2020

CHARACTERIZING AND PREDICTING THE ANTIMICROBIAL PROPERTIES OF LIGNIN DERIVATIVES

Ryan Kalinoski

University of Kentucky, rmka232@uky.edu
Author ORCID Identifier:
 https://orcid.org/0000-0001-8774-275X
Digital Object Identifier: https://doi.org/10.13023/etd.2020.455

CHARACTERIZING AND PREDICTING THE ANTIMICROBIAL PROPERTIES OF
LIGNIN DERIVATIVES

_____

DISSERTATION
_____

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
Colleges of Agriculture and Engineering
at the University of Kentucky

By
Ryan Kalinoski
Lexington, Kentucky
Advisor: Dr. Jian Shi, Assistant Professor of Biosystems and Agricultural Engineering
Lexington, Kentucky
2020

ABSTRACT OF DISSERTATION

CHARACTERIZING AND PREDICTING THE ANTIMICROBIAL PROPERTIES OF
LIGNIN DERIVATIVES

Due to the overuse of antibiotics in our society, there has been a steady rise in highly antimicrobial-resistant bacteria in the last decade. This has created a renewed interest in natural phenolic compounds for antimicrobial discovery amongst the scientific community. To this end, lignin is the most abundant naturally occurring phenolic polymer on earth and has already been known to have antimicrobial properties due to its polyphenolic structure. In addition, lignin is considered a major waste product for lignocellulosic biorefineries, and its valorization into value-added products will generate extra profit for a biorefinery, making biofuels less expensive, increasing their marketability as an alternative to fossil fuels. However, the retention of lignin's antimicrobial properties in different materials, as depolymerized products, or even the prediction of their antimicrobial properties is not well understood in the literature.

Much work has utilized lignin as a functional polymer in a variety of composites and materials, but their antimicrobial properties have not been as widely explored. Therefore, ionic liquids were used in the facile preparation of cellulose-based hydrogels, and the addition of different lignocellulosic components (lignin and xylan) or the use of whole biomass (poplar and sorghum) were evaluated for their effects on hydrogel properties (mechanical and antimicrobial). The addition of both lignin and xylan improved hydrogel mechanical strength/stiffness, and lignin-containing hydrogels showed retained antimicrobial properties when screened against the target organism (*Escherichia coli*). Utilizing raw biomass provided increased mechanical strength (poplar), similar water retention abilities (poplar and sorghum), and retained antimicrobial properties (poplar). These results indicate that the different components of lignocellulose can be used to fine tune the properties of cellulose-based hydrogels and that lignin can confer its antimicrobial properties when incorporated into hydrogels.

The antimicrobial properties of different lignin depolymerization products were explored using a reductive and oxidative depolymerization method to produce phenolic rich lignin-based bio-oils. Purified alkali-enzymatic corn stover lignin (AEL) was depolymerized by catalytic transfer hydrogenolysis using supercritical ethanol and a Ru/C catalyst, generating a bio-oil stream at high yields. Sequential extraction was used to fractionate the bio-oil into five fractions with different phenolic compositions using

hexane, petroleum ether, chloroform, and ethyl acetate. Antimicrobial properties of the bio-oils were screened against Gram-positive/negative bacteria and yeast by examining microbial growth inhibition. The monomers in the bio-oil fractions contained primarily alkylated phenols, hydrogenated hydroxycinnamic acid derivatives, syringol and guaiacol-type lignins created from reductive cleavages of ether linkages. After sequential extraction, the lignin derived compounds were fractionated into groups depending on solvent polarity. Results suggest that the total monomer concentration and the presence of specific monomers (i.e., syringyl propane) may correlate to the antimicrobial activity of lignin depolymerization products, but the exact mode of action or antimicrobial activity caused by the complex mixtures of monomers and unidentified oligomers remains unclear.

The same AEL lignin was depolymerized through oxidative procedures using peracetic acid, and its applications as an antibiotic replacement in the fuel ethanol industry were explored. The resulting bio-oil had a low degree of depolymerization that mostly produced unidentifiable lignin oligomers. Nonetheless, this bio-oil displayed highly selective antimicrobial properties, with up to 90% inhibition of commercially sampled lactic acid bacteria (LAB) at 4 mg/ml and no inhibition of yeast. Using the bio-oil (4 mg/ml) as an alternative antibiotic treatment during simultaneous-saccharification and fermentation of raw corn starch showed an 8% increase in ethanol production at a yeast to LAB ratio of 1:100, compared to untreated contaminated controls. The ability of the bio-oil to improve ethanol yields clearly shows its efficacy as an alternative antibiotic and that depending on depolymerization method lignin derivates can display a variety of useful antimicrobial properties/applications.

The final study was the first attempt in the literature to predict the antimicrobial properties of lignin derivatives using quantitative structure−activity relationship (QSAR) models. First, the open-access database ChEMBL, with non-lignin specific compounds, was used to create datasets of compounds with MIC activity measurements against both *B. subtilis* and *E. coli*. Machine learning algorithms were used to develop the QSARs for the large ChEMBL datasets and were found to underpredict the antimicrobial activity of actual lignin compounds. Conversely, as metanalysis of the literature containing MIC data of lignin derivatives were used to build QSAR models with ordinary least square regressions (OLS). An accurate QSAR model for *E. coli* was not found, but a satisfactory model was obtained for the *B. subtilis* metanalysis dataset. Molecular Operation Environment (MOE)-type descriptors and the number of aliphatic carboxylic acid groups showed strong correlations to the MIC values (R^2 of 0.759). Comparatively, an additional dataset was experimentally derived by screening 25 lignin monomers and three dimers against *B. subtilis* by measuring bacterial load difference (BLD). This datasets QSAR, using OLS, found that MOE-type descriptors and the number of aromatic hydroxyl groups were better predictors of BLD (R^2 of 0.831). Thus, the smaller datasets highlighted how the variability in antimicrobial measurements and the specific compounds used will impact the predictive nature of the resulting QSARs. Overall, this entire work provides critical knowledge and guidance on using lignin as an antimicrobial source in different industrial processes/products and the identification of lignin derivatives with enhanced activity.

Ryan Kalinoski

(Name of Student)

11/23/2020

Date

CHARACTERIZING AND PREDICTING THE ANTIMICROBIAL PROPERTIES OF
LIGNIN DERIVATIVES

By
Ryan Kalinoski

Jian Shi	
Director of Dissertation	

Donald Colliver	
Director of Graduate Studies	

11/23/2020	
Date	

DEDICATION

I dedicate this dissertation to my mother, Donna Kalinoski, for her endless and unconditional love, encouragement, and support. I could not have done any of this without you.

ACKNOWLEDGMENTS

Firstly, I would like to thank and express my deepest appreciation for my advisor Dr. Jian Shi, whose ceaseless support, guidance, and encouragement made this work possible. The skills and personal growth I have gained through Dr. Shi's excellent mentorship will help propel my professional career and has made me a more confident researcher/person. I hope to exemplify the high quality of scholarship, mentorship, and pure kindness Dr. Shi has shown me to any future mentees I might have in my professional career.

I would like to thank my Dissertation Committee, and outside reader, respectively, Dr. Michael Montross, Dr. Sue Nokes, Dr. Qing Shao, Dr. Michal Flythe, and Dr. Jesse Hoagg. All of whom provided timely and instructive comments and evaluation at every stage of the dissertation process, allowing me to complete this project on schedule. I would also like to thank Dr. Bert Lynn, Dr. Justin Mobley, and their students for their constant support and guidance on the chemistry aspects of my dissertation work.

I would like to thank my current and former lab colleagues, Dr. Joseph Stevens, Dr. Wenqi Li, Dr. Lalitendu Das, Enshi Liu, Jameson Hunter, Abisola Olayeye, Makua Vin-Nnajiofor, Can Liu, Dr. Binling Ai, Yuxuan Zhang, and Joseph Woomer for their invaluable support and friendship. I would like to also thank all the faculty, staff, and graduate/undergraduate students in the Biosystems and Agricultural Engineering Department for their assistance and kindness that made the department feel like a family and such a joy to be apart of.

Finally, I cannot thank my family enough for getting me through this wild ride we call life. For my mother and fathers' support, love, and guidance that made me the

person I am today. To my sister, Sara, for her unyielding love and support throughout my life, especially for taking on the role of "older" sibling on occasion despite being younger than me. To my dogs Babei and Dexter for their unconditional companionship and joy that brightened my days. Also, to my beautiful wife, Devan, for having the patience to follow me through my career in academia, for her constant love/support, and willingness to be my partner in the good times and bad.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Lignin: origin, chemistry, and extraction

Chemistry, Structure, and Sources of Lignocellulose

Lignocellulosic biomass is produced most abundantly as a crop or residue from perennial herbaceous plants and woody crops. Forest residues and municipal organic wastes from agriculture and pulp/paper industries are other significant sources of biomass [1]. The primary polymers that comprise lignocellulosic biomass are located in the plant cell wall and are classified as cellulose, hemicellulose and lignin [2]. Total lignocellulosic dry matter consists of 30-50% cellulose, 20-40% hemicellulose, and 15-25% lignin [3], which varies depending on the lignocellulosic feedstock. The three major polymers of lignocellulose are intertwined together to form a complex structure used for plant rigidity, flexibility, and defense [4]. **Table 1.1** summarizes the differences in chemical composition and structure of cellulose, hemicellulose, and lignin.

Table 1.1: Structure and chemical composition major polymers in plant cell walls, adapted from Chen [4]

	Cellulose	Hemicellulose	Lignin
Subunits	D-Pyran glucose unit	D-Xylose, mannose, L-arabinose, galactose, glucuronic acid	p-coumaryl (H), coniferyl (G), and sinapyl (S) alcohols
Bonds between subunits	β-1,4-Glycosidic bonds	β-1,4-Glycosidic bonds in main chains; β-1.2-, β-1.3-, β-1.6- glycosidic bonds in side chains	Mostly β-O-4 (β-aryl ether), β−β (resinol), and β-5 (phenylcoumaran) and other various C-C and ether bonds
Bonds between three components	Chemical bond with hemicellulose hydrogen bond with cellulose	Chemical bond with lignin and hydrogen bond with cellulose	Hydrogen bonding with cellulose/hemicellulose and chemical bonding with hemicellulose
Composition	Linear three-dimensional with crystalline and amorphous region	Three-dimensional, heterogenous	Amorphous, nonlinear, heterogenous, three dimensional
Degree of polymerization	10,000 and above	100-200	~4,000

Cellulose is a linear polymer composed of glucose (C_6 sugar) residues bonded by β-1,4-glycosidicbonds, with cellobiose as the basic coupling unit. This allows for cellulose's structure to be a flat sheet that can be packed as several strands into crystallin fibrils [5]. Cellulose has the highest degree of polymerization compared to the other components of lignocellulose, which makes it less flexible and insoluble in most solvents [5]. Hemicellulose is a branched polymer comprised of C_5 and C_6 sugars that have acetyl, methyl, cinnamic, glucuronic and galacturonic acid functional groups that make up a variety of subunits: D-xylose, mannose, L-arabinose, galactose and glucuronic acid [4, 6]. Hemicelluloses are non-covalently bound to the surface of cellulose fibrils, acting as a matrix material holding them together [5, 7]. Hemicellulose and cellulose are both

carbohydrate polymers that have repeating units of monomers that form a linear and nonlinear three-dimensional structure, while lignin is an amorphous structure that is highly inhomogeneous and nonlinear.

Lignin is an amorphous polyphenolic complex that is primarily composed of polymerized *p*-coumaryl (H), coniferyl (G), and sinapyl (S) alcohols [8]. These base units are randomly bonded together by ether linkages such as β-O-4, α-O-4, and 4-O-5, as well as condensed linkages (i.e. 5-5, β-β, β-5 and β-1) [9-11]. **Figure 1.1** illustrates the main phenylpropanoid units of lignin and the variety of inter-unit linkages within lignin's structure. β-O-4 (β-aryl ether) linkages are the most abundant and can account for more than 50% of all linkages [9]. Lignin is primarily bound to hemicellulose through chemical linkages on the galactose and arabinose residues on the side chains of hemicellulose molecules [4], with limited hydrogen bond interactions with cellulose. Lignin is produced as plant growth ceases and provides reinforcement of cellulose fibrils playing a crucial role in the structural integrity of the plant while also taking part in plant defense by deterring plant pathogens through its aromatic structure [12]. When examining the structure of lignocellulose, specific feedstocks have major differences in the amount of each of the three major polymers and the subunits in both lignin and hemicellulose. Therefore, lignocellulosic biomass has been classified into three categories that have similar lignocellulosic structures: hardwood, softwood, grasses.
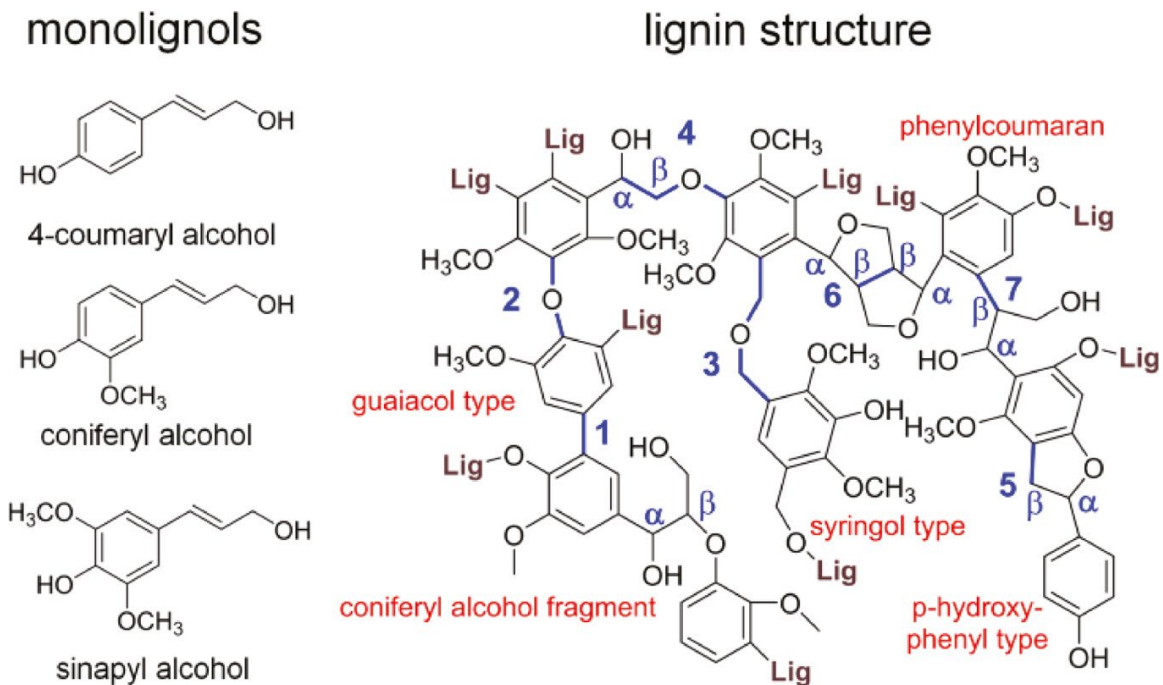
Figure 1.1: The three phenylpropanoid units of lignin and their main inter-unit linkages: biphenyl (1), diphenly ether (2), dibenzyl ether (3), β–O–4 (4), β–5 (5), β–β (6) and β–1 (7); taken from [13].

