-propenyl)-, cis	0.242	#
20675-95-0	Syringol, 4-(1-propenyl)-, trans	0.802	#
2478-38-8	Acetosyringone	0.291	c
	Syringyl acetone	0.066	#
	CARBOHYDRATES		
	Sugars	**2.089**	
498-07-7	Anhydro-ß-D-glucopyranose, 1,6- (Levoglucosan)	1.741	c
4451-31-4	Dianhydro-α-D-glucopyranose, 1,4:3,6-	0.348	#
	OTHER ORGANIC COMPOUNDS		
	N-compounds	**0.374**	
	unknown N-Compound (no NIST spectrum found)	0.039	#
	unknown N-Compound (no NIST spectrum found)	0.118	#
	4-Piperidinone, 2,2,6,6-tetramethyl- (NIST MQ 87)	0.121	#
	unknown N-Compound (no NIST spectrum found)	0.032	#
	1H-Indole, 4-methyl- (NIST MQ 82)	0.064	#
	Acetates	**0.000**	
	Terpenes	**0.000**	
	unknown compounds	**0.000**	
	Miscellaneous	**0.000**	

c = calibrated compound
n.q. = not quantified
# = estimated response factor

Table S4. Detailed composition of aqueous condensate (AC) from Miscancell lignin quantified by GC-MS/FID.

CAS No.	Compound	wt.% wet	
	NONAROMATIC COMPOUNDS		
	Acids	**1.021**	
64-19-7	Acetic acid	0.585	c
79-09-4	Propionic acid	0.436	c
	Nonaromatic Esters	**0.000**	
	Nonaromatic Alcohols	**0.020**	
71-23-8	Propanol, 1-	0.003	#
	2-Propen-1-ol (NIST MQ 91)	0.016	#
	Nonaromatic Aldehydes	**0.000**	
	Nonaromatic Ketones	**0.600**	
123-42-2	2-Pentanone, 4-hydroxy-4-methyl- = Diacetone alcohol = impurity of Acetone	0.078	c
116-09-6	Acetol (Hydroxypropanone)	0.124	c
78-93-3	Butanone, 2-	0.062	c
5077-67-8	Butanone, 1-hydroxy-2-	0.027	c
120-92-3	Cyclopentanone	0.028	c
930-30-3	Cyclopenten-1-one, 2-	0.122	c
1121-05-7	Cyclopenten-1-one, 2,3-dimethyl-2-	0.013	c
1120-73-6	Cyclopenten-1-one, 2-methyl-2-	0.047	c
2758-18-1	Cyclopenten-1-one, 3-methyl-2-	0.035	c
80-71-7	Cyclopenten-1-one, 2-hydroxy-3-methyl-2-	0.025	c
	2-Pentanone (NIST MQ 94)	0.016	#
	3-Pentanone (NIST MQ 92)	0.007	#
	3-Penten-2-one, 4-methyl- (NIST MQ 92)	0.009	#
	Isomere of 2-Cyclopenten-1-one, 3-methyl-	0.005	#
	2-Cyclopenten-1-one, x,y-dimethyl-	0.004	#
	Aliphatic Hydrocarbons	**0.000**	
	HETEROCYCLIC COMPOUNDS		
	Furans	**0.102**	
98-00-0	Furfuryl alcohol, 2-	0.027	c
1192-62-7	Ethanone, 1-(2-furanyl)-	0.009	c
22122-36-7	Furan-2-one, 3-methyl-, (5H)-	0.008	c
	Furan-2-one, 2,5-dihydro-3,5-dimethyl-	0.008	#
96-48-0	Butyrolactone, γ-	0.050	c
	Pyrans	**0.000**	
	AROMATIC COMPOUNDS		

	Benzenes	**0.000**	
	Catechols	**0.000**	
	Aromatic Alcohols	**0.000**	
	Aromatic Aldehydes	**0.000**	
	Aromatic Ketones	**0.000**	
	Aromatic Esters	**0.000**	
	Lignin derived Phenols	**0.234**	
108-95-2	Phenol	0.144	c
95-48-7	Cresol, o-	0.035	c
106-44-5	Cresol, p-	0.024	c
108-39-4	Cresol, m-	0.017	c
105-67-9	Phenol, 2,4-dimethyl-	0.007	c
123-07-9	Phenol, 4-ethyl-	0.009	c
	Guaiacols (Methoxy phenols)	**0.218**	
90-05-1	Guaiacol	0.175	c
93-51-6	Guaiacol, 4-methyl-	0.035	c
2785-89-9	Guaiacol, 4-ethyl-	0.008	c
	Syringols (Dimethoxy phenols)	**0.006**	
91-10-1	Syringol	0.006	c
	CARBOHYDRATES		
	Sugars	**0.000**	
	OTHER ORGANIC COMPOUNDS		
	N-compounds	**0.457**	
	Propanenitrile (NIST MQ 94)	0.008	#
	Oxazole, 4,5-dihydro-2,4,4-trimethyl- (NIST MQ 82)	0.269	#
	unknown N-Compound (no NIST spectrum found)	0.071	#
	unknown N-Compound (no NIST spectrum found)	0.005	#
	unknown N-Compound (no NIST spectrum found)	0.067	#
	4-Piperidinone, 2,2,6,6-tetramethyl- (NIST MQ 87)	0.003	#
	unknown N-Compound (no NIST spectrum found)	0.009	#
	1H-Pyrazole, 3,4,5-trimethyl- (NIST MQ 88)	0.025	#
	Acetates	**0.000**	
	Terpenes	**0.000**	
	unknown compounds	**0.000**	

	Miscellaneous	**0.000**

c = calibrated compound

n.q. = not quantified

= estimated response factor

Supporting Information – Chapter 3

Figure S1. Scheme of the process to simulate the NMR 1H spectra for the mixture with ACD Labs and Python

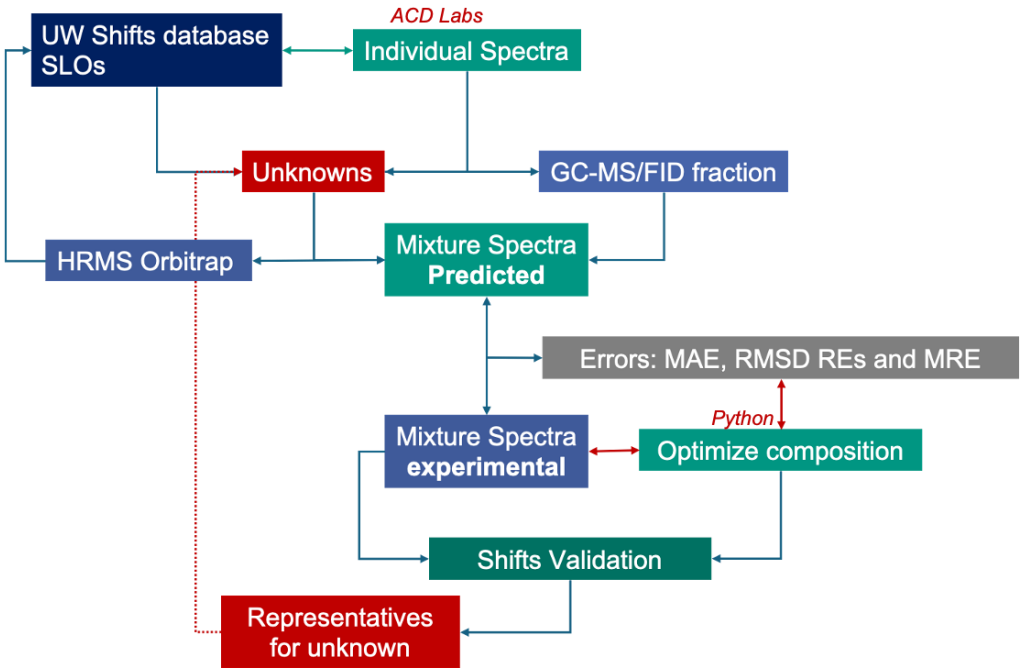


Table S5. GC-MS detailed list of Kraft (Indulin AT) lignin bio-oil.

Group	Peak#	Ret.Time	Area	Conc.	Name
PH	2	15.96	11782955.00	2.68	Phenol
PH	5	20.27	17246843.00	3.51	Phenol, 2-methyl-
OxyPH	7	21.81	76075688.00	14.29	Phenol, 2-methoxy-
OxyPH	8	22.30	8696066.00	1.63	Phenol, 2-methoxy-
PH	9	22.59	6111738.00	1.15	p-Cresol
PH	10	23.32	4583600.00	0.86	Phenol, 2,5-dimethyl-
PH	11	26.08	1017403.00	0.81	Phenol, 2-ethyl-
PH	13	26.54	10664831.00	2.99	Phenol, 2,4-dimethyl-
OxyPH	15	27.96	4529221.00	0.85	Phenol, 4-methoxy-3-methyl-
PH	16	28.25	3907998.00	0.73	Phenol, 4-ethyl-
OxyPH	19	28.89	45555424.00	8.56	Creosol
PH	21	32.00	2942465.00	0.55	Phenol, 2-ethyl-4-methyl-
OxyAR	22	32.31	5204516.00	0.98	3,4-Dimethoxytoluene
PH	23	32.53	10996347.00	2.07	Phenol, 4-ethyl-3-methyl-
PAH	26	33.81	18055058.00	3.39	Naphthalene, 2,6-bis(1,1-dimethylethyl)-
OxyPH	29	34.65	33258971.00	6.25	Phenol, 4-ethyl-2-methoxy-
OxyAR	30	35.23	2678176.00	0.50	Benzene, 4-ethyl-1,2-dimethoxy-
NIT	31	35.67	3013551.00	0.57	p-Methoxybenzamide
AC	32	36.33	3240073.00	0.61	1,3-Benzodioxole-5-carboxylic acid
OxyPH	34	37.08	14955840.00	2.81	2-Methoxy-4-vinylphenol
OxyPH	36	37.84	4335451.00	0.81	Ethanone, 1-(2-hydroxy-5-methoxyphenyl)-