Table 1.2 illustrates the differences in cellulose, hemicellulose and lignin content between hardwoods, softwoods, and grasses. When comparing the two wood categories, we see hardwoods (i.e., oak and poplar) can have slightly higher cellulose and hemicellulose content compared to softwoods, but that softwoods (i.e., pine) have significantly greater lignin content. The hemicellulose fraction in hardwoods is primarily composed of xylose and xylan, where softwood contains mostly mannose [14]. Softwoods are considered more recalcitrant than hardwood to separate its lignocellulosic fractions because softwood lignin is present at higher percentages and is primarily composed of G lignin [15, 16]. Hardwood lignin contains high amounts of both G and S lignin subunits, which makes it less recalcitrant. Furthermore, grasses have the lowest lignin and cellulose content compared to softwood and hardwood, but it does have a very high amount of

hemicellulose. Grass hemicellulose is primarily composed of xylose subunits, and its lignin is derived of G, S and H subunits [15, 16]. By having such low lignin content and high hemicellulose content, grasses are widely used for biofuel production due to the milder conditions required to release its sugary fractions for fermentation purposes. The composition of the major feedstocks not only vary in their general classifications, but each plant species, and individual plant within a species has major differences. Nonetheless, understanding the general structure of cellulose, hemicellulose, and lignin and how they are bound together can help determine the processes by which lignocellulosic polymers are extracted and used to create more sustainable fuels, chemicals, and materials.

Table 1.2: Lignocellulose composition of hardwoods, softwoods, and grasses; adapted from [17].

	Lignocellulose Composition (wt%)		
	Cellulose	Hemicellulose	Lignin
Hardwoods	40-55	24-40	18-25
Softwoods	45-50	25-35	25-35
Grasses	25-40	35-50	10-30

Lignin Extraction Methods

The deposition of lignin on the surface of cellulose hinders the extraction of cellulose's usable sugars and fibers for use in fuels, chemicals, and paper production. Therefore, the main goal of lignocellulosic biorefineries and the paper/pulping industry is to utilize a variety of pretreatment processes that allow for the separation of lignin from the usable polysaccharides in lignocellulose. **Table 1.3** summarizes the more traditional extraction methods utilized by the paper/pulp and biorefinery industries.

Table 1.3: Traditional lignin extraction methods from paper/pulp and biorefinery industries [18].

Lignin sources	Lignin extraction methods	Conditions and procedures	Lignin features	References
Lignin from pulp and paper industry	Kraft Pulping	150~180 °C, H_2O, NaOH, Na_2S	Kraft lignin: oligomers with highly condensed structures and -HS group (1.5~3wt% S), low purity.	[19]
	Sulfite pulping	140~170 °C, H_2O, sulfites (e.g. Na_2SO_3, $NaHSO_3$, $(NH_4)_2SO3$, $MgSO_3$, $CaSO_3$).	Lignosulfonate: oligomers with highly condensed structures and $-SO_3$ group (4~8 wt% S), low purity.	[19]
	Soda pulping	160~170 °C, H2O, NaOH, (Anthraquinone).	Soda lignin: sulfur-free, oligomers, with low purity	[20]
	Organosolv pulping	100~110 °C, H_2O, Alcohol or alcohol/water mixtures. Formic acid, acetic acid, H_2SO_4	Organosolv lignin: sulfur-free, with relatively high purity.	[21]
Lignin from biorefineries	Dilute acid hydrolysis of biomass	120°C ~300°C, H_2O, 0~5 wt.% HCl, H_2SO_4, H_3PO_4, or HF	Lignin oligomers: with less condensed structures, partial preservation of β-O-4 linkages.	[22]
	Concentrated acid hydrolysis of biomass	Two-step: 1) 20°C ~30°C, H_2O, concentrated mineral acid, *e.g.* 72% H_2SO_4; 2) 121 °C, diluted H2SO4	Klason lignin: highly degraded oligomers with condensed structures.	[23]
	Alkaline hydrolysis of biomass	40°C ~160°C, H_2O, NaOH, $Ca(OH)_2$, ammonia.	Lignin monomers and oligomers: with low condensation structures; N incorporation when ammonia was used.	[24]
	Enzymatic hydrolysis of biomass	30°C ~60°C, H_2O, cellulase, hemicellulase, addition pretreatments (e.g. dilute acid,	Enzymatic lignin residue: with less condensation structures and low purity,	[25, 26]

	steam explosion, ammonia fiber explosion).	large variety.	
Steam explosion	0.69~4.83 MPa, 160~260°C	Condensed lignin structures	[27]
Ammonia fiber expansion (AFEX)	100~150°C	Acetylated lignin, highly degraded structures	[28]

The pulp and paper industry represents the most prevalent source of commercial technical lignins on the market. The processes developed by this industry rely on the removal of lignin to separate the cellulose fibers for use in the production of paper products. Kraft, sulfite, soda, and organosolv pulping work by fragmenting lignin polymers into smaller water or alkaline-soluble fragments that can be removed from the solid cellulose fibers [18]. These four techniques can produce either sulfur-containing or sulfur-free lignin. Both the kraft and sulfite pulping processes utilize HS^- or SO_3^{2-} nucleophiles to cleave the ether linkages in lignin. The resulting highly reactive lignin will repolymerize (condense) due the replacement of C-O bonds with stronger C-C bonds, which will lead to highly variable lignin intermediates and properties [20]. Additionally, the sulfides introduce a significant amount of sulfuric functional groups onto lignins surface [19]. Both of which significantly impede the subsequent valorization of lignin into downstream co-products. On the other hand, the soda and organosolv pulping processes produce sulfur-free lignins with less modified structures.

The soda process is similar to krafting, but only involves the use of NaOH, and due to its weaker nucleophile (OH^-) it generates a vinyl ether from lignin's ether bonds instead of cleaving them [19]. Since annual plants (i.e. bagasse, straw, and hemp) are impregnated more easily than wood and have lower lignin content, they require less chemicals making

them the preferred feedstock for the soda processes [29]. This results in significantly lower lignin yields, but with less modified structure compared to the sulfur-containing lignins. Organosolv removes lignin by cleaving lignin-carbohydrate ether linkages and some inter-unit ether linkages via a mixture of water, organic solvents (i.e. methanol, ethanol, and acetone), and either acid or base catalysts [20, 30]. The resulting lignin is of higher purity due to lower carbohydrates, ash, and no sulfur content compared to the other pulping methods. This makes organosolv a more preferred treatment for downstream lignin valorization technologies. While the paper/pulp industries want to utilize the cellulose from lignocellulose, biorefineries utilize all components of lignocellulose to produce renewable chemicals and fuels.

Typical biorefinery models aim to pretreat lignocellulose by altering or removing structural and compositional impediments (lignin) to expose lignocellulosic polysaccharides to hydrolysis, in hopes of increasing the yield of fermentable sugars for fuel (ethanol) production. Because of the variations found in the different pretreatment chemistries and conditions, there will be significant effects to the structural and compositional properties of the extracted lignin, even among the same feedstocks. More traditional pretreatment techniques involve physical methods like steam explosion or thermochemical methods like acid or alkaline hydrolysis **(Table 1.3).** Steam explosion involves high pressure and temperatures to expose the lignin to extraction with alkali or organic solvents, that has high levels of condensation [30]. Acid pretreatments aim to completely solubilize (>90%) hemicellulose, thereby exposing the cellulose to enzymatic hydrolysis [31]. Little lignin is removed from the cellulose solids during this process, so the purified lignin after enzymatic hydrolysis can contain large amounts of residual sugars.

In contrast, alkaline-based pretreatments primarily dissolve the lignin fractions with residual hemicellulose, exposing the solid cellulose to further enzymatic hydrolysis [24]. The resulting solubilized lignin can then be recovered in a more pure form (>95%) by performing acid-base precipitation, which removes most of the residual sugars [32]. While the above-mentioned pretreatment technologies are the traditional "standards" for biorefineries, more recent research has shown the usefulness of biological pretreatments and "green solvents" such as ionic liquids or deep eutectic solvents [33].

Biological pretreatment focuses on the use of microorganisms like white or brown-rot fungi to selectively degrade lignin from lignocellulosic biomass, thereby reducing the severity of other pretreatment processes or increasing cellulose surface area for direct hydrolysis [3]. However, these microorganisms also utilize the polysaccharides for their own metabolism, so even though the use of thermochemical pretreatment can be avoided or reduced, there are more drastic losses to available sugars. Ionic liquids (ILs) are a category of molten salts at room temperature that offer several desirable features such as low-toxicity, no vapor pressure, strong polarity, reusability, and high stability as compared to other organic solvents [34]. Ionic liquids are often referend to as "designer solvents", due the near infinite number of combinations of anions and cations that form ionic liquids [35]. In terms of biomass pretreatment, this means ionic liquids can be tuned to selectively dissolve and fractionate specific components of lignocellulose with reduced chemical modifications [36]. Despite ILs potential in preferentially extracting lignocellulosic components, there remain high solvent and processing costs associated with using ILs at industrial scales [37]. Therefore, this has resulted in research involving the use of deep eutectic solvents (DES) for lignin extraction. DES are a mixture of chemicals acting as

hydrogen bond acceptor (HBA) or hydrogen-bond donors (HBD). They offer a cheaper alternative to ILs with comparable lignin extraction yields and purity (>95%) at mild operating conditions [38]. While using DES provides low-molecular-weight lignin with retained native lignin activities [38], DES is not as tunable to dissolving all the different fractions of lignocellulose compared to ILs. The pros and cons of each of the above pretreatment methods can be more extensively reviewed in other works [26, 36, 39-43]. No matter the pretreatment, the cost-effectiveness of lignocellulosic fuels still cannot compete with its fossil fuel counterparts. Therefore, the biorefinery model needs to incorporate lignin as a value-added product and commodity to make lignocellulosic fuels more cost-effective, as lignin's natural aromatic structure has a plethora of industrial applications.

Lignin Depolymerization and Fractionation

The socio-economic impact of lignin valorization cannot be understated as creating value from lignin by utilizing it as a source of natural phenolics will generate extra profit for a biorefinery, making biofuels less expensive and increasing their marketability as an alternative to fossil fuels. Recently, a considerable amount of research has shown the variety of applications for using waste lignin's natural phenolic structure to produce polymers, cement additives, resins, battery components, fuels and chemicals [44]. To produce most of these biochemicals from lignin, it is necessary to fragment the larger oligomeric structure of technical lignin into smaller compounds, known as depolymerization. The production of lignin monomers for use as a source of aromatic building blocks can help offer suitable alternatives to petroleum-derived aromatics that saturate the market [45]. To this end, it is important to understand methodologies for

10

depolymerizing lignin's larger structure into usable monomers or oligomers that can help valorize lignin waste streams. Some of the most popularly reviewed depolymerization methods are pyrolysis, acid/base/metal catalyzed hydrolysis, hydrogenolysis and oxidation [9, 45, 46].

Pyrolysis is a thermal treatment in the absence of oxygen, with or without catalysts. The products of pyrolysis are solid char, liquid oil and gases, and their yields depend on temperature and heating rates [47]. During pyrolysis of lignin, there is cleavage of the OH functional group in the aliphatic side chain, breakage of the alkyl side chain, and cleavage of aryl ether linkages between aromatic rings as temperatures increase [46]. This typically produces vinyl phenols, guaiacol, syringol, and catechol, but as temperatures exceed 500°C aromatic ring cracking can occur, reducing monomer yields [48]. Since pyrolysis is a highly complex reaction with low selectivity in the bonds it breaks, increased lignin condensation and repolymerization due to reactive phenolic monomers and free-radical reactions can further reduce bio-oil and monomer yields [49]. This degree of low selectivity is also pronounced in acid/base catalyzed hydrolysis of lignin.

Base-catalyzed lignin hydrolysis is typically carried out utilizing NaOH at temperatures above 300°C and high pressure. The main bond cleavage occurs on the aryl-alkyl linkages which produces mostly catechol, syringol, and guaiacol monomers [50-52]. Acid catalyzed hydrolysis utilizes formic acid and ethanol solutions at temperatures above 300°C and high pressure, with methoxyphenols, catechol, and phenol produced as the major products due to β-O-4 cleavage [53, 54]. Alternatively, metallic catalyzed depolymerization has an advantage over both acid/base catalyzed depolymerization because it has a higher degree of selectivity to certain monomeric compounds and milder

reaction conditions that don't require as high of pressure, temperature or pH [46]. Metallic catalyzed hydrolysis reactions can take place in ethanol, formic acid and water solutions with a variety of metal catalysts (i.e. Pt, Pd, Ni, Ru etc.) at temperatures between 100-300°C [46]. The bonds targeted are C-O and C-C linkages, and while major products produced are dependent on catalyst and feedstocks, phenol, 4-proplyguaicol, guaiacol and pyrocatechol are primarily formed [55-57]. Even though metal catalyzed depolymerization has been shown to produce more selective lignin monomers under less severe conditions, other methods like hydrogenolysis have much higher total monomer yields.

Hydrogenolysis has received increased attention due to its reductive bond cleavage of lignin linkages that are hydrogenated and less reactive, which can produce significantly higher monomer yields compared to pyrolysis or hydrolysis [58, 59]. While more traditional hydrogenolysis methods utilize H_2 gas as a hydrogen donating source to cleave ether linkages [60], catalytic transfer hydrogenolysis (CTH) offers the use of inexpensive organic alcohols and catalysts to generate hydrogen molecules at lower temperatures and pressures [61]. A variety of hydrogen donating agents have been utilized (i.e. formic acid, methanol, ethanol, teralin, water, isopropyl alcohol, acetonitrile, and acetone) to depolymerize lignin substrates [62]. Hydrogenolysis produces a variety of phenolics with a preference in forming alkylated phenols and providing retention of C-C double bonds due to less radical development [57, 62]. Since the reaction conditions during hydrogenolysis are at high temperatures and pressures, like base/acid/metallic catalyzed hydrolysis, there remain issues with the energy intensiveness of the process. Thus, oxidative procedures that use even milder conditions may be more attractive.

Oxidative strategies for lignin depolymerization can offer a more economically feasible valorization scheme due to the already widely employed oxidative procedures in pulp bleaching [45]. This includes the use of oxygen, hydrogen peroxide, or peroxyacids at milder reaction conditions (below 200°C). However, oxidative procedures can also be less selective, prone to overoxidation, cause aromatic ring destruction, add complex functional groups, decrease product yield, and induce repolymerization of monomers due to free radical generation [45]. More recent literature has focused on peracetic acid as an oxidizer, due to its ability to cleave C-C and ether bonds, its higher monomer selectivity, rapid reactivity under mild conditions, and minimal side reaction and by-product formation [63-65]. Many of the compounds produced during oxidation of lignin are hydroxylated phenolics (i.e. dihydroxybenzenes) and phenolic acids (i.e. coumaric acid). In conclusion, each depolymerization method has its pros and cons, but to produce higher monomeric yields or have very mild reaction conditions, hydrogenolysis and oxidative procedures may be more useful in valorizing lignin into usable monomeric phenolics.

Nevertheless, many of the above depolymerization methods will produce a highly complex mixture that can contain more than 300 different compounds, most of which are unidentifiable oligomers. Therefore, physicochemical techniques are used to separate high molecular weight chains from lower molecular weight fractions (i.e. monomers), known as fractionation. Separation methodologies are important as many of the applications using lignin phenolics will require high purity levels and even specific phenolics (i.e. pharmaceutical applications). Some of the most popular techniques are selective precipitation at varying pH, distillation, chromatography, and liquid-liquid extraction by partial suspension in organic solvents [66-73].

Alkaline extraction methods utilize strong alkaline solutions that react with phenols to form phenolates that precipitate out of raw solutions, and then mineral acids can be used to solubilize and purify the extracted phenolics [74-76]. This method does not have high specificity and is more so used to separate phenolics from other compounds such as residual sugars and furans. Alkaline extraction of pyrolysis liquids has also been shown to form amorphous residues and caustic soda precipitates that reduce recovery and creates additional waste streams [77, 78]. Distillation processes can help separate phenolics based on their boiling points, increasing specificity, but can result in significant losses due to degradation (22 wt%) [79, 80]. Chromatography and membrane filtration technologies can obtain incredibly high yields (87-93 wt%) of specific phenolics [79, 80], but their operating costs are extremely high. Liquid-liquid extraction (LLE) is a method of separating compounds based on their solubilities in two immiscible liquids. Due to its relatively low material cost and easy operation, LLE has become an attractive option for separating aromatic/phenolic compounds from lignocellulosic derived bio-oils. Previous work has shown good performance in extracting phenolic compounds from bio-oil utilizing solvents like chloroform, hexane and ethyl acetate individually and sequentially [66, 67]. They found that by using chloroform and ethyl acetate sequentially to extract compounds from pyrolytic oils created improved phenolic extraction yields compared to utilizing the solvents individually or using only non-polar solvents [66]. Even though LLE can require large amounts of solvents, due to their volatility, they can easily be recovered and reused after drying. Therefore, even though LLE is the simpler of the fractionation methods, it offers great versatility and low operating costs that could be applied at large scales.

Lignin Antimicrobial Properties

Due to the overuse of antibiotics in our society, there has been a steady rise in highly antimicrobial-resistant bacteria in the last decade. This has created a renewed interest in natural compounds for antimicrobial discovery amongst the scientific community [81, 82]. Plant-based phenolics have a wide spectrum of antimicrobial activity and a variety of ring scaffolds that makes them an ideal source of potential antimicrobial replacements [82, 83]. To this end, lignin is one of the most abundant naturally occurring sources of phenolic polymers on earth and is currently considered a major waste product in the paper and pulp industries and lignocellulosic biorefineries [44]. Therefore, utilizing lignin phenolics as a natural source of antimicrobials can help create new antibiotic replacements and facilitate the valorization of lignin. The following discussion will review the antimicrobial properties of lignin when used to create hydrogels/composites/materials, in its more natural polymeric state (technical lignin), and the monomers that can be formed after depolymerization.

Lignin Polymers

Hydrogels, Composites and Materials

The randomly cross-linked network of reactive functional groups (i.e. methoxy and hydroxyl) that comprise lignin's three-dimensional structure has allowed for its preparation into highly versatile materials. These materials are used as tissue engineering scaffolds, wound dressings, drug delivery systems, bio-sensors, adhesives, supercapacitors, bio-plastics, slow-release fertilizes/herbicides, active food packaging, and absorbents [30, 84, 85]. However, the literature is very inconsistent with its definition of lignin-based component materials, where "blends" and "composites" are often used interchangeably. Also, the true role of lignin as a matrix material, additive, filler, reinforcement, or bioactive

molecule is often unclear. Therefore, we will focus on hydrogels, films, fibers, and nanoparticles where lignin is attributed to its known antimicrobial properties. **Table 1.4** summarizes literature studies with lignin materials that have tested antimicrobial properties. Amongst the variety of lignin-based materials, hydrogels are one of the most widely researched [84, 86].

Table 1.4: Antimicrobial activities of lignin composites [84]. MIC: minimum inhibitory concentration; MBC: minimum bactericidal concentration.

Material	Application	Test Method	Tested bacterial strains		Results	Reference
			Gram +	Gram -		
Lignin model dehydrogenate polymer (DHP), alginate (Alg)	Biomedical	MIC/MBC	*B. cereus* *L. Monocytog enes* *M. flavus* *S. aureus*	*E. cloacae* *E. coli* *P. aeruginosa* *S. enterica*	DHP: MICs of 0.002–0.90 mg/mL and MBCs of 0.004–1.25 mg/ mL	[87]
Lignin model dehydrogenate polymer (DHP), bacterial cellulose	Biomedical	MIC/MBC	*S. aureus* *L. Monocytog enes* *S. typhimuriu m*	*P. aeruginosa*	DHP: MICs of 0.22-0.88 mg/mL and MBCs of 0.22-0.88 mg/ mL	[88]
PVA/lignin/silv er nanofibers	Biomedical	Agar well diffusion	*B. circulans*	*E. coli*	Inhibition zone for *E. coli*: 1.1 ± 0.05 cm *B. circulans* 1.3 ± 0.08 cm	[89]
Chitosan and lignosulfonate nanoparticles	Biomedical	Turbidimet ric	*S. aureus* *B. subtilis*	*E. coli*	Bacterial growth decrease	[90]
Lignin, PVA nanocomposite fiber webs	Biomedical or Packaging	ASTM E 2149-10	*S. aureus*	*E. coli*	99.9% reduction rate against *S. aureus*	[91]
Lignin decorated thin multi-walled carbon	Biomedical	Dilution method	*S. aureus*	-	68.7% bacterial growth decrease after	[92]

16

nanotubes in poly(vinyl alcohol) nanocomposites					18 h	
Gelatin/lignin films	Biomedical	Dynamic shake flask method (ASTM E 2149–2010)	*B. subtilis*	-	Excellent bacterial growth inhibition	[93]
Polyvinyl Alcohol, chitosan, lignin hydrogels	Biomedical and packaging	Optical density	*S. aureus*	*E. coli*	*E. coli* > 95% *S. aureus* > 85% Cell reduction after 24 h	[94]
Artocarpus heterophyllus peel lignin, chitosan biocomposites	Biomedical	Disc diffusion	-	*E. coli Klebsiella*	Inhibition zone for *E. coli*: ~0.9 mm *Klebsiella* ~1.1 mm	[95]
Cellulose-lignin beads	Biomedical	Plate count	*S. aureus*		Greater than 90% inhibition	[96]
Lignin nanoparticles	Biomedical	Optical density	*S. typhimurium*	*E. coli*	30-50% inhibition in growth	[97]
PLA films with lignin nanoparticles	Packaging	Plate counts	-	*X. axonopodis X. arboricola*	Up to 2 log units of inhibition	[98]
Lignin nanoparticles	Agricultural	Broth dilution	-	*P. syringae X. axonopodis X. arboricola*	2-3 log units after 24hrs of growth	[99]

Hydrogels are a three-dimensional hydrophilic network of polymers that are crosslinked to form a matrix that can swell with the absorption of water. To form a hydrogel the polymer chains are crosslinked together, this prevents dissolution in an aqueous environment, and the polymer chains must contain hydrophilic functional groups that bind with water causing swelling [100]. There are two main classifications for the formation of hydrogels: physical and chemical crosslinking [101]. Physical crosslinking involves ionic,

H-bonding, hydrophobic, hydrophilic, or molecular entanglements between polymer chains that are reversible. Chemical crosslinking involves permanent chemical bonds between the polymer chains that form the hydrogel matrix, which is usually performed by the addition of a crosslinking agent that links the polymer chains [102]. Hydrogels can also be made from a variety of materials that are generally classified as synthetic (acrylic polymers) or bio-polymer (protein, collagen, and plant material) based. Due to hydrogels variety of substrates and forms, they can be applied to many fields such as hygiene [103], agricultural water retention [104], CO_2 capture [105], and biomedical materials (wound dressing, drug carriers, and tissue engineering) [106-108]. These fields utilize hydrogels because of their high-water absorbency and unique mechanical properties (elasticity and strength). However, the increased use of hydrogels has highlighted the need for reducing problems associated with solubility, high crystallinity, biodegradability, unfavorable thermal properties, and unreacted toxic monomers and crosslinking agents [109]. Therefore, there has been much attention in utilizing lignin as a hydrogel material source.

Lignin-based hydrogels have been shown to improve water absorbency, mechanical properties, thermal stability, antioxidant potential, and provide a functional mechanism for up-taking and releasing antibiotics, all while replacing synthetic materials [86, 110]. However, in terms of demonstrating lignin as the main antimicrobial component, very few studies have been performed, and often utilize lignin along with other components to create a synergistic antimicrobial mechanism. For example, lignin nanoparticle incorporated PVA/chitosan hydrogels showed significant reductions in the viability of both *Escherichia coli* and *Staphylococcus aureus*, which was attributed to a synergetic effect of lignin and chitosan [94]. Other works have utilized dehydrogenative polymer of coniferyl

alcohol(DHP), an enzymatically synthesized lignin model compound from coniferyl alcohol (CA), in hydrogels comprised of either alginate or bacterial cellulose [87, 88]. While these hydrogels do not use lignin as a structural feature, they do utilize a lignin derivate as a slow-release antimicrobial agent. These works represent a small fraction of the lignin-based hydrogel literature and signify a need for increased research/testing of lignin-based hydrogels for antimicrobial properties. Despite this lack in hydrogel research, there have been increased developments of lignin nanoparticles, fibers, and thin films with antimicrobial properties.