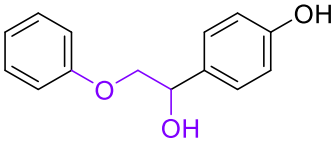
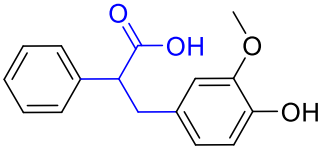
AC	37	38.19	6079901.00	1.14	3-(4-Methoxyphenyl)propionic acid
OxyPH	40	39.76	10092022.00	1.90	Eugenol
AL	41	39.91	7133354.00	1.34	5-Hepten-3-yn-2-ol, 6-methyl-5-(1-methylethyl)-
OxyPH	42	40.38	7689743.00	1.44	Phenol, 2-methoxy-4-propyl-
NIT	43	40.71	2743553.00	0.52	2-Benzothiazolamine, 4-methoxy-
OxyPH	48	42.96	6849931.00	1.29	trans-Isoeugenol
ALD	49	43.34	3033100.00	0.57	4-Ethoxy-3-anisaldehyde
PH	51	44.11	7906281.00	1.49	5-Methyl-2-allylphenol
OxyAR	53	44.69	1421782.00	0.27	4-tert-Butylcatechol, dimethyl ether
OxyPH	56	45.50	21194581.00	3.98	trans-Isoeugenol
OxyPH	60	48.13	10297037.00	1.93	Apocynin
KET	65	50.74	3836759.00	0.72	Spiro[5.6]dodecane-1,7-dione
OxyPH	73	56.33	3595902.00	0.68	4-(1-Hydroxyallyl)-2-methoxyphenol
PAH	74	56.60	3008343.00	0.57	2-Naphthalenol, 3-methoxy-
NIT	96	68.55	3200202.00	0.60	4-(Piperidin-1-yl)-5-amino veratrole
PAH	100	69.55	8661627.00	1.63	Retene
PAH	104	70.90	2873965.00	0.54	Triphenylene, 1,2,3,4-tetrahydro-
AC	112	75.74	2684821.00	0.50	4-Methoxy-4',5'-methylenedioxybiphenyl-2-carboxylic acid

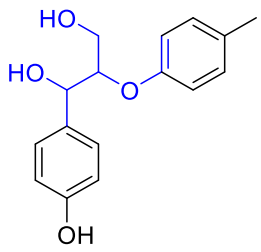
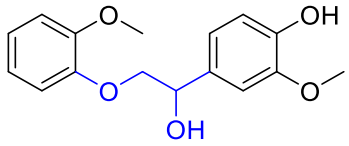
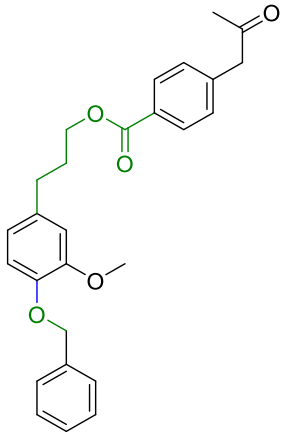
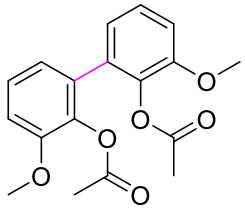
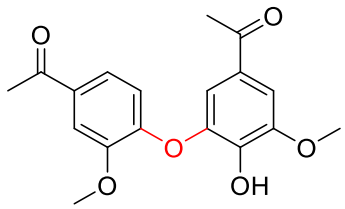
Table S6. GC-MS detailed list of alkali (Miscancell) lignin bio-oil.

Group	Peak#	Ret.Time	Area	Peak Area (%)	Name
PH	4	16.00	3876842.00	2.13	Phenol
KET	7	18.85	2201635.00	0.66	2-Cyclopenten-1-one, 2,3-dimethyl-
PH	8	20.55	5280777.00	1.60	Phenol, 2-methyl-
OxyPH	10	21.99	7393379.00	2.23	Phenol, 2-methoxy-
OxyPH	11	22.10	17078325.00	5.16	Phenol, 2-methoxy-
PH	12	23.41	1910780.00	0.58	Phenol, 2,3-dimethyl-
PH	14	26.30	1415246.00	0.43	Phenol, 2-ethyl-
PH	15	26.85	4480625.00	1.35	Phenol, 2,4-dimethyl-
PH	16	28.02	18445044.00	5.57	Phenol, 4-ethyl-
PH	17	28.55	1829195.00	0.55	Phenol, 4-ethyl-
PH	18	28.66	3582326.00	1.08	Phenol, 3-ethyl-
OxyPH	19	29.04	8739362.00	2.64	Creosol
NIT	22	32.18	2263579.00	0.68	2-Amino-N-(4-nitrophenyl)-3-phenylpropanamide
EST	23	32.40	1837004.00	0.55	Methyl salicylate
PH	24	32.64	4060199.00	1.23	p-Cumenol
PH	27	34.23	2810474.00	0.85	Phenol, 3-propyl-
OxyPH	28	34.72	13181358.00	3.98	Phenol, 4-ethyl-2-methoxy-
OxyPH	34	37.13	7266893.00	2.19	2-Methoxy-4-vinylphenol
NIT	35	37.73	2452027.00	0.74	N-(1-Cyclohexen-1-yl)piperidine
NIT	36	37.99	3762100.00	1.14	Pyrazole, 4-ethyl-3,5-dipropyl-
KET	39	38.79	5096885.00	1.54	4-Acetylphenyl ether
OxyPH	41	39.54	12326802.00	3.72	Phenol, 2,6-dimethoxy-
OxyPH	42	39.78	6604119.00	1.99	3-Allyl-6-methoxyphenol

OxyPH	43	39.96	5270629.00	1.59	2',6'-Dihydroxy-3'-methylacetophenone
PH	44	40.40	5831522.00	1.76	1,4-benzenediol, 2-(1-methylpropyl)-
NIT	46	41.59	2490172.00	0.75	1H-Indole, 6-methyl-
ALI	47	41.92	1977251.00	0.60	2-Tetradecene, (E)-
OxyPH	49	43.01	5332335.00	1.61	trans-Isoeugenol
PH	52	44.20	1761108.00	0.53	2-Allyl-4-methylphenol
OxyPH	57	45.53	15828823.00	4.78	trans-Isoeugenol
OxyAR	58	45.84	3084145.00	0.93	Benzene, 2-(1,1-dimethylethyl)-1,4-dimethoxy-
KET	59	46.61	2287705.00	0.69	(2,5-Dimethoxyphenyl)acetone
EST	62	48.22	2024114.00	0.61	2'-Hydroxy-6'-methoxyacetophenone, acetate
NIT	63	48.49	2439437.00	0.74	thiazolo[5,4-b]pyridin-2-amine, N-ethyl-
AR	67	50.19	6223810.00	1.88	4-Ethylbiphenyl
OxyPH	75	54.32	2795889.00	0.84	Phenol, 2,6-dimethoxy-4-(2-propenyl)-
AR	76	54.59	2697509.00	0.81	Benzene, 1-ethyl-3-(phenylmethyl)-
KET	79	56.31	2306543.00	0.70	2-Hydroxy-4-isopropyl-7-methoxytropone
OxyAR	80	56.55	2249964.00	0.68	Benzene, 2-(1,1-dimethylethyl)-1,4-dimethoxy-
AR	84	57.55	1316973.00	0.40	(E)-Stilbene
AC	85	57.79	1374354.00	0.42	4-Isopropylphenoxyacetic acid
OxyPH	86	58.01	4544993.00	1.37	(E)-2,6-Dimethoxy-4-(prop-1-en-1-yl)phenol
ALI	87	58.72	2432449.00	0.73	Cyclopropane, 1-methyl-2-pentyl-
OxyPH	88	59.16	6511262.00	1.97	Ethanone, 1-(4-hydroxy-3,5-dimethoxyphenyl)-
ALI	91	60.49	1626614.00	0.49	1-Octadecene
ALI	98	62.96	1714129.00	0.52	1-Nonadecene
EST	102	63.74	1171180.00	0.35	Hexadecanoic acid, methyl ester
EST	119	76.84	2193488.00	0.66	Triacontyl acetate

Table S7. Synthetic lignin oligomers (SLOs) representatives for the heavy fraction in FPBO based on the database [122].

SLO	Database No., full name and structure	Information
SLO-1	(125) 1-(4-hydroxyphenyl)-2-phenoxyethanol 	Chemical Formula: C14H14O3 Exact Mass: 230.09 Molecular Weight: 230.26 m/z: 230.09 (100.0%), 231.10 (15.4%), 232.10 (1.7%) Elemental Analysis: C, 73.03; H, 6.13; O, 20.84
SLO-2	(298) 3-Methoxy-4-hydroxy stilbene carboxylic acid 	Chemical Formula: C16H16O4 Exact Mass: 272.10 Molecular Weight: 272.30 m/z: 272.10 (100.0%), 273.11 (17.6%), 274.11 (2.3%) Elemental Analysis: C, 70.58; H, 5.92; O, 23.50

SLO-3	(268) : 1-(4-hydroxyphenyl)-2-(p-tolyloxy)propane-1,3-diol 	Chemical Formula: C16H18O4 Exact Mass: 274.12 Molecular Weight: 274.32 m/z: 274.12 (100.0%), 275.12 (17.5%), 276.13 (1.5%) Elemental Analysis: C, 70.06; H, 6.61; O, 23.33
SLO-4	(126) 4-(1-hydroxy-2-(2-methoxyphenoxy)ethyl)-2-methoxyphenol 	$C_{16}H_{18}O_5$ Chemical Formula: C16H18O5 Exact Mass: 290.12 Molecular Weight: 290.32 m/z: 290.12 (100.0%), 291.12 (17.7%), 292.12 (2.5%) Elemental Analysis: C, 66.20; H, 6.25; O, 27.55
SLO-5	(251) 3-(4-(benzyloxy)-3-methoxyphenyl)propyl 4-(2-oxopropyl)benzoate 	Chemical Formula: C27H28O5 Exact Mass: 432.19 Molecular Weight: 432.52 m/z: 432.19 (100.0%), 433.20 (29.7%), 434.20 (5.3%) Elemental Analysis: C, 74.98; H, 6.53; O, 18.50
SLO-6	(247) 3,3'-dimethoxy-[1,1'-biphenyl]-2,2'-diyl diacetate 	Chemical Formula: C18H18O6 Exact Mass: 330.11 Molecular Weight: 330.34 m/z: 330.11 (100.0%), 331.11 (19.7%), 332.12 (1.9%), 332.11 (1.2%) Elemental Analysis: C, 65.45; H, 5.49; O, 29.06
SLO-7	(300) 1-(4-(5-acetyl-2-hydroxy-3-methoxyphenoxy)-3-methoxyphenyl)ethan-1-one 	$C_{18}H_{18}O_6$ Chemical Formula: C18H18O6 Exact Mass: 330.11 Molecular Weight: 330.34 m/z: 330.11 (100.0%), 331.11 (19.7%), 332.12 (1.9%), 332.11 (1.2%) Elemental Analysis: C, 65.45; H, 5.49; O, 29.06

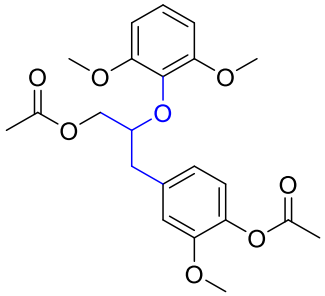
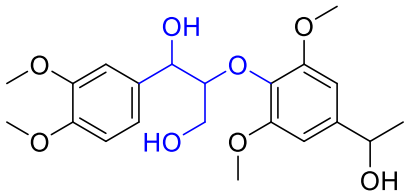
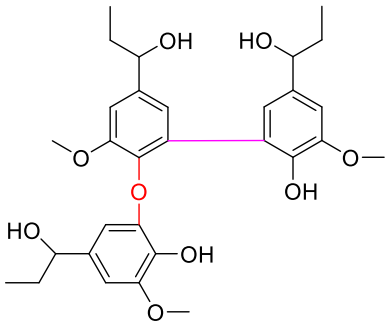
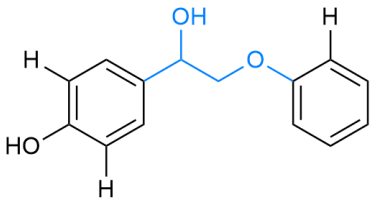
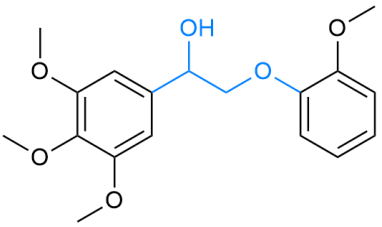
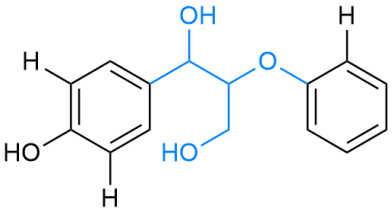
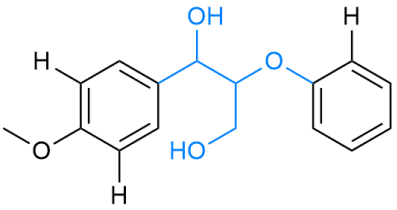
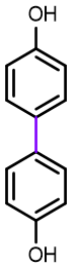
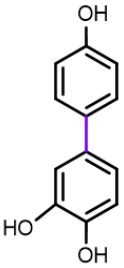
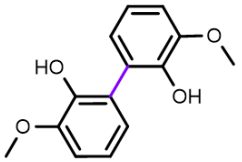
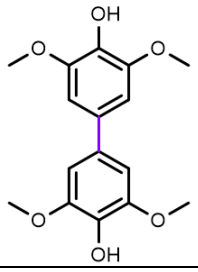
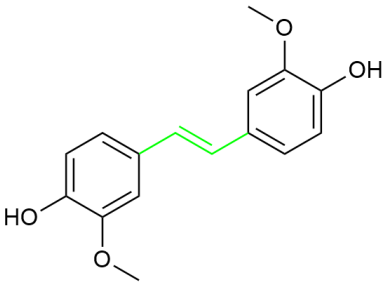
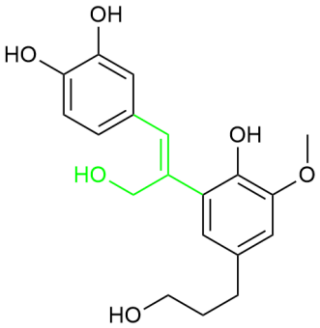
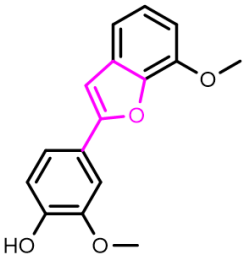
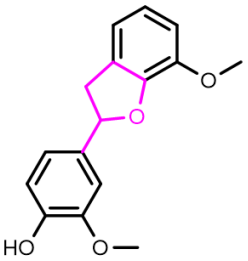
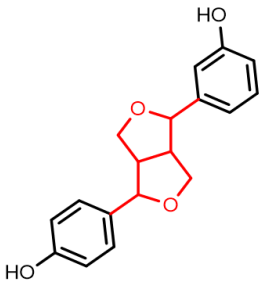
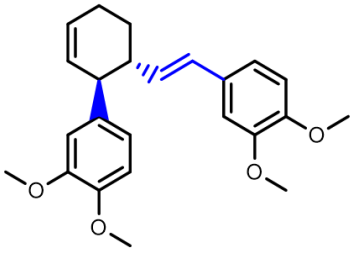
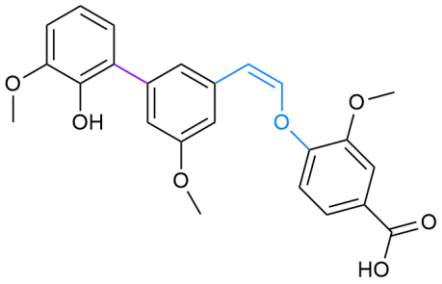
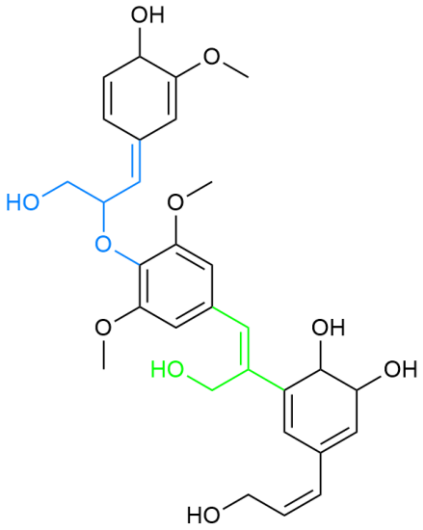
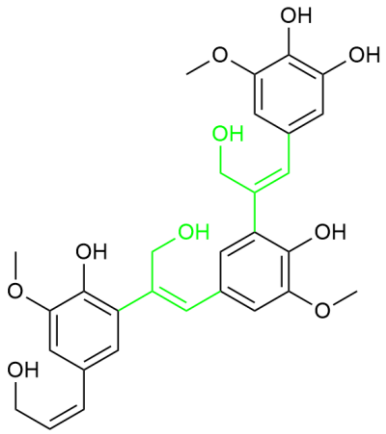
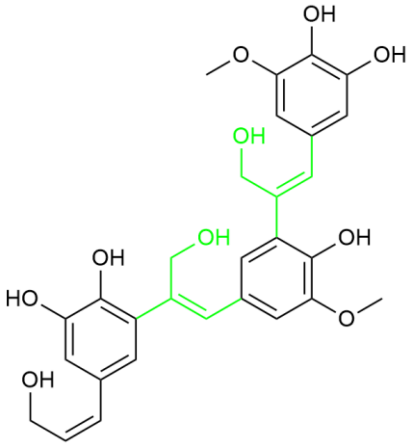
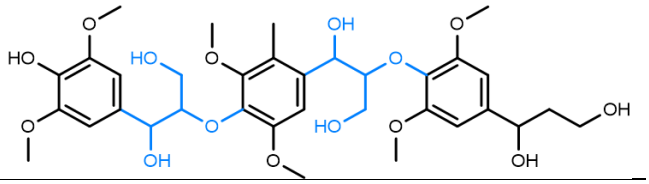
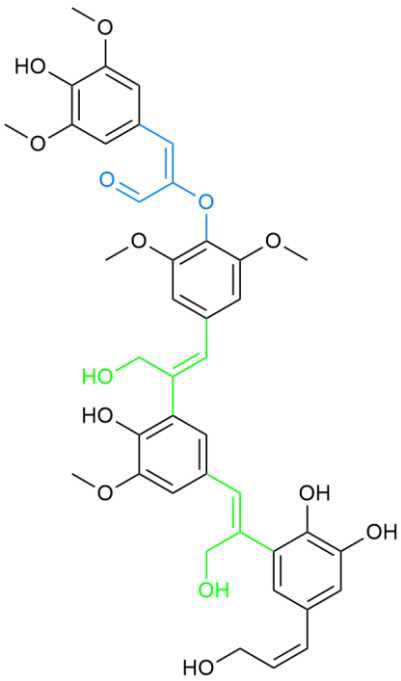
SLO-8	(218) 4-(3-acetoxy-2-(2,6-dimethoxyphenoxy)propyl)-2-methoxyphenyl acetate 	Chemical Formula: C22H26O8 Exact Mass: 418.16 Molecular Weight: 418.44 m/z: 418.16 (100.0%), 419.17 (24.4%), 420.17 (4.5%) Elemental Analysis: C, 63.15; H, 6.26; O, 30.59
SLO-9	(2) 1-(3,4-dimethoxyphenyl)-2-(4-(1-hydroxyethyl)-2,6-dimethoxyphenoxy)propane-1,3-diol 	Chemical Formula: C21H28O8 Exact Mass: 408.18 Molecular Weight: 408.45 m/z: 408.18 (100.0%), 409.18 (23.3%), 410.19 (2.6%), 410.18 (1.6%) Elemental Analysis: C, 61.75; H, 6.91; O, 31.34
SLO-10	(284) 1,1'-(6-hydroxy-6'-(2-hydroxy-5-(1-hydroxypropyl)-3-methoxyphenoxy)-5,5'-dimethoxy-[1,1'-biphenyl]-3,3'-diyl)bis(propan-1-ol) 	Chemical Formula: C30H38O9 Exact Mass: 542.25 Molecular Weight: 542.63 m/z: 542.25 (100.0%), 543.25 (32.4%), 544.26 (7.2%), 545.26 (1.2%) Elemental Analysis: C, 66.40; H, 7.06; O, 26.54