Recent works have produced lignin nanoparticles from technical lignins that have been found to be effective antimicrobial agents for plant/fruit pathogens in solution, when fixed in active packaging, and as a pesticidal treatment on plants [98, 99, 111]. Nanofibers composed of polyvinyl alcohol (PVA)-Acacia wood lignin-silver nanoparticles and just PVA/lignin were developed through electrospinning and showed antimicrobial properties [91] [89]. The silver nanoparticle incorporated nanofibers showed inhibition to both *Bacillus cirulans* and *E. coli*, which was attributed to a synergistic effect of the polyphenolic subunits of lignin and the presence of silver nanoparticles [89]. On the other hand, the PVA/lignin fibers only showed inhibition against *S. aureus* and not *E. coli* [91]. Furthermore, active packaging was developed from different lignin sources to create hydroxypropylmethylcellulose (HPMC)-lignin thin films with or without chitosan. The film's antimicrobial activity against Gram-positive and negative bacteria had the following trend with different lignins: organosolv of softwood > kraft of softwood > organosolv of grass. All lignin-containing films showed up to 8 log reductions in growth for *E. coli*, *S. aureus*, *B. thermosphacta,* and *P. fluorescens*, and the addition of chitosan enhanced these

properties. Based on these outcomes, the antimicrobial lignin materials discussed here can be useful candidates for biomaterials, healthcare, food, and agricultural products. Yet, a prevalent trend in many of these materials is the use of synthetic and bio-based materials in conjunction (i.e. PVA and lignin), or the presence of multiple antimicrobial sources (i.e. chitosan, lignin and silver). Thus, there remains gaps in the use of completely bio-based materials when using lignin for antimicrobial applications. Nonetheless, these materials have clearly showed the antimicrobial potential of lignin when incorporated in a functionalized material but purified technical lignins and their derivatives have shown greater antimicrobial properties on their own.

Technical Lignins

Due to lignin's high phenolic content, much research has observed antimicrobial activity of technical lignins. Technical lignins are the direct by-products of the industrial processing of wood, energy corps, or agricultural residues (i.e. biorefineries or paper and pulp industries). Most of the earlier work focused on the use of lignin derived from the paper and pulp industry. This work tested softwood and hardwood lignin as well as oxidized lignins (organosolv, sulfite, and Kraft lignins) on a series of yeast (*C. tropicalis, T. cutaneum*, and *C. albicans*), of which the oxidized lignins had lower antimicrobial properties than the unmodified lignin [112]. Kraft lignin and spruce hydrolysis lignin were found to be an effective antimicrobial agent against various phytopathogenic bacteria at concentrations ranging from 0.25%-2% on nutrition media [113]. The same study found that modification with quaternary ammonium increased antibacterial properties and was arranged in the following order of increasing potential: Kraft lignin<quaternized hydrolysis lignin<quaternized Kraft lignin. Due to the antimicrobial properties found in Kraft lignin,

20

agricultural researchers have even utilized Kraft lignin as a dietary supplement in cattle and broiler chickens, which showed a prebiotic effect by increasing beneficial bacteria that helped improve weight gain [114].

In terms of biorefinery lignins, the black liquors derived from bagasse and cotton stalks were found to be effective against Gram-positive bacteria (*B. subtilis* and *B. mycoides*), but ineffective against *E. coli* and *A. niger* [115]. Comparatively, cotton stalk lignin derived from alkaline methods were found to have antimicrobial properties against both Gram-positive and negative bacteria (*E. coli* and *B. pumilus)* [116]. Sugarcane bagasse and lignin from oil palm have shown biological activities against Gram-negative bacteria including *Klebsiella* sp. [117], and *E. coli/S. thyphirmurium* [118]. Additionally, corn stover lignin extracts from ethanol production were found to be antimicrobial against Gram-positive bacteria (*L. monocytogenes* and *S. aureus*) and yeast (*C. lipolytica*), but not Gram-negative bacteria (*E. coli* O157:H7 and *S. Enteritidis*) or the bacteriophage MS2 [119]. Food processing fungi (i.e. fermentation yeasts and *A. niger*) were also found to be inhibited by lignin produced from apple tree cutting residues [120]. Therefore, technical lignins' antimicrobial activity depends on the lignin origin, extraction method, chemical structure, concentration, and the tested organisms.

Moreover, studies have shown that by using successive ethanol-water fractionation of hydrolysis lignin, the antimicrobial activity can be concentrated by extracting the lower molecular weight fractions [121]. These lower molecular weight fractions were found to have lower MIC values for *S. aureus, B. subtilis, E coli, and S. enterica* compared to the larger molecular weight fractions and starting lignin. Collectively, this shows that the

highly heterogenous nature of technical lignin greatly impacts its antimicrobial potential, and that to increase its activity smaller units may need to be created.

Lignin Monomers and Depolymerization Products

While a variety of technical lignins have had notable antimicrobial properties, lignin monomers and depolymerization products have been shown to have greater antimicrobial properties compared to the larger and not well defined polyphenolic structures comprising technical lignins [122]. Early research involving *S. cerevisiae, B. licheniformis*, and *A. niger* found that wood lignin phenolic fragments containing a double bond in the α-, and β-positions of the side chain with a methyl group in γ-position have more antimicrobial properties than compounds containing oxygenated functional groups (i.e. -OH, -CO, -COOH) on the side chain [123]. An examination of lignin intermediates from the phenylpropanoid pathways comparing hydroxycinnamaldehydes, hydroxycinnamic acids and hydroxycinnamyl alcohols, had varying antibacterial (*B. subtilis, E. coli* and *P. syringe*) and antifungal (*S. cerevisiae, S. pombe* and *S. roseus*) properties. Hydroxycinnamaldehydes had the most notable antifungal and antibacterial activity, hydroxycinnamic acids displayed some antibacterial activity, and hydroxycinnamyl alcohols possessed little to no antimicrobial activity [124]. Greenberg, Dodds [125] found that naturally and synthetically produced phenolic monomers and dimers had antimicrobial properties against oral bacteria (*S. mutans* and *F. nucleatum*). Specifically, phenolics with C-C double bonds in the side chain (i.e. eugenol) had the greatest antimicrobial properties and that increasing alkyl chain length also corresponded to greater antimicrobial properties (i.e. 4-ethyphenol compared to 4-nonylphenol). Further research revolving around the antimicrobial properties of pyrolysis oils/acids, liquid

smoke, wood vinegars and plant extracts comprised of phenolic fragments (derived from lignin) have shown a variety of antimicrobial properties against bacteria, fungi, and molds as well [126-132].

Pyrolysis oils, liquid smoke and wood vinegars are derived from the liquid fraction obtained from the incomplete combustion of wood and other lignocellulosic materials. These products have been used extensively in human history to preserve food by smoking and creating a protective barrier on wood for building applications [133, 134]. More recently, pyroligneous acid from the slow pyrolysis of hardwood has shown significant antimicrobial activity against multi-antibiotic resistant strains of *E. coli*, *P. aeruginosa, S. aureus, Candida albicans* and *Cryptococcus neoformans,* based on agar diffusion tests [128]. The pyrolysis oil from pine trees has also been shown to have antimicrobial properties against the foodborne pathogens, *B. cereus* and *L. monocytogenese*, at concentrations ranging from 500-1000 ug/ml [135]. The main antimicrobial components of these products have been attributed to phenolics, furans, formaldehyde, and organic acids. Wood vinegars from sapwood were found to have significant antimicrobial activity against *Ralstonia solanacearum*, *Phytophthora capsici*, *Fusarium oxysporum*, and *Pythium splendens*. This study even evaluated all the individual compounds found within the wood vinegar for antimicrobial properties. Their results showed that while organic acids (i.e. acetic acid) and alcohols (i.e. methanol) comprised most of the wood vinegar, that they had little to no antimicrobial activity compared to phenols and guaiacols [127]. The antimicrobial activity of furans and cyclic hydrocarbons were also found to be weaker than the tested phenolics. This work highlights the antimicrobial potential of lignin phenolics and how they contribute significantly to pyrolytic product's antimicrobial properties.

However, other studies have shown that the primary active components of pyrolytic oils were not phenolics. For example, a study found that pyrolytic bio-oils showed significant toxicity to insects (*Leptinotarsa decemlineata*, *Trichoplusia ni,* and *Acyrthosiphon pisum*), fungi (*Pythium ultimum*, *Rhizoctonia solani,* and *Sclerotinia sclerotiorum*), and bacteria (*Clavibacter michiganensis* subsp. *michiganensis*, *Streptomyces scabies,* and *Xanthomonas campestris* pv. *vesicatoria*), but that polycyclic aromatic hydrocarbons (PAHs) were the primary active component [136]. Nonetheless, while these biomass-derived extractives represent a complex mixture of compounds, these works have shown that guaiacol, syringol, hydroxycinnamate, and vanillin derived lignin phenolics provide a higher contribution to their antimicrobial properties [123, 127, 137].

Even though the above products utilize pyrolysis, a popular depolymerization method, they focus on using whole biomass and not just lignin, so the applicability of these products to a lignin valorization scheme may not be entirely comparable. Therefore, there exists a gap in the literature for using other depolymerization strategies employed for lignin valorization to produce lignin-based antimicrobial products. Moreover, while the above discussions have clearly illustrated the antimicrobial potential of lignin products, their exact mode of action is not well understood or studied.

Mode of Action

In general, it is believed that technical lignin's mode of action centers around its hydroxyl function groups that interact with and damage cell membranes, causing leakage of intracellular components and cell death [60]. Hydroxyl groups on phenolics are known to promote electron movement in the membrane, acting as an electron exchanger that reduces the electron gradient across the membrane [138]. This causes a collapse of the

cell's proton-driving force, a decrease of adenosine triphosphate (ATP) and ultimately cell death. While a membrane disruption type of mode of action is commonly reported for larger technical lignins, there is a much wider variety of mechanisms reported for monomeric phenolics.

Phenolic monomers have been shown to have a variety of mode of actions involving the destabilization and permeability of cell membranes, enzyme inhibition from oxidized products through reaction with sulfhydryl groups, nonspecific interactions from the formation of reactive oxygen species that react with proteins, and inhibition of nucleic acid synthesis for both Gram-positive and negative bacteria [139-144]. Specifically, phenolics with increased hydrophobic characters, such as alkylated phenolics (i.e. carvacrol, thymol, and eugenol), alkyl gallates, and phenolic acids with alkyl esters (hydroxycinnamates), have been shown to interact directly with the outer membranes of bacteria [145-149]. This interaction disintegrates the lipopolysaccharides layer through the alteration of the dynamics of phospholipid chains and increases the permeability of cytoplasmic ATP and solutes, resulting in cell death [140]. Through this mode of action, these types of phenolics have been shown to have increased activity against Gram-negative bacteria compared to Gram-positive, due higher lipid content of Gram-negative cell walls [140, 150]. Their individual antimicrobial activities have also been shown to increase with the length of their alkyl chain [125, 148]. Furthermore, even though more polar hydroxybenzoic acids (i.e. gallic acid), have also shown similar modes of action to hydroxycinnamic acids by altering cell membrane structure and rigidity [140], the propenoid side chain of hydroxycinnamic acids increases their antimicrobial properties by facilitating transport through cell membranes [143, 151]. Conversely, increases in the hydroxylated function groups of

phenolics, which aids in quinone formation, provides increased reactivity with enzymes/amino acids and their subsequent inhibition [152]. Phenolics such as carvacrol have also been shown to have ionophoric activity by acting as a trans-membrane carrier of monovalent cations by exchanging its hydroxyl group for cationic salts (i.e. K^+) [153]. Therefore, while at the larger technical lignin level it seems hydroxyl groups are responsible for lignins mode of action, at the monomer level differences in function groups/structure can drastically change its mode of action and antimicrobial potential.

This causes issues when considering the antimicrobial applications of lignin depolymerization products that are complex cocktails of monomers (<50% w/w) and larger oligomers. Due to the wide variety of structures present, it would be hard to predict what the active compounds are and what organism they would be most effective against, without intensive and time-consuming experimentation. Similar concerns would be faced when comparing different depolymerization strategies, as reductive processes tend to create lignin derivatives with alkyl functional groups (i.e. syringyl propane, 4-ethylphenol, and 4-propylguaiacol), while oxidative procedures produce highly hydroxylated and acidic functional groups (i.e. hydroquinone, p-coumaric acid, and 2,6-dimethoxyhydroquinone). Thus, the use of predictive modeling could expedite the search for lignin derivatives with enhanced antimicrobial properties by simply examining their chemical structure and circumventing exhaustive experimental procedures.

Predicting Phenolic Antimicrobial Activity

Quantitative structure−activity relationship (QSAR) models are an indispensable tool in drug design and discovery. They work by finding relationships between the

variations in calculated molecular descriptors (properties) or fingerprints (functional groups) with the biological activity of a group of compounds so that the biological activity of new chemical entities can be assessed more quickly [154]. There has already been much work on utilizing QSAR models for predicting the antimicrobial properties of natural phenolics, but most of these studies are concerned with essential oils or flavonoids and not lignin specifically. Nonetheless, these studies have shown that the number and position of OH groups [155], the size and type of alkyl chains [125, 156], the presence of acetate or aldehyde groups [157], and the hydrophobic/amphiphilic character of the molecule contributes significantly to the antibacterial efficacy of natural phenolic compounds [158]. Even larger polyphenols (dimers, trimers, and tetramers) have been studied through QSARs, which found that number of hydroxyl groups, electronic/charge effects, and lipophilicity were the most common descriptors influencing their antimicrobial activity [159, 160]. While the general trends of the above studies may have some applicability to the compounds formed after lignin depolymerization, to date there has not been a complete study utilizing QSARs or other models to predict the antimicrobial activity of lignin. Therefore, this highlights a large gap in our knowledge and ability to truly understand the antimicrobial potential of lignin and its derivatives for future product development.