Table S8. Pyrolytic lignin (PL) surrogate molecules used in VLE calculations.

Molecule	Description	Formula	MW (g/mol)	Structure
D-A1	Dimer - β-O-4	$C_{14}H_{14}O_3$	230.26	
D-A2	Dimer - β-O-4	$C_{18}H_{22}O_6$	334.37	
D-A3	Dimer - β-O-4	$C_{15}H_{16}O_4$	260.29	
D-A4	Dimer - β-O-4	$C_{16}H_{18}O_4$	274.32	
D-B1	Dimer - Biphenyl	$C_{12}H_{10}O_2$	186.21	
D-B2	Dimer - Biphenyl	$C_{12}H_{10}O_3$	202.21	

D-B3	Dimer - Biphenyl	$C_{14}H_{14}O_4$	246.26	
D-B4	Dimer - Biphenyl	$C_{16}H_{18}O_6$	306.31	
D-C1	Dimer - Stilbene	$C_{16}H_{16}O_4$	272.3	
D-C2	Dimer - Stilbene	$C_{19}H_{20}O_6$	344.36	
D-D1	Dimer - Phenylcoumaran	$C_{16}H_{14}O_4$	270.28	
D-D2	Dimer - Phenylcoumaran	$C_{16}H_{16}O_4$	272.3	

D-E1	Dimer - Resinol	$C_{18}H_{18}O_4$	298.34	
D-F1	Dimer - Bridging double bond	$C_{24}H_{28}O_4$	380.48	
TR1	Trimer - mixed links	$C_{24}H_{22}O_7$	420.46	
TR2	Trimer - mixed links	$C_{30}H_{36}O_{10}$	556.61	

TR3	Trimer - stilbene links	$C_{30}H_{32}O_{10}$	552.58	
TR4	Trimer - stilbene links	$C_{29}H_{30}O_{10}$	538.55	
TR5	Trimer - β-O-4 links	$C_{34}H_{46}O_{15}$	694.73	
TE1	Tetramer - mixed links	$C_{41}H_{43}O_{14}$	758.77	

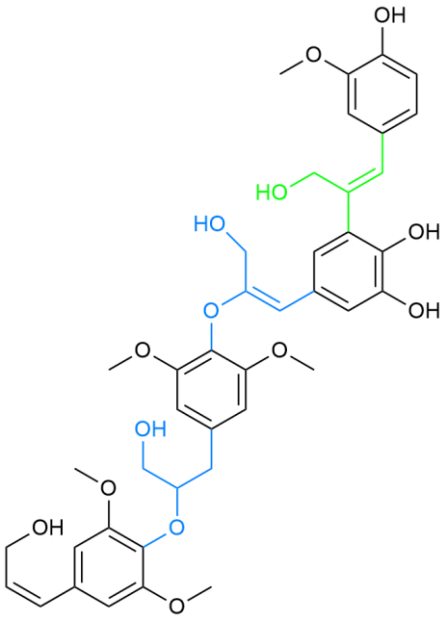
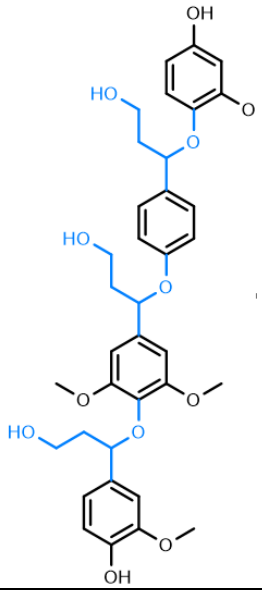
TE2	Tetramer - mixed links	$C_{41}H_{46}O_{14}$	762.81	
TE3	Tetramer - β-O-4 links	$C_{37}H_{44}O_{12}$	680.7	

Table S9. Functional group information for UNIFAC-DMD model of tested PL surrogates for VLE calculations (continued on the next page).

Group code Aspen	1005	1010	1030	1035	1055	1065	1100	1105	1150	1155	1160
Molecule/FC Group	>CH-	>CH2	Cyclic -c-CH	Cyclic -c-CH2	-CH=C<	-CH=CH-	Ar >C=	Ar -CH=	Ar C-CH<	Ar C-CH2	Ar C-CH3
D-A1	0	0	0	0	0	0	1	9	1	0	0
D-A2	0	0	0	0	0	0	5	6	1	0	0
D-A3	0	0	0	0	0	0	1	9	1	0	0
D-A4	0	0	0	0	0	0	2	9	1	0	0
D-B1	0	0	0	0	0	0	2	8	0	0	0
D-B2	0	0	0	0	0	0	2	7	0	0	0
D-B3	0	0	0	0	0	0	4	6	0	0	0
D-B4	0	0	0	0	0	0	6	3	0	0	0
D-C1	0	0	0	0	0	1	4	6	0	0	0
D-C2	0	2	0	0	1	1	4	5	0	0	0
D-D1	0	0	0	0	0	0	6	7	0	0	0
D-D2	0	0	0	1	0	0	5	4	0	0	0
D-E1	0	0	2	0	0	0	2	8	0	0	0
D-F1	1	2	0	0	0	2	5	6	1	0	0
TR1	0	0	0	0	0	0	4	9	1	0	0
TR2	0	3	0	0	1	2	7	7	0	0	0
TR3	0	3	0	0	2	1	8	6	0	0	0
TR4	0	3	0	0	2	1	7	6	0	0	0
TR5	0	4	0	0	0	0	8	5	3	0	1
TE1	1	3	0	0	2	1	12	8	0	0	0
TE2	1	4	0	0	1	1	11	9	0	1	0
TE3	0	2	0	0	0	0	10	12	0	0	0

Table S9 (continued).

Group code Aspen	1200	1210	1350	1450	1605	1610	1615	1625	1630	1955
Molecule/FC Group	1º -OH	2º -OH	Phenolic =C-OH	Aldehyde O=CH-	Ether -CH-O-	Ether -CH2-O-	Ether -CH3-O-	Cyclic c-CH2-O-(CH2)5	Cyclic c-(CH2)5-O-(CH2)5	Acid -COOH
D-A1	0	1	1	0	0	1	0	0	0	0
D-A2	0	1	0	0	0	1	4	0	0	0
D-A3	1	1	1	0	1	0	0	0	0	0
D-A4	1	1	0	0	1	0	1	0	0	0
D-B1	0	0	2	0	0	0	0	0	0	0
D-B2	0	0	3	0	0	0	0	0	0	0
D-B3	0	0	2	0	0	0	2	0	0	0
D-B4	0	0	2	0	0	0	4	0	0	0
D-C1	0	0	2	0	0	0	2	0	0	0
D-C2	2	0	3	0	0	0	1	0	0	0
D-D1	0	0	1	0	0	0	2	0	1	0
D-D2	0	0	1	0	0	0	2	1	0	0
D-E1	0	0	2	0	0	0	0	2	0	0
D-F1	0	0	0	0	0	0	4	0	0	0
TR1	0	0	1	0	1	0	3	0	0	1
TR2	3	0	3	0	1	0	3	0	0	0
TR3	3	0	4	0	0	0	3	0	0	0
TR4	3	0	5	0	0	0	2	0	0	0
TR5	3	3	1	0	0	0	6	0	0	0
TE1	3	0	4	1	1	0	5	0	0	0
TE2	4	0	3	0	2	0	5	0	0	0
TE3	3	0	2	0	3	0	4	0	0	0

Table S10. Calculated FPBO properties for each PL surrogate molecule.

	Resulting elemental composition for each PL surrogate			Density (kg/m³)	
Molecule	**C (wt.%)**	**H (wt.%)**	**O (wt.%)**	**Rackett**	**PR-BM**
D-A1	60.05%	7.11%	32.84%	1138.2656	932.3424
D-A2	54.93%	7.41%	37.66%	1191.7363	1165.7032
D-A3	57.72%	7.15%	35.13%	1070.0465	1023.6504
D-A4	57.08%	7.61%	35.31%	1081.0097	1048.0148
D-B1	62.74%	6.67%	30.60%	1132.3596	933.4997
D-B2	58.98%	6.40%	34.61%	1166.9004	968.2771
D-B3	57.15%	6.86%	35.99%	1022.0521	865.5375
D-B4	53.75%	6.98%	39.27%	1110.0320	899.7613
D-C1	58.55%	6.98%	34.47%	1110.0760	913.2757
D-C2	56.14%	6.60%	37.26%	986.6422	874.0527
D-D1	58.87%	6.55%	34.58%	1003.0837	931.1446
D-D2	58.55%	6.98%	34.47%	1110.0760	913.2757
D-E1	59.71%	7.08%	33.21%	1021.3365	17.4220
D-F1	61.73%	7.90%	30.37%	854.2145	9.5101
TR1	57.12%	6.57%	36.31%	885.8089	796.0708
TR2	55.56%	6.50%	37.94%	813.1505	6.4580
TR3	55.41%	6.72%	37.87%	842.8522	7.7575
TR4	55.24%	6.35%	38.41%	867.2863	9.8901
TR5	50.17%	7.25%	42.58%	513.9193	4.3976
TE1	55.19%	6.61%	38.20%	1290.0497	6.6934
TE2	55.08%	6.77%	38.15%	36.0257	6.6510
TE3	55.31%	7.34%	37.35%	442.6522	4.9659
No-PL	39.53%	8.65%	51.82%	1138.2656	932.3424

Table S11. Mean absolute relative error (MARE) for VLE predictions of FPBO mixture with each PL surrogate.

	MARE Temperature			MARE Water Content		
Molecule	**IDEAL**	**UNIF-DMD**	**PR-BM**	**IDEAL**	**UNIF-DMD**	**PR-BM**
D-A1	33.0196	38.6434	40.1687	47.5533	28.4737	41.1730
D-A2	31.2078	34.6312	41.9792	70.3222	39.6143	42.9974
D-A3	28.6250	34.5833	39.5473	66.2733	52.6287	36.8064
D-A4	27.2610	35.5405	38.6714	61.0266	32.2640	32.3980
D-B1	24.0941	28.0067	36.4523	46.3359	42.5731	38.5895
D-B2	26.4744	29.1407	36.3762	53.8411	72.7062	33.3512
D-B3	26.3998	30.7013	47.7660	35.8046	25.6718	47.2395
D-B4	27.7626	32.8646	45.6408	61.6630	39.5163	41.6625
D-C1	30.9028	34.9199	43.9684	47.6739	36.7282	42.5675
D-C2	29.0210	33.4733	43.7720	64.0354	74.1724	36.1208
D-D1	27.1661	33.8156	43.8632	59.4255	32.0042	41.2142
D-D2	30.9013	35.9461	43.9670	47.6647	29.3025	42.5662
D-E1	48.6654	56.6471	56.6666	99.2540	99.5703	45.4840
D-F1	27.5876	36.2979	54.8913	52.1719	42.2971	45.4357
TR1	34.3887	40.3181	61.1881	76.4166	58.2254	43.8478
TR2	30.4264	35.4021	63.4770	39.0273	27.5920	48.9785
TR3	30.2304	33.4787	58.3260	52.7307	52.0340	45.9892
TR4	27.8004	35.0450	58.9034	62.4978	75.8585	46.2730
TR5	27.8428	33.5875	-	58.3777	38.7538	-
TE1	30.7157	35.7312	64.7103	60.5150	50.2330	47.5993
TE2	129.3392	-	64.0232	76.6204	-	47.3106
TE3	29.9557	34.7979	-	50.6119	29.5275	-

Figure S2. Temperature and water content profiles for the rest of the top five best PL surrogates.

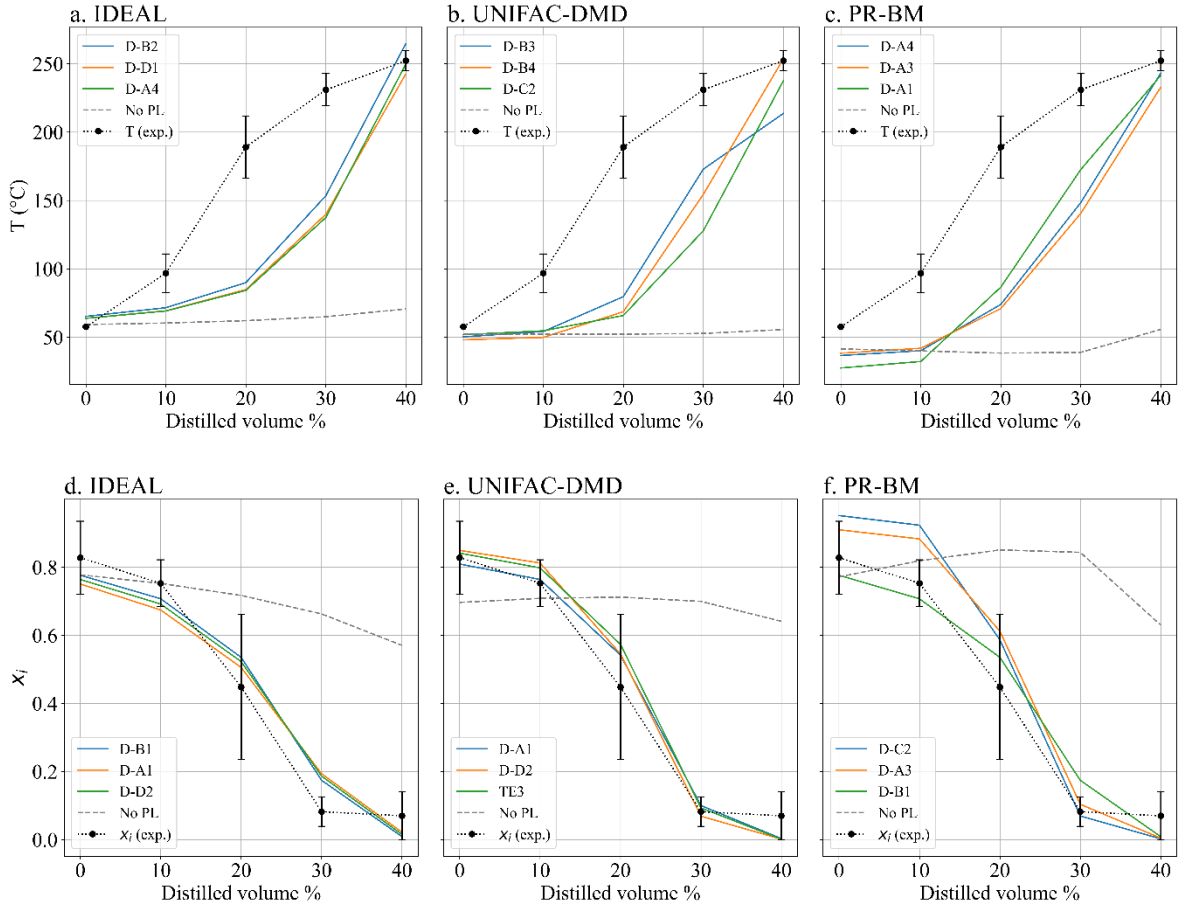


Figure S3. Unknowns profile for other selected PL surrogate molecules.

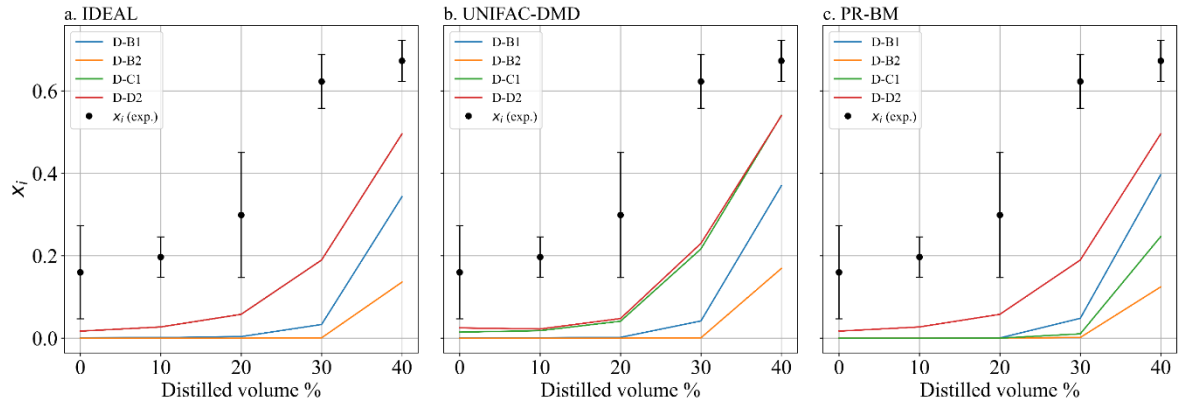


Table S12. Functional groups of PL surrogate molecules used in LLE calculations.

Group code Aspen	1005	1010	1030	1035	1055	1065	1100	1105	1150	1155	1160
Molecule/FC Group	>CH-	>CH2	Cyclic -c-CH	Cyclic -c-CH2	-CH=C<	-CH=CH-	Ar >C=	Ar -CH=	Ar C-CH<	Ar C-CH2	Ar C-CH3
A1	0	0	0	0	0	0	6	7	0	0	0
A2	0	0	0	1	0	0	6	4	0	0	0
B1	0	0	0	0	0	0	4	6	0	0	0
B2	0	0	0	0	0	0	3	7	0	0	0
B3	0	0	0	0	0	0	6	4	0	0	0
B4	0	0	0	0	0	0	2	8	0	0	0
B5	0	0	0	0	0	0	2	7	0	0	0
C1	0	0	2	0	0	0	2	8	0	0	0
D1	0	0	0	0	0	0	1	10	1	0	0
D2	0	0	0	0	0	0	5	6	1	0	0
E1	0	0	0	0	0	1	6	4	0	0	0
E2	0	2	0	0	1	1	4	5	0	0	0
E3	0	2	0	0	1	1	5	4	0	0	0
F1	0	4	0	0	0	0	8	5	3	0	1
G1	0	3	0	0	2	1	8	6	0	0	0
G2	0	3	0	0	2	1	7	6	0	0	0
H1	0	0	2	2	0	2	6	6	0	0	0
I1	1	3	0	0	2	1	12	8	0	0	0
I2	1	4	0	0	1	1	11	9	0	0	0
I3	0	2	0	0	0	0	10	12	0	0	0

Table S12 (continued).