Conclusions and Research Motivations

The potential of lignocellulosic biorefineries are hindered by its high operating costs and competition with fossil fuels, but the valorization of lignin could resolve the marketability of lignocellulosic based fuels/chemicals. Much research in recent years has focused on developing a variety of lignin valorization strategies. Additionally, the overuse of antimicrobials and spikes in the evolution of resistant organisms has renewed the search

for novel antimicrobials using natural phenolic compounds. Since lignin is considered a waste product from different industrial sectors and has a polyphenolic structure, lignin has the capacity to become a future source of natural antimicrobial agents. Even though lignin has been shown to have antimicrobial properties, there remain gaps in how it can be effectively incorporated into different materials, what specific lignin derivatives retain antimicrobial properties in materials, as depolymerization products, and which have increased activity.

Therefore, in the present study, we aim to explore how lignin-containing polymers and depolymerized lignin bio-oils can be utilized as antimicrobial agents. Specifically, we will examine how different lignocellulosic components (lignin and hemicellulose) affect the formation and properties of physically cross-linked cellulose-based hydrogels. Can whole biomass-based hydrogels be formed based on these results and if lignin will retain its antimicrobial properties when incorporated into these hydrogels. Due to the lack of knowledge in the antimicrobial properties of lignin depolymerization products, we will explore the use of a reductive and oxidative depolymerization method to produce antimicrobial lignin-based bio-oils. The reduction process of catalytic transfer hydrogenolysis (CTH) will be used to depolymerize lignin biorefinery waste into a phenolic rich bio-oil. The antimicrobial properties of this bio-oil and liquid-liquid extracted fractions will be examined to better understand the antimicrobial potential of different lignin derivatives. Furthermore, an oxidative depolymerization strategy using peracetic acid will be used on the same biorefinery lignin to create a bio-oil with antimicrobial applications in the fuel ethanol industry. Finally, quantitative structure−activity relationship (QSAR) models will be developed to predict the antimicrobial properties of

lignin derivatives. Therefore, expediting the search for highly active lignin phenolics in future depolymerization strategies. The hope of this research is to provide critical knowledge and guidance on using lignin as an antimicrobial source in different industrial processes/products and to identify lignin derivatives with enhanced activity.

Chapter Organization

The first chapter serves as a literature review.

The second chapter explores the use of ionic liquids as solvents for creating physically crosslinked hydrogels from mixtures of cellulose, xylan and lignin to examine how each component affects hydrogel formation and its physical properties (i.e. mechanical strength and swelling kinetics). Whole biomass-based hydrogels were also formed from the same methods using sorghum bagasse and poplar wood. Additionally, the ability of lignin-containing hydrogels to retain antimicrobial properties were examined. These physically cross-linked hydrogels, which are completely bio-based, were also compared to a synthetically chemically crosslinked hydrogel using epichlorohydrin as a crosslinking agent. This study provides insights into using lignocellulosic biomass for hydrogel production and how these novel hydrogels have tunable mechanical and antimicrobial properties as compared to chemically crosslinked cellulose hydrogels.

The third chapter depolymerized biorefinery corn stover lignin by catalytic transfer hydrogenolysis (CTH) in supercritical ethanol with a Ru/C catalyst. The lignin-derived bio-oil was then sequentially extracted utilizing hexane, petroleum ether, chloroform, and ethyl acetate as solvents in order of less polar to polar. Antimicrobial properties of the bio-oils were screened against Gram-positive (*Bacillus subtilis*, *Lactobacillus*

amylovorus, and *Staphylococcus epidermidis)*, Gram-negative (*Escherichia coli*) bacteria and yeast (*Saccharomyces cerevisiae*) by examining microbial growth inhibition. This study provides insights into using sequential extraction to fractionate lignin-derived compounds and correlations between the properties of the extracted compounds and their antimicrobial activity.

The fourth chapter examined the unique properties of depolymerized corn stover lignin, from peracetic acid oxidation, to selectivity inhibit lactic acid bacteria (LAB) compared to fermentation yeast. We also examined the effects of the lignin bio-oil on enzyme function for both α-amylase and glucoamylase, determined the efficacy of using the lignin bio-oil as an antibiotic during the simultaneous saccharification and fermentation (SSF) of corn starch into fuel ethanol, and its ability to reduce contamination associated with LAB. This study shows the potential of using lignin depolymerization products as an antimicrobial replacement in industrial processes.

The fifth chapter aimed to develop QSAR models for predicting the antimicrobial properties of lignin monomers and dimers. The objectives of this chapter were to: 1) determine if open-source libraries of bioactive compounds (not lignin specific) could be used in conjunction with machine learning algorithms to develop predictive QSARs for lignin specific compounds, 2) develop more traditional QSARS using ordinary least square (OLS) regressions using antimicrobial activity measurements for lignin monomers from a metanalysis of available literature, and 3) an experimentally derived dataset using commercially available lignin monomers and dimers with screened antimicrobial properties. This study is the first attempt at predicting the antimicrobial properties of lignin compounds using QSAR models. Overall the results from this study will provide insights

into using different types of databases (open access, metanalysis, experimentally derived, and lignin specific/non-specific) to develop QSAR models with the potential to predict the antibacterial activity of future lignin derivatives.

The sixth chapter provides conclusionary statements regarding the full body of this work and insights into future work.

CHAPTER 2. HYDROGELS DERIVED FROM LIGNOCELLULOSIC COMPOUNDS:

EVALUATION OF THE COMPOSITIONAL, STRUCTURAL, MECHANICAL AND

ANTIMICROBIAL PROPERTIES

*This Chapter in whole has been published in *Industrial Crops and Products,* February
2019

Abstract

Hydrogels derived from lignocellulosic biomass and its constituent components have attracted growing interests due to the applications in agriculture, material, environment and biomedical fields. Some ionic liquids (i.e. 1-n-butyl-3-methylimidazolium chloride ($[C_4C_1im][Cl]$ and 1-ethyl-3-methylimidazolium acetate ($[C_2C_1im][OAc]$) can solubilize all factions of lignocellulosic biomass, while generating hydrogel materials without additional chemical crosslinkers such as epichlorohydrin. The present study explored the use of ionic liquids as solvents for creating physically crosslinked hydrogels from mixtures of cellulose, xylan and lignin to examine how each component affects hydrogel formation. The chemical, physical and mechanical properties of generated hydrogels were characterized using FT-IR, SEM, XRD, compositional analysis, swelling kinetics, and stress-strain analysis then compared against a chemically crosslinked cellulose hydrogel. We further tested hydrogels formed directly from poplar wood and biomass sorghum and examined the antimicrobial properties of the lignin containing hydrogels. The hydrogels with xylan had significantly higher elastic moduli at 0.1 MPa compared to other hydrogels, while poplar-based hydrogel had the highest strain of 65.3% and a stress of 0.12 MPa prior to rupture. The biomass-based hydrogels exhibited swelling ratio comparable to the chemically crosslinked cellulose hydrogel. All lignin containing hydrogels, besides the sorghum hydrogel, resulted in an 80% reduction in *E. coli* colony growth, indicating retained antimicrobial activities. This study provides insights into using lignocellulosic biomass for hydrogel production and how these novel hydrogels have tunable mechanical and antimicrobial properties as compared to chemically crosslinked cellulose hydrogels.

Introduction

Hydrogels are considered cross-linked polymeric materials that are resistant to dissolution and contain a large number of hydrophilic groups that allow for massive absorption of water molecules within its porous structure [102]. These cross-linked polymeric materials come in a variety of physical forms such as membranes, beads, and gels. Hydrogels can also be made from a variety of materials that are generally classified as synthetic (acrylic polymers) or biopolymer (protein, collagen, and plant material) based. Due to hydrogels variety of substrates and forms, they can be applied to many fields such as hygiene [103], agricultural water retention [104], CO_2 capture [105], and biomedical materials (wound dressing, drug carriers, and tissue engineering) [106-108]. These fields utilize hydrogels because of their high-water absorbency and unique mechanical properties (elasticity and strength). Due to the increasing environmental effects caused by the fossil fuel industry, it is important to examine biopolymer-based hydrogels. Specifically, biopolymer-based hydrogels have the advantageous properties of biocompatibility and biodegradability compared to synthetic based hydrogels. Biocompatibility and biodegradability aspects are favored by industries like agriculture that employ hydrogel's swelling capabilities for water irrigation and retention purposes [102, 161]. However, some bio-based hydrogels may have loss of mechanical properties due to formation type [102], and thus researchers are taking great steps towards obtaining novel bio-based hydrogels that have improved mechanical properties while maintaining their biocompatibility.

Lignocellulose is one of the most abundant renewable macromolecules on earth [162], and it is a great source of polymeric materials that can be used in hydrogel formation. All three of the major macromolecules found in lignocellulose (cellulose, hemicellulose,

and lignin) can be used to form hydrogels [163]. However, because cellulose is more abundant and has a large amount of hydroxyl groups that aid in the hydrogels structural properties [164], it can be seen as a more viable at industrial scales. Additionally, cellulose-based hydrogels can be formed through both chemical and physical cross-linking methods. However, due to celluloses high crystallinity, it can be hard to find the proper solvents to dissolve cellulose, especially for physical cross-linking methodologies that rely on the non-covalent interactions between polymeric backbones to form hydrogels [102]. Ionic liquids, which are a category of molten salts at room temperature, offer several desirable features, such as low-toxicity, no vapor pressure, strong polarity, high stability as compared to other organic solvents; they are reusable, and have a very high cellulose dissolution rate [34]. In addition, using ionic liquid as solvent can eliminate the needs of chemical cross linkers such as epichlorohydrin, glyoxal, silane, glutaraldehyde, sodium tetraborate etc., and thus simplify the hydrogel making process [102, 164]

While cellulose provides a more viable polymeric backbone for hydrogel formation, the addition of both lignin and/or hemicellulose have been shown to increase mechanical strength [165], porosity [166, 167], and aid in release of polyphenols and lipase immobilization when used in cellulosic hydrogels [168, 169]. Thus, their addition can help combat losses in mechanical strength when making bio-based hydrogels. Furthermore, lignin has been found to have a high degree of antimicrobial properties due to the large number of polyphenolic compounds that comprise lignin's structure [170]. These polyphenolic compounds can damage the cell membranes, causing lysis, of both gram positive and negative bacteria [122]. While the exact mechanism for cell lysis is not well understood, it is thought that the phenolic compounds in lignin can act as ionophores which

are known to increase the ion permeability of the cell membrane causing cell death [171, 172]. In the literature, there are a variety of factors that have been found to affect the antimicrobial properties of lignin that include the concentration of lignin, the chemical structure of the lignin monomers/polymers being used, the origin of the extracted lignin, and the type of microorganisms tested [12, 122]. However, the antimicrobial properties of lignin-based hydrogels were not well understood. Raw biomass based hydrogels have also been shown to create novel hydrogels by utilizing the whole dissolved biomass as a source of polymeric material for the hydrogels formation [173-175]. There are gaps in the literature on how lignin plays a role in biomass hydrogels and if it can confer its antimicrobial properties. These antimicrobial properties would be integral to biocompatible hydrogels in the biomedical fields by helping reduce infections and agricultural fields by preventing premature degradation or aiding in biocontrol [176, 177]. The use of raw biomass is more environmentally friendly, as no energy is needed to fractionate lignocellulose into its individual components for hydrogel formation and a variety of feedstocks can be utilized.

In this investigation, a simple and convenient physical crosslinking method for preparing cellulose-based hydrogels with the aid of ionic liquids was compared with known methods of cellulose based hydrogel formation using NaOH/Urea as a solvent with the aid of a chemical crosslinker. The effects of lignin and xylan were also examined in the formation of cellulose hydrogels to see the effects on mechanical strength and antimicrobial properties. Additionally, novel lignocellulosic biomass-based hydrogels were synthesized by directly solubilizing biomass (sorghum, and poplar) in an ionic liquid, and then they were compared to the pure cellulose-based hydrogels.

Materials and Methods

Materials

Microcrystalline cellulose (MCC, Avicel), Kraft lignin, xylan from beechwood, and 1-butyl-3-methylimidazolium chloride ([C_4C_1im][Cl], 99% pure) were purchased from Sigma-Aldrich. Hybrid poplar and biomass sorghum bagasse samples were provided by the Idaho National Laboratory, U.S. Department of Energy (Idaho Falls, ID). Both biomass types were ground by a Wiley mill to pass through 1 mm screen and kept at room temperature before use.

The IL, 1-ethyl-3-methylimidazolium acetate [C_2C_1im][OAc] was prepared by mixing 30% 1-ethyl-3-methylimidazolium methyl carbonate in methanol, purchased from Proionic, VTU Holding GmbH, in equal parts (on molar basis) with acetic acid drop-wise for 24h, then heating in a rotary evaporator at 70°C for 4h while rotating at 65 rpm. The resulting IL was further dried in a vacuum oven at 26°C for 3 days or until use.

Cholinium lysinate [Ch][Lys] was synthesized according to a method described elsewhere (Sun et al., 2014), where 1 mole equivalent of choline hydroxide was added to 1.2 equivalents of L-lysine at 4°C and stirred for 48hrs in the dark. Then the excess water was removed in a vacuum oven at 55°C, and the excess lysine removed with a solution of acetonitrile-methanol (9:1, v/v) through precipitation and centrifugation. The supernatant was then concentrated using a rotary evaporated and dried (<5% water content) in a vacuum oven for 48 hr at 70°C.

The deep eutectic solvents (DES) choline chloride-ethylene glycol and choline chloride-acetic acid were prepared by mixing choline chloride with either acetic acid or ethylene glycol in its solid state in a mass ratio of 1:2, followed by heating them at 65°C

for 2h in an oil bath as previously reported [178]. The solid mixture was stirred until no solid particles were left, leading to final transparent liquid; the mixture was cooled down in a desiccator for further use.