Group code Aspen	1200	1210	1350	1450	1605	1610	1615	1625	1630
Molecule/FC Group	**1º -OH**	**2º -OH**	**Phenolic =C-OH**	**Aldehyde O=CH-**	**Ether -CH-O-**	**Ether -CH2-O-**	**Ether -CH3-O-**	**Cyclic c-CH2-O-(CH2)5**	**Cyclic c-(CH2)5-O-(CH2)5**
A1	0	0	1	0	0	0	2	0	1
A2	0	0	1	0	0	0	2	1	0
B1	0	0	2	0	0	0	2	0	0
B2	0	0	2	0	0	0	1	0	0
B3	0	0	2	0	0	0	4	0	0
B4	0	0	2	0	0	0	0	0	0
B5	0	0	3	0	0	0	0	0	0
C1	0	0	2	0	0	0	0	2	0
D1	0	1	1	0	0	1	0	0	0
D2	0	1	0	0	0	1	4	0	0
E1	0	0	2	0	0	0	2	0	0
E2	2	0	3	0	0	0	1	0	0
E3	2	0	3	0	0	0	1	0	0
F1	3	3	1	0	0	0	6	0	0
G1	3	0	4	0	0	0	3	0	0
G2	3	0	5	0	0	0	2	0	0
H1	0	0	0	0	0	0	4	0	0
I1	3	0	4	1	1	0	5	0	0
I2	4	0	3	0	2	0	5	0	0
I3	3	0	2	0	3	0	4	0	0

Figure S4. The scheme of simulation flowsheet used in Aspen Plus™ V14. The mass flow rates displayed are for the BE_1 system.

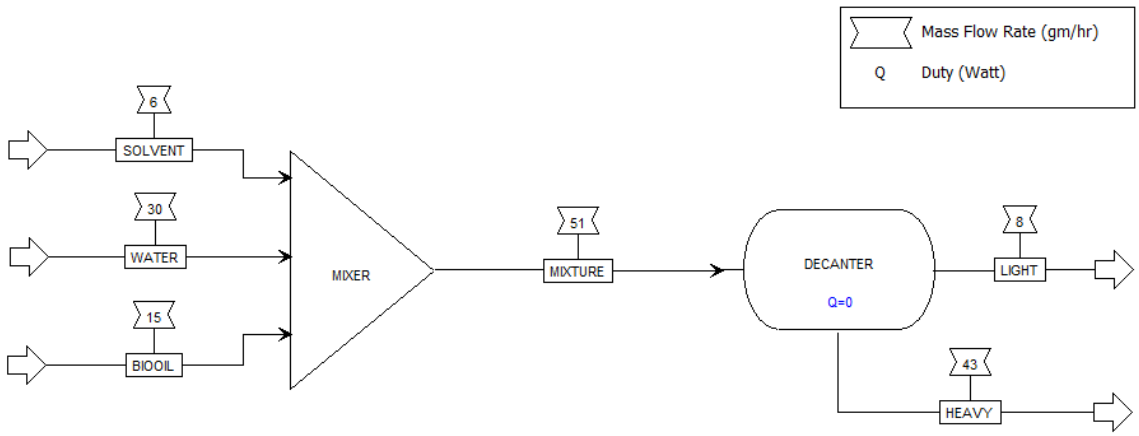


Figure S5. Variation of solvent (BE) and water (WE) vs Ratio O/A for case 1: a-d, and case 2: e-h.

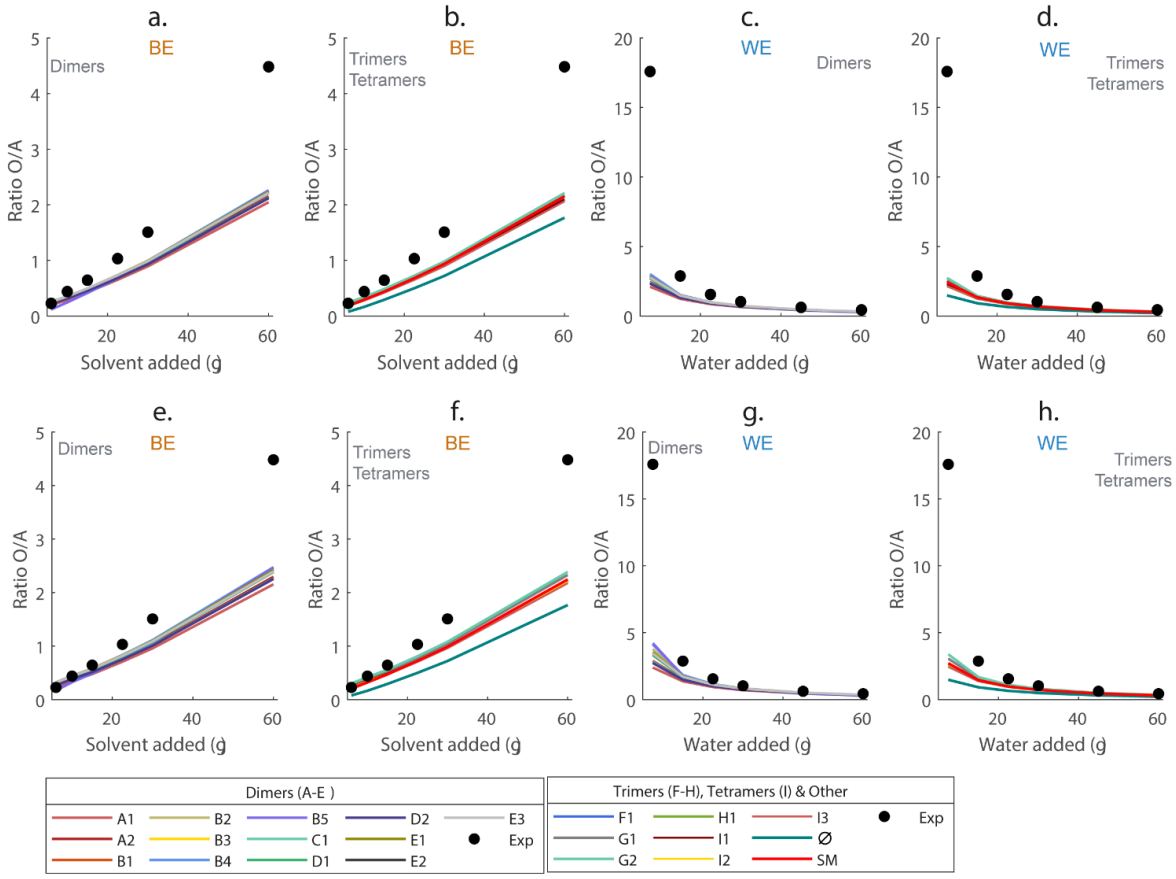


Figure S6. Variation of solvent (BE) and water (WE) vs K_{ow} of Bio-oil for case 1: a-d, a and case 2: e-h.

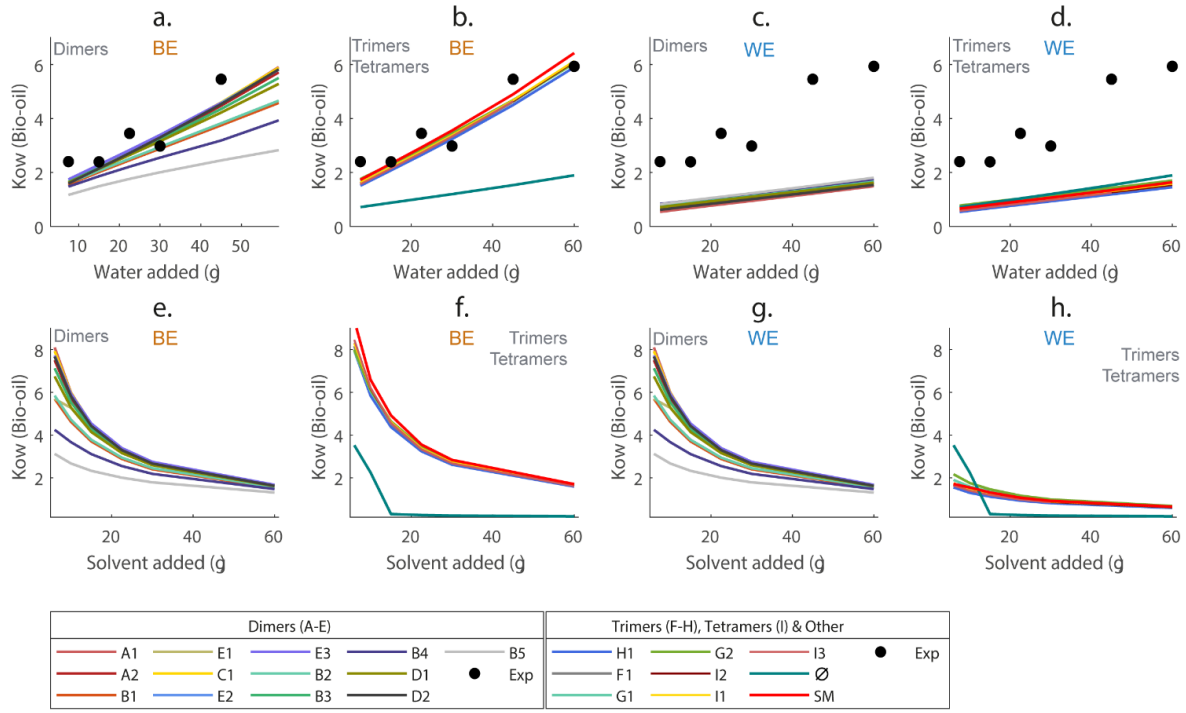


Figure S7. Variation of solvent (BE) and water (WE) vs Kow variation of Water for case 1: a-d, and case 2: e-h.

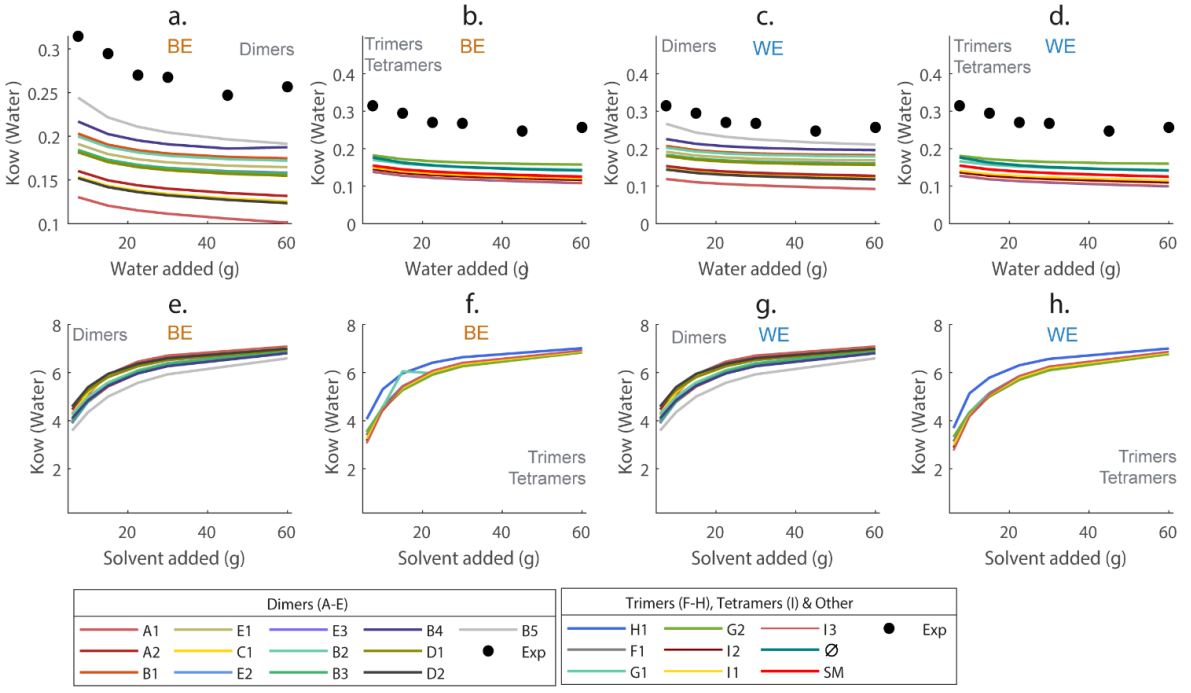


Figure S8. Variation of solvent (BE) and water (WE) vs Kow variation of n-butanol (solvent) for case 1: a-d, and case 2: e-h.

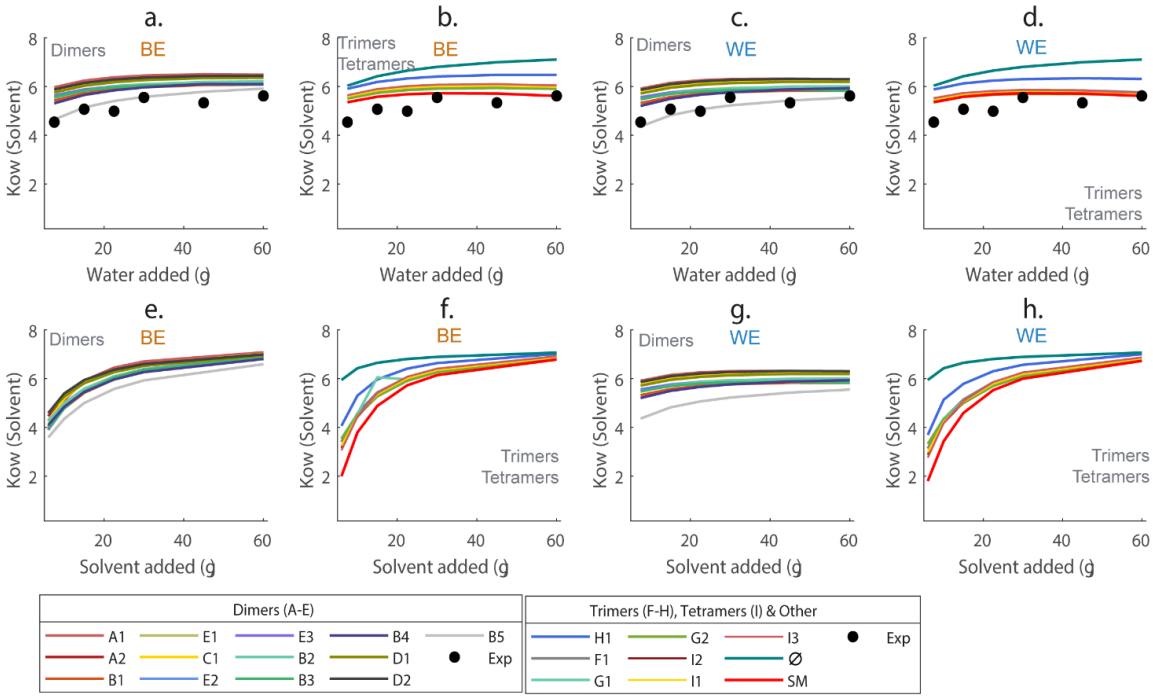


Figure S9. Variation of solvent (BE) and water (WE) vs Kow variation of Pyrolignins (PL represented by the LOs structures) for case 1: a-d, and case 2: e-h.

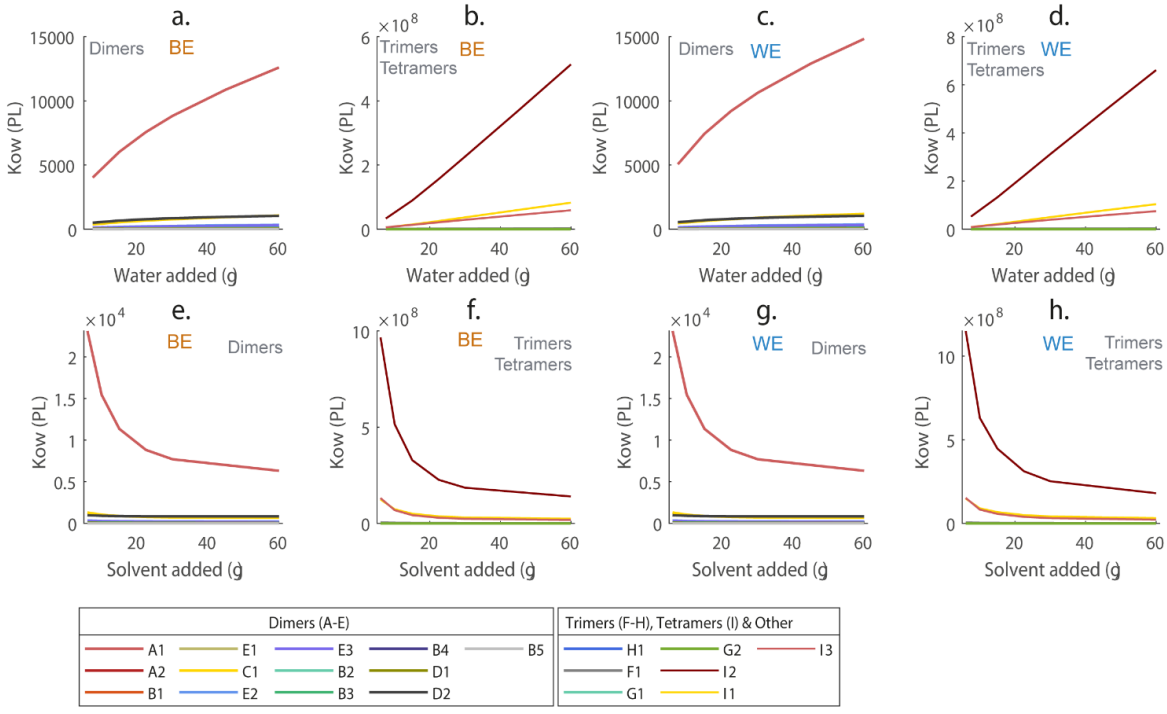


Figure S10. Variation of solvent (BE) and water (WE) vs variation of concentration of Pyrolignins (PL represented by the LOs structures) in the Organic phase for case 1: a-d and case 2: e-h.