Solubility of Lignocellulose in Solvents

The ILs and DESs along with deionized (DI) water and 1:1 ethanol-water solution were examined for their ability to dissolve lignocellulosic compounds. Either MCC, xylan, or Kraft lignin was added to the solvent stepwise at a ratio of 1%(w/v) every 30 min with stirring every 10 min at 100°C until particles remained undissolved. A small sample of the mixture was plated on a microscope slide with a cover slip and the undissolved particles were examined visually either by naked eye or under a microscope if necessary. The solubility of each compound was recorded as the highest concentration until which undissolved particles were seen in the mixture.

Fabrication of Hydrogels

The fabrication of IL based hydrogels was accomplished via the following procedure. Firstly, a certain amount of MCC, xylan, or Kraft lignin was dissolved in the IL, $[C_2C_1im][OAc]$, at 80 °C for 30 min or until full dissolution with stirring every five minutes. For biomass based hydrogel formation, a certain amount of biomass was added to $[C_2C_1im][OAc]$ at 140°C for 4h with constant stirring. The following concentrations and combinations of cellulose, lignin, xylan, and biomass (in % w/v on dry basis) were added to the IL: 1) 4% cellulose; 2) 4% cellulose and 2% xylan; 3) 4% cellulose and 2% Kraft lignin; 4) 4% cellulose, 2% xylan and 2% Kraft lignin; 5) 4% ground poplar; and 6) 4% ground sorghum.

The resulting hydrogel solutions were then cast in a mold and immersed in a bath with 1:1 (v/v) ethanol-water solution for 24 h. Next, the cast hydrogels were immersed in a new batch of 1:1 (v/v) ethanol-water solution for 24 h, and then in pure DI water solution for two additional times with a duration of 24 h each to remove any residual ILs. The resulting hydrogels were then stored in DI water until further analysis.

The fabrication of hydrogels using NaOH/Urea as a solvent was performed following a pre-established method [179]. In brief, an aqueous solution consisted of 60 g/L NaOH and 40 g/L urea was filtered through a glass fiber filter (1.2 μm) to be used as solvent of cellulose. The cellulose solution was prepared by dispersing 4g of MCC into 100 g of solvent, stirred for 5min and then stored in a refrigerator (-18°C) for 12 h to allow full dissolution. The frozen solid was thawed and stirred extensively at room temperature for 4 h. Next, 5ml of epichlorohydrin was added dropwise to the thawed cellulose solution, and the mixture was stirred at 25 °C for 1 h. Then, the mixture was cast into a mold and heated at 50 °C for 20 h. The resulting hydrogels were then immersed in a DI water bath that was changed after 24 hrs for a total of 4 times, and then stored in DI water until analysis.

Characterization

Water swollen hydrogel samples were freeze dried using a FreeZone 6 liter console freeze dry system (Labconco, Kansas City, MO) at -50°C under 0.1-0.2 mBar vacuum for 24 hrs. Structural carbohydrates and lignin composition of raw sorghum, poplar, and freeze dried hydrogel samples were determined by a two-step acidolysis method according to NREL laboratory analytical procedure [180]. The sugar concentration was analyzed by

HPLC (Ultimate 3000, Dionex Corporation, Sunnyvale, CA) equipped with a refractive index detector and using a Bio-Rad Aminex HPX-87H column and guard assembly.

The freeze-dried hydrogels were characterized by Fourier Transform Infrared Spectroscopy (FTIR) using a Thermo Nicolet Nexus 870 ESP ATR–FTIR spectrometer. The samples were pressed to 12 psi under a spring-loaded jack onto the attenuated total reflection crystal (ATR), and analyzed using an average of 64 scans between 400 and 4000 cm^{-1} with a resolution of 1.928 cm^{-1}. The raw FTIR spectra were then normalized and baseline corrected using Omnic 6.1a software and compared in the range 800-2000 cm^{-1}.

The morphologies of the freeze-dried hydrogels were observed using scanning electron microscope (SEM) with a FEI Quanta 250 FEG instrument (Hillsboro, Oregon). The instrument operated at SE mode under low vacuum with beam accelerating voltages of 2kV after samples were sputter-coated in gold.

X-ray powder diffraction (XRD) patterns of raw MCC, freeze-dried hydrogels made of MCC (C), poplar (P), and sorghum (S) using IL, and freeze-dried hydrogels (N/U-C) using NaOH/Urea were attained using a Bruker D8 X-ray powder diffractometer (Billerica, MA). The freeze-dried hydrogel sample was ground using a blender and then 0.5 g of the sample was pressed under 17 MPa for 30s into a 40 mm wafer. Scans were collected at a speed of $1°$ min^{-1} in the 2θ range between $10°$ and $40°$. The crystallinity index (CI) was calculated using a method described previously [181]. Where CI was defined as the ratio between the intensity of the crystalline peak ($I_{002}-I_{AM}$) and the total intensity (I_{002}).

Swelling Kinetics

The swelling kinetics of hydrogels reflecting the water holding capacity was measured according to the following procedure [164]: the hydrogels were freeze-dried under -50°C for 24h and accurately weighed as m_d; the dry hydrogel samples were soaked in excess of DI water at room temperature and weighed at 20 min time intervals for 2 h to monitor the swelling. The surface of hydrogels was wiped off with filter papers to remove the excess water and the weight of the swollen samples was recorded as m_s. The swelling ratio (SR) of the hydrogels was determined using the following equation: $SR = (m_s - m_d)/m_d$. The swelling test was repeated three times for all samples and the averages values were recorded.

Antimicrobial Properties

The antimicrobial properties of prepared hydrogels were measured by examining their effects on *E. coli* (NRRL-409, obtained from the ARS Culture Collection) growth using colony counting. Colony counting of *E. coli* was performed by growing a 10^{-13} serial dilution of an *E. coli* stock inoculum on the hydrogel cast in a 35mm petri plate and a control tryptone yeast-extract glucose (TGY) agar plate. Each hydrogel was soaked in sterile TGY liquid growth medium for 12 h prior to inoculation. The extra medium solution was wiped off using sterile cheese cloth. Hydrogels and control plates were then inoculated with 50 μl of the of *E. coli* inoculum using a colony spreader, and incubated at 37°C for 24 hrs. After incubation, the number of *E. coli* colonies were counted. Subsequently, the *E. coli* colony number of hydrogel samples and the control plates (in triplicates) were compared using one-way ANOVA and Tukey's pairwise test in SigmaPlot (Systat Software Inc., San Jose, CA).

Mechanical Properties

The mechanical properties of the hydrogels were tested by the compressive stress-strain measurement using a TA.XT plus Texture Analyzer (Texture Technologies Corp. and Stable Micro Systems, Ltd, Hamilton, MA). Hydrogels were cast into a 24 flat bottom well plate that created cylindrical hydrogels with average diameter of 9 mm and height of 7 mm. The gels were set on the lower plate of the texture analyzer and then compressed using a 25 mm cylindrical plexi-glass load cell, with a loading strain of 80% at a compression rate of 1 mm/sec. The strain is defined as the change in the thickness divided by the thickness of the sample at the free-standing state. The stress is defined as the force divided by the area of the samples vertical to the direction of the force. The elastic modulus, E, was calculated from the average slope of the initial linear portion from the stress-strain curve, generally within the stress percent range of 0-10% [182].

Results and Discussion

Solubility of Lignocellulosic Compounds in Various Solvents

Since cellulosic based hydrogels are of interest due to their sustainable nature, we examined a variety of solvents to assess their ability to dissolve all factions of lignocellulose to determine the most viable solvent for our study. Specifically, we aimed to examine solvents that could give cellulose hydrogels with an amorphous structure, which is beneficial in hydrogel formation (Shen et al., 2015). We first examined three commonly used ILs, $[C_2C_1im][OAc]$, $[C_4C_1im][Cl]$, and $[Ch][Lys]$ because of their use as pretreatments agents to deconstruct lignocellulose for subsequent hydrolysis and fermentation of cellulose sugars (Brandt et al., 2013; Liu et al., 2017; Sun et al., 2014).

Results show that both imidazolium ILs dissolved high amounts of each lignocellulosic faction with [C₂C₁im][OAc] had the greatest ability to dissolve cellulose, xylan and Kraft lignin at values greater than 30% w/v (**Table 2.1**). [Ch][Lys] exhibited high solubility on Kraft lignin (~26%) while much lower solubility on cellulose and xylan, consistent with previous report that [Ch][Lys] selectively solubilize lignin component in cellulosic biomass [183, 184].

Table 2.1: Solubility measurements for various solvents dissolving the three factions of lignocellulose at 100°C *.

Solvent	Cellulose %(w/v)	Xylan %(w/v)	Kraft Lignin %(w/v)
Water	<1	~1	<1
1:1 Ethanol: water	<1	<1	<1
[C₂C₁im][OAc]	>30	>30	>30
[C₄C₁im][Cl]	~15	~17	~26
[Ch][Lys]	~6	~7	~20
[ChCl][ethylene glycol]	<1	~1	>30
[ChCl][acetic acid]	<1	~1	>30

*Solubility of the main components in lignocellulose in various solvents at 100°C. Cellulose is microcrystalline cellulose, xylan was derived from beech wood. [C₂C₁im][OAc] = 1-ethyl-3-methylimidazolium acetate, [C₄C₁im][Cl] = 1-n-butyl-3-methylimidazolium chloride, [Ch][Lys] = Cholinium Lysinate, and ChCl = choline chloride.

Deep eutectic solvents (DES) have received increasing attention recently as a new class of agents for biomass deconstruction and subsequent lignin extraction [178, 185-187]. As a eutectic mixture of two or more hydrogen-bond donors (HBD) and hydrogen-bond acceptors (HBA), DES shares some similar solvent characteristics of ionic liquids (ILs). However, DES can be prepared at high purity by simple mixing, thus avoiding tedious

purification and dehydration steps in synthesis of ILs [186, 188]. These characteristics make DES a promising candidate for multiple applications including biomass deconstruction. We examined two DESs, [ChCl][ethylene glycol] and [ChCl][acetic acid]) (both at 1:2 ratio) for their ability to dissolve lignocellulosic factions. These DESs have been tested in the fractionation of lignocellulose for biofuel production [189, 190]. From **Table 2.1**, the two tested DESs did not effectively dissolve MCC or xylan, but it did dissolve large amounts of Kraft lignin which is congruent with literature [190]. In this study, the IL [C_2C_1im][OAc] was chosen for further test because it can dissolve the largest amounts of each part of lignocellulose, which will benefit our goal of creating cellulosic based hydrogels that can retain lignin and its antimicrobial properties.

Compositional Analysis

Compositional analysis was performed on all hydrogel samples to determine final glucan, xylan, and lignin contents after gel formation, which can be found in **Table 2.2**. When examining the xylan and lignin content in the C+X, C+L, and C+X+L hydrogels, two times more xylan than Kraft lignin is present despite xylan and Kraft lignin being added at the same concentration during hydrogel formation. This suggests that the Kraft lignin did not bind as much as xylan to the cellulosic hydrogel during formation and the free portion was consequently washed away during the washing process, which was visually confirmed as initial baths were brown in color. For the biomass-based hydrogels, sorghum gel had a higher xylan composition than poplar gel; while poplar gel had a higher cellulose and lignin content than sorghum gel, which correlates to the lignin and xylan compositions in the raw biomass. However, in both the sorghum and poplar-based hydrogels the composition of xylan did not severely decrease compared to the raw biomass,

but the amount of lignin in both hydrogels was about half as much compared to the raw biomass. This coincides with the observations found in the MCC hydrogels as the lignin content was smaller than xylan content and the initial washing baths were brown in color, which further suggests that lignin does not bind as well to cellulosic hydrogels as xylan and the unbound portion is washed away.

Table 2.2: Chemical composition of dried raw biomass and hydrogel samples[#].