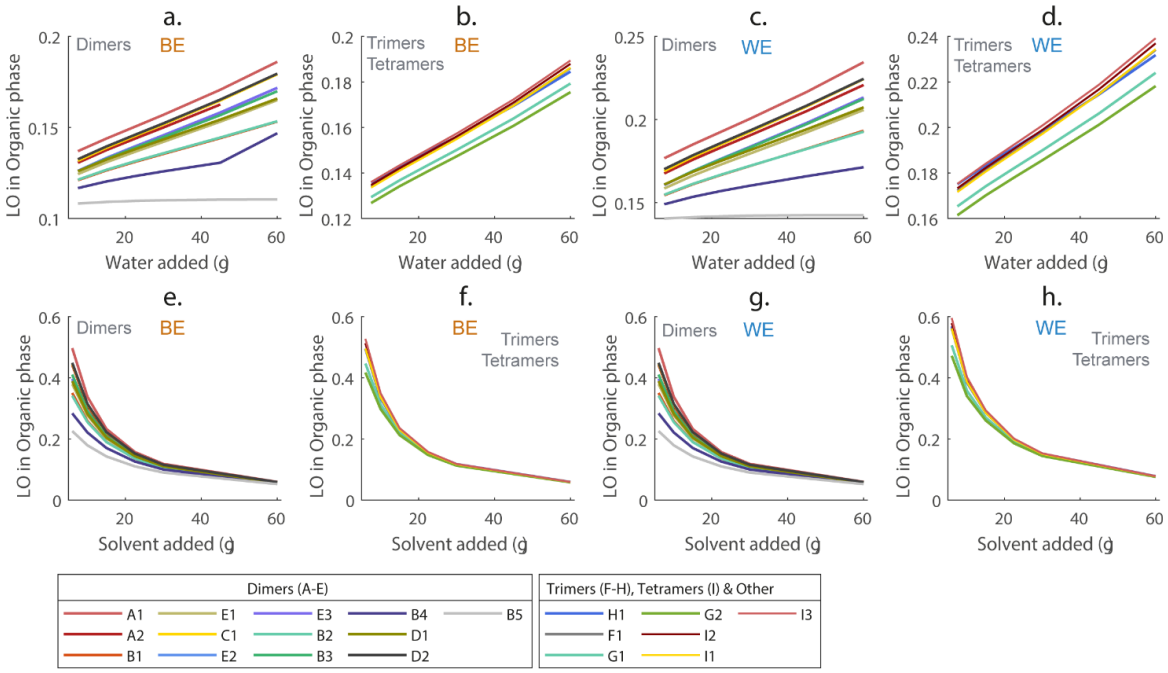


Figure S11. Variation of solvent (BE) and water (WE) vs variation of concentration of Pyrolignins (PL represented by the LOs structures) in the aqueous phase for case 1: a-d, and case 2: e-h.

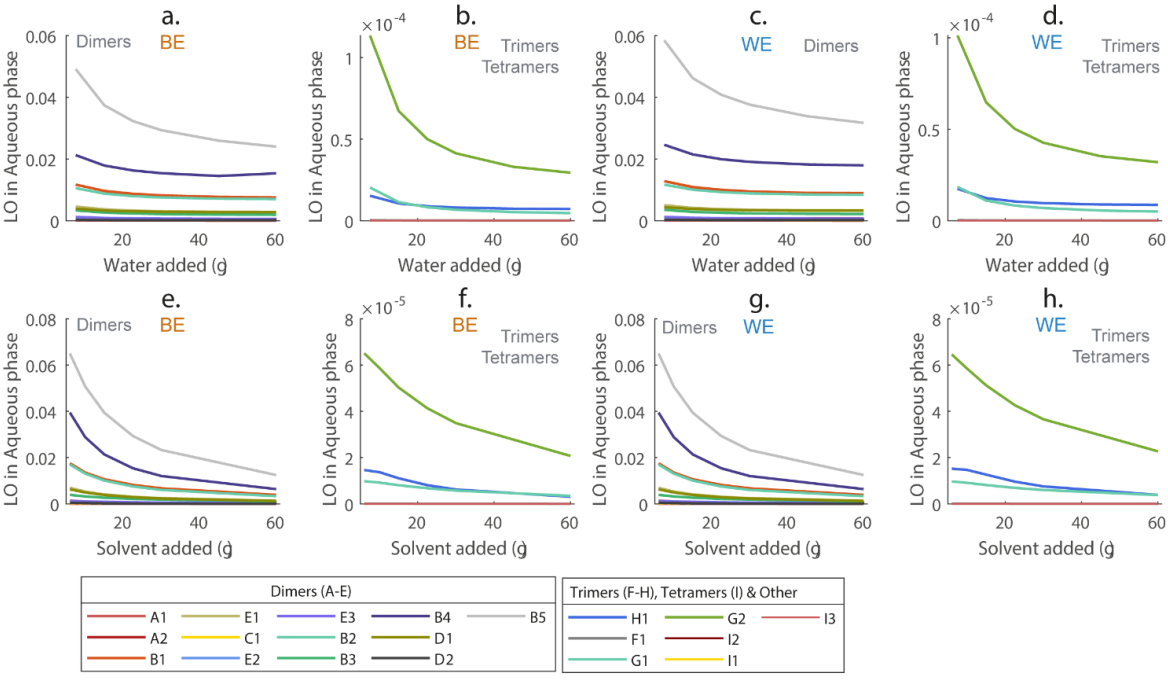


Figure S12. Variation of solvent (BE) and water (WE) Kow of Levoglucosan for case 1: a-d, and case 2: e-h.

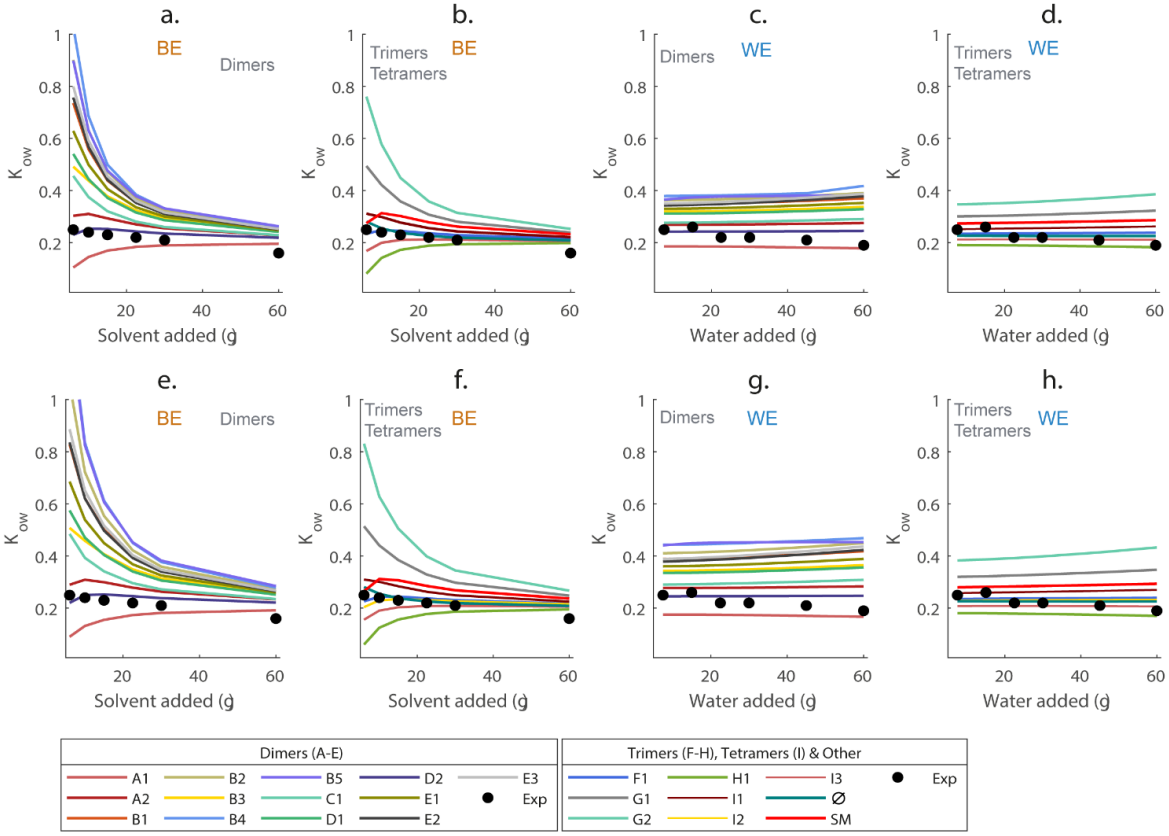
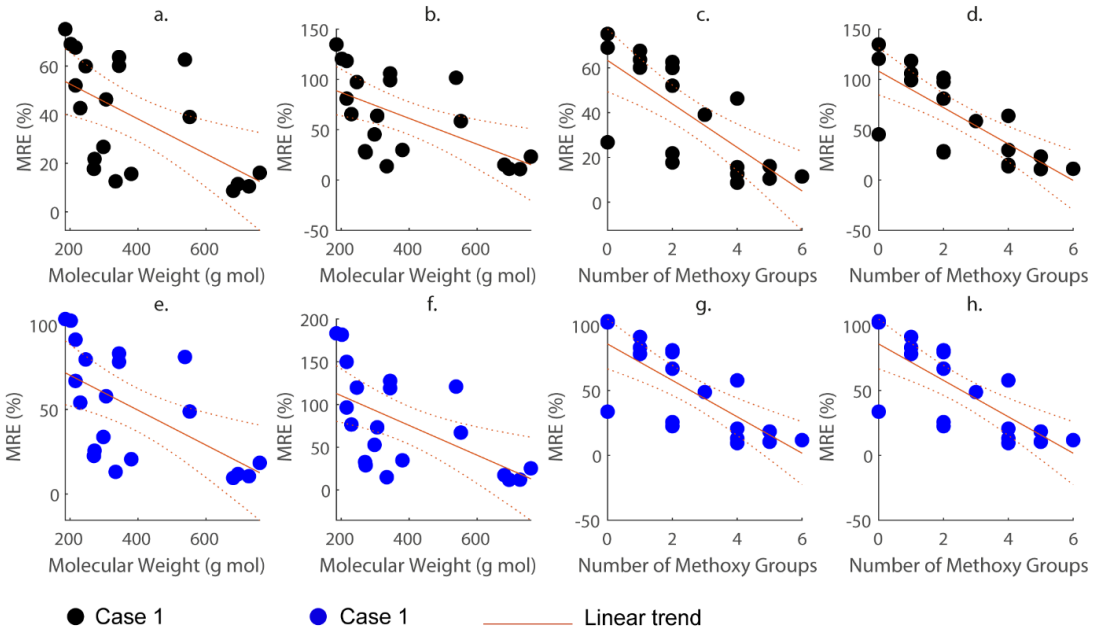


Figure S13. Trends of Types and Number of functional groups on Kow of Levoglucosan prediction for case 1: a-d, and case 2: e-h.



List of Figures

List of Tables

List of Figures and Tables in Supporting Information