Sample	Glucan %	Xylan %	Lignin %	Others %
Raw sorghum*	34.8	20.4	16.0	28.7
Raw poplar*	43.8	13.3	25.7	17.2
C	83.6 ± 2.8^{a}	0.0	0.0	16.4
C+X	75.1 ± 0.7^{ab}	14.5 ± 0.8^{a}	0.0	10.4
C+L	83.7 ± 1.6^{a}	0.0	8.4 ± 1.5^{a}	7.9
C+X+L	70.8 ± 1.1^{bc}	12.3 ± 0.3^{a}	8.5 ± 2.3^{a}	8.4
N/U-C	44.8 ± 1.1^{d}	0.0	0.0	55.2
S	50.3 ± 2.6^{d}	20.3 ± 1.4^{b}	8.8 ± 0.3^{a}	20.6
P	73.2 ± 3.8^{ac}	13.4 ± 0.8^{a}	11.9 ± 0.8^{a}	1.5

*Chemical composition adapted from ("Bioenergy Feedstock Library," 2017).
[#]Letter difference indicate differences at 95% confidence where values are mean ± SE (n=3), using one-way ANOVAs and Tukey's pairwise comparisons

Structure and Morphology of Hydrogels

The structural and chemical changes of the hydrogels were investigated by FTIR using characteristic peaks associated with lignin and carbohydrates, which are shown in **Figure 2.1A**. The peak at 900 cm^{-1} represents the amorphous cellulose [191], and can be seen to have a high intensity in all hydrogels samples, which is greater in intensity than the pure MCC FTIR spectra found in **Figure 2.1B**. Conversely, the peak at 1098 cm^{-1} refers to C-O vibration of the crystalline region in cellulose [191], and can be seen to have a very

low intensity in all hydrogel samples when compared to the pure MCC sample in **Figure 2.1B**. This increase in amorphous cellulose and decrease in crystalline cellulose in the hydrogels, indicating a decrease in crystallinity index, is also supported by the XRD data found in **Figure 2.4**. The peaks at 1329 cm^{-1} and 1510 cm^{-1} represent the syringyl and guaiacyl condensed lignin and aromatic skeleton of lignin [191], respectively; only the hydrogels formed with the presence of lignin (C+L, C+X+L, S, and P) exhibited a peak at these wavenumbers, which indicates the retention of lignin after solidification of the hydrogels [168]. When examining the peak at 1056 cm^{-1}, which represents the C-O stretching in cellulose and hemicellulose [191], only the hydrogels containing hemicellulose/xylan (C+X, C+X+L, S, and P) had a relatively intense peak at that wavenumber. This increase in intensity indicated the presence of hemicellulose in those hydrogels during their formation. Even though not all hydrogels contained hemicellulose (xylan), they all exhibited an intense peak at 1375 cm^{-1}, which represents the C-H deformation of cellulose and hemicellulose [191]. Additionally, when examining the FTIR spectra for pure [C$_2$C$_1$im][OAc] the characteristic peaks at 1378 cm^{-1} and 1567 cm^{-1}, ascribed to the symmetric and asymmetric O-C-O stretches of the acetic anion [192], were not present in any of the hydrogel samples, thus supporting the complete removal of the IL from the hydrogels during washing. Taken together, these results clearly indicate that all hydrogels were primarily comprised of cellulose, and when lignin and hemicellulose were present in the hydrogels formation, they were retained upon solidification.
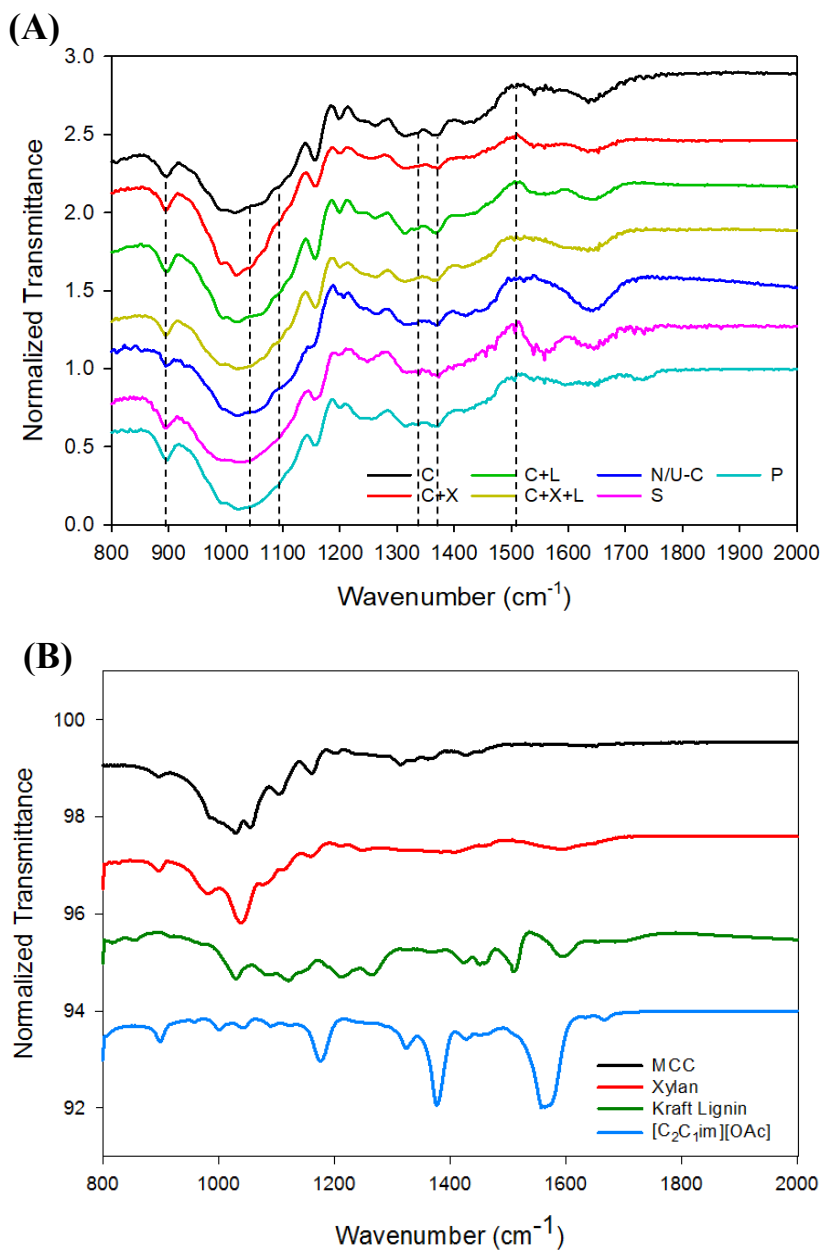
Figure 2.1: FTIR spectra for **A)** freeze-dried hydrogels made of MCC (C), MCC and xylan hydrogel (C+X), MCC and Kraft lignin hydrogel (C+L), MCC with xylan and Kraft lignin hydrogel (C+X+L), NaOH/Urea based MCC hydrogel (N/U-C), sorghum hydrogel (S), and poplar hydrogel (P); **B)** FTIR spectra for raw MCC, xylan, Kraft lignin, and $[C_2C_1im][OAc]$.

The morphological properties of all hydrogels were evaluated using SEM, as shown

in **Figure 2.2**. The figure shows that all hydrogels formed a porous structure with various

pore sizes, which are spherical in nature. Examining the N/U-C hydrogel, we saw larger

pore sizes compared to the other hydrogels, which could be attributed to its higher water retention (**Figure 2.4**) that caused the expansion of the pores [193]. Comparatively, the C+X, C+L, C+X+L, sorghum (S), and poplar (P) hydrogels presented much more compacted honeycomb like structures, compared to larger pores of MCC hydrogel (C), which may be contributed to the addition of other lignocellulosic compounds that filled the pores of the cellulosic hydrogel [164]. Furthermore, this homogeneous porous architecture of the hydrogels with added lignocellulosic compounds suggest good miscibility between the cellulose and xylan and/or lignin [193].
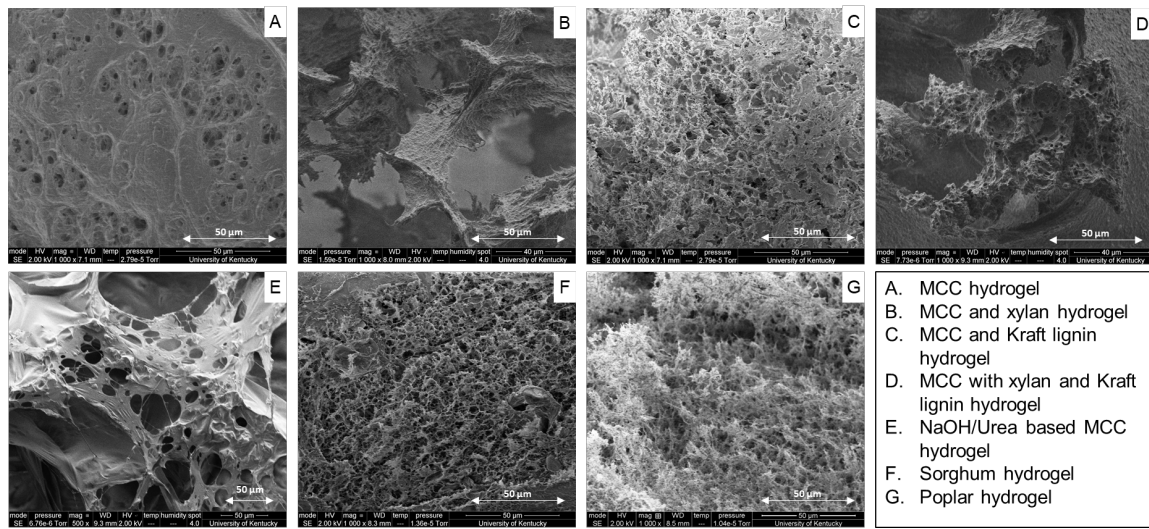
Figure 2.2: SEM images of the cross section of freeze-dried hydrogels: MCC (C), MCC and xylan hydrogel (C+X), MCC and Kraft lignin hydrogel (C+L), MCC with xylan and Kraft lignin hydrogel (C+X+L), NaOH/Urea based MCC hydrogel (N/U-C), sorghum hydrogel (S), and poplar hydrogel (P).

Crystallinity

XRD studies were performed on the MCC hydrogel (C), the sorghum hydrogel (S), poplar hydrogel (P), the N/U-C hydrogel, and MCC to reveal their crystallization behaviors. The XRD patterns shown in **Figure 2.3**. for MCC, reveals a large diffraction peak at $2\theta=22.6°$ relative to (020) crystal faces of cellulose I and diffraction peaks at

2θ=15.5, 20.5, and 34.5° were corresponded to (110), (101), and (004) crystal faces, respectively [194]. The MCC showed a CI of 78.2%. However, in both the MCC hydrogel and the N/U-C hydrogel we see that all three of the MCC peaks disappear and that a flat and wide diffraction peak appeared instead at 2θ=21° (**Figure 2.3**). This implies that when the hydrogels were formed the crystallinity of the cellulose changed from crystalline (cellulose I) to a more amorphous pattern because the new diffraction peak at 2θ=21° belongs to the (020) crystal faces of cellulose II [164, 194]. This coincided with a significant decrease in the CI for the MCC hydrogel, 24.1%, and for the N/U-C hydrogel at 30.1%. For the poplar hydrogel we saw two diffraction peaks for (020) and (110) crystal face that appeared flat and wide at 2θ=21 and 17°, respectively. These changes attributed to a more amorphous cellulose structure in the poplar hydrogel with a very low CI of only 16.41%. The sorghum hydrogel contained one flat and wide peak at 2θ=21° corresponding to the (020) crystal face (**Figure 2.3**). This suggests that the sorghum hydrogel also had an amorphous structure with a decrease in its CI 36.84% when compared to MCC. Additionally, it can be seen that the sorghum hydrogel had an intense unknown peak at 2θ=26.8°, which was similarly found in other studies that examined the XRD patters of sorghum [195, 196], but the cause of the peak was not explained.
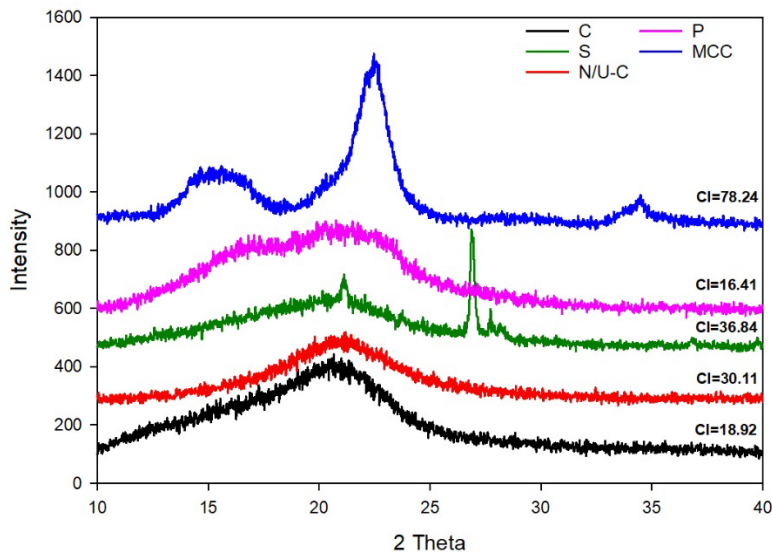
Figure 2.3: XRD patterns and the calculated crystallinity index (CI) of MCC, MCC hydrogel (C), NaOH/Urea based MCC hydrogel (N/U-C), sorghum hydrogel (S), and poplar hydrogel (P).

Swelling Properties and Mechanical Properties

The swelling kinetics of all hydrogels are shown in **Figure 2.4**, where all IL-MCC based hydrogels are seen to have similar swelling properties compared to the N/U-C and biomass based hydrogels. For example, the MCC, C+X, C+L, and C+X+L hydrogels using IL as a solvent showed approximately the same swelling ratio of ~120% at the 2hr time point. This supports that the addition of xylan and Kraft lignin to MCC hydrogels has no effect on their ability to absorb water. However, the poplar, sorghum, and NaOH/Urea based MCC hydrogels did show higher swelling ratios at 1065, 1155, and 1430%, respectively. These increases could be because the N/U-C hydrogel had the largest pore sizes that were less densely distributed (**Figure 2.2**), which has been shown to correlate to higher swelling ratios [197]. Additionally, since the biomass based hydrogels have less cellulose content (**Table 2.2**) compared to MCC based hydrogels, they will have less

diffusion resistance and a less compacted cellulose network that will cause the increase in swelling ratio [164, 198].
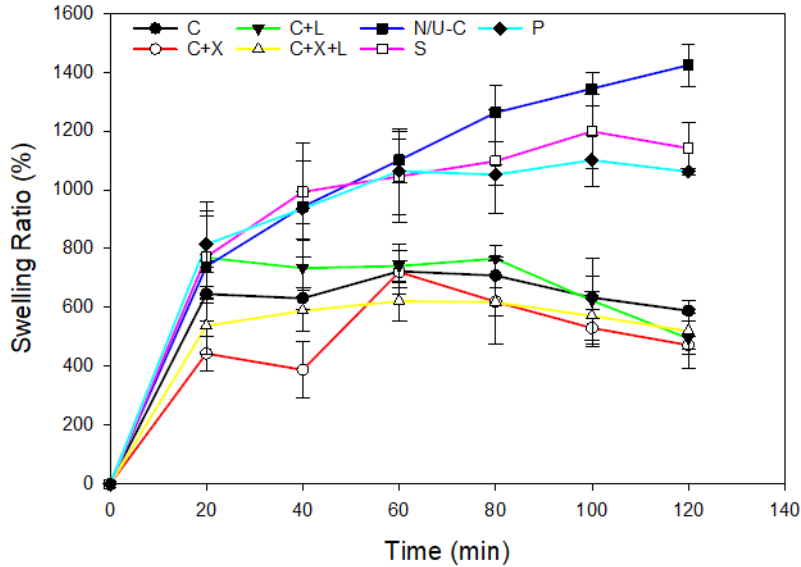
Figure 2.4: Swelling kinetics of freeze-dried MCC hydrogel (C), MCC and xylan hydrogel (C+X), MCC and Kraft lignin hydrogel (C+L), MCC with xylan and Kraft lignin hydrogel (C+X+L), NaOH/Urea based MCC hydrogel (N/U-C), sorghum hydrogel (S), and poplar hydrogel (P).

To assess the mechanical strength of the hydrogels, compressive force stress-strain curves for all hydrogels were measured (**Figure 2.5A**) and their elastic moduli calculated (**Figure 2.5B**). The MCC hydrogel (C) was seen to be relatively brittle and non-elastic as it had a low compressive strength with a minimal strain of 22% prior to rupture. When examining its elastic modulus (**Figure 2.5B**), we see that it is low but not significantly different from the other hydrogels besides the C+X hydrogel. For the C+X hydrogel, the compressive strength of rupture increased dramatically from 0.034 MPa in the MCC hydrogel to 0.075 MPa, but the elasticity did not change as they both broke at a strain of ~22%. This dramatic increase in strength was also seen in Lopez-Sanchez et al., (2015),

52

who suggested that the xylan crosslinks are less resistance to compression compared to cellulose, thus enabling the cellulose crosslinks to compact on each other increasing the density and compression strength of the network. The elastic modulus of the C+X hydrogel also supports its strength as it had a significantly higher elastic modulus at 0.1 MPa compared to all other hydrogels besides the C+L hydrogel. This indicates a highly stiff composite that requires high loads to elastically deform its structure. The N/U-C hydrogel was seen to have a higher compressive strength at 0.078 MPa and elasticity at a strain of 56% prior to rupture, compared to the MCC hydrogel. This increase in compressive strength and elasticity compared to the MCC hydrogel may be due to the chemical crosslinking method used to make the N/U-C, which has been found to increase mechanical properties of hydrogels [199]. While the N/U-C hydrogel's elastic modulus was not significantly lower than any of the hydrogels, it is considerably low at 0.006 MPa, which supports its highly elastic and rubber like features (**Figure 2.5B).**

(A)

(B)
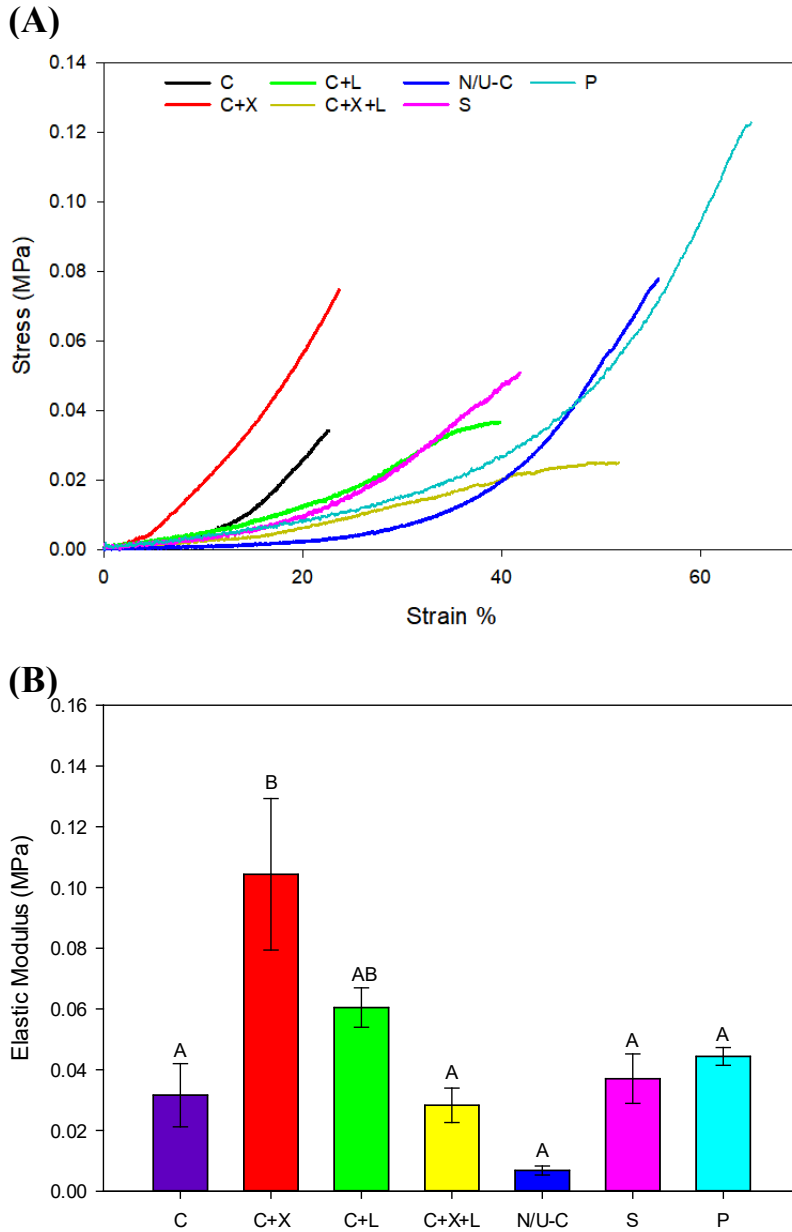
Figure 2.5: **A)** Compressive stress-strain curves and **B)** calculated elastic moduli of MCC hydrogel (C), MCC and xylan hydrogel (C+X), MCC and Kraft lignin hydrogel (C+L), MCC with xylan and Kraft lignin hydrogel (C+X+L), NaOH/Urea based MCC hydrogel (N/U-C), sorghum hydrogel (S), and poplar hydrogel (P). Letters in **B** indicate differences at 95% confidence where values are mean ± SE (n=3), using one-way ANOVAs and Tukey's pairwise comparisons.

When Kraft lignin was added to the MCC hydrogel (C+L), the strength of the

hydrogel did not change, but the elasticity increased to a strain of 40%. Additionally, the

elastic modulus of the C+L hydrogel was comparable to the C+X hydrogel, which indicates

a higher stiffness due to the presence of lignin. With the addition of both Kraft lignin and xylan (C+X+L), the elasticity is seen to increase to a strain of 52%, but strength decreased as seen by a compressive strength of 0.025 MPa at rupture and a low stiffness with an elastic modulus of only 0.026 MPa. The work by Nakasone and Kobayashi [200] shows that increasing lignin content subsequently increases the strength and elasticity of cellulosic hydrogels by reinforcing the cellulosic network, which supports our results with increase in elasticity in both the C+L and C+X+L hydrogels. For the sorghum-based hydrogel that had both xylan and lignin, we observed a higher compressive strength at a stress of 0.0493 MPa, but a lower elasticity at a strain of 41.2% compared to the C+X+L hydrogel. This increase in strength could be due to the higher composition of xylan in the sorghum hydrogel (Table 2.2). When examining the poplar-based hydrogel, the largest compressive strength at 0.123 MPa and strain of 65% at rupture were observed, representing a highly elastic and strong material despite having a similar elastic modulus to the other hydrogels. The improvement in strength and elasticity of the poplar hydrogel is likely due to the higher cellulose and slightly higher lignin content as compared to the other hydrogels (Table 2.2), which was shown to significantly reinforce and strengthen the cellulosic network [200].

Taking together, results from the present study suggest that adding lignin and/or xylan to cellulosic hydrogels led to significant changes in their mechanical properties. While the mechanical properties of these cellulosic hydrogels are weak compared to some cellulose hydrogels, those properties might fit applications in the biomedical field such as cartilage replacement [182]. It also noted that mechanical properties of the physically crosslinked hydrogels appear rather alike polyacrylic hydrogels than the chemically crosslinked N/U-C hydrogel [201]. Results also suggest that by altering the ratios of xylan

and lignin to cellulose one can potentially fine-tune the mechanical properties of cellulosic hydrogels [202]. It is also possible to use chemical crosslinkers when preparing the physically crosslinked hydrogels, especially in view of improving the mechanical properties of the hydrogels. However, the physical crosslinking conditions (temperature and pH, etc.) did not appear suitable for chemical crosslinkers [203]. As our focus is to create physically crosslinked hydrogels using ionic liquid without utilizing chemical crosslinkers, using chemical crosslinkers along with physical crosslinking methods may warrant future study.

Antimicrobial Properties

To assess antimicrobial properties of the prepared hydrogels, *E. coli* (NRRL-409, obtained from the ARS Culture Collection) was inoculated on the hydrogels (presoaked with TGY medium) and the colonies were counted after 48 hour cultivation. *E. coli* colony counts for the MCC hydrogel, C+X hydrogel, N/U-C hydrogel, sorghum hydrogel were not significantly different when compared to the TGY agar plate (control), suggesting no antimicrobial properties in those hydrogels (**Figure 2.6**). However, *E. coli* colony counts for the C+L hydrogel, C+X+L hydrogel, and poplar hydrogel were significantly lower than the control, N/U-C hydrogel, and MCC hydrogel (**Figure 2.6**). This suggests the addition of lignin to the MCC based hydrogels does offer some antimicrobial activity. When examining the lignin monomer ratios (S:G:H) in the sorghum (33:63:4) and poplar (54:46:0) ("Bioenergy Feedstock Library," 2017), we see that sorghum has a higher ratio of G-lignin (coniferyl alcohol), which according to previous work [123] has a higher amount of antimicrobial properties due to the presence of a double bond in the $C_\alpha = C_\beta$ position of the side chain and a methyl group in the γ position. However, these lignin

monomer ratios are for the raw biomass and Varanasi, Singh [204] has shown that ILs can preferentially degrade certain lignin monomers depending on the biomass during treatment, which will change the type of lignin moieties present that can confer antimicrobial properties. Thus, the differences in antimicrobial activity of sorghum and poplar hydrogels may be due to the differences in the lignin moieties present after gel formation. Additionally, Kraft lignin supplemented to the MCC hydrogels may have underwent a higher degree of depolymerization than the lignin derived from the sorghum, leading to increasing amount of phenolic fragments necessary for microbial inhibition [122], and Kraft lignin in free solution has already been shown to have antimicrobial properties towards gram-negative bacteria in previous research (Cazacu et al., 2013). Therefore, the results support that lignin can retain its antimicrobial properties when introduced into a cellulosic hydrogel, and that depending on the biomass source, native lignin in the biomass can also be a source of antimicrobial activity. As the antimicrobial mechanism of lignin derived compounds is still not well understood, further research should be performed to identify the lignin moieties present in the hydrogels based on factors like molecular weight, distribution of functional groups, and inter-unit linkages to determine the exact source of the antimicrobial properties from lignin present in the hydrogels.
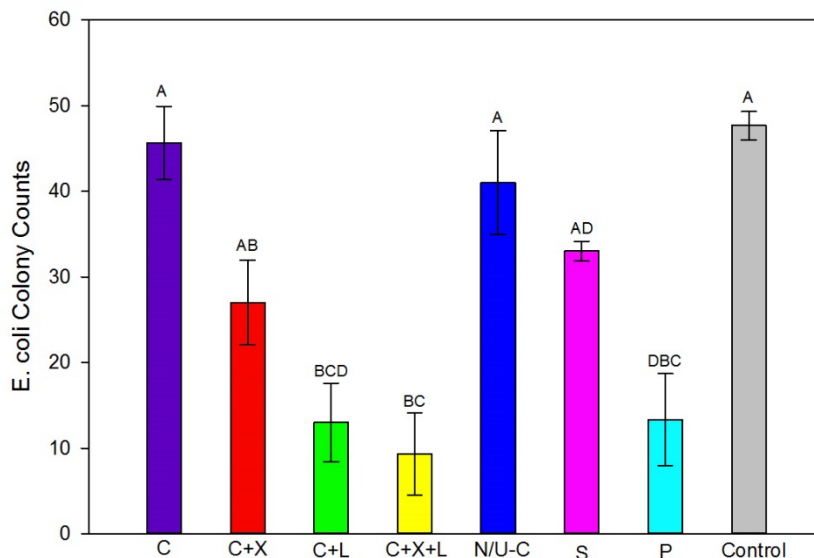
Figure 2.6: *E. coli* colonies counted after growing on MCC hydrogel (C), MCC and xylan hydrogel (C+X), MCC and Kraft lignin hydrogel (C+L), MCC with xylan and Kraft lignin hydrogel (C+X+L), NaOH/Urea based MCC hydrogel (N/U-C), sorghum hydrogel (S), poplar hydrogel (P), and bacterial growth agar (control). Letters indicate differences at 95% confidence where values are mean ± SE (n=3), using one-way ANOVAs and Tukey's pairwise comparisons.

Conclusions

The potential of using ILs and lignocellulose for hydrogel formation was evaluated. The addition of both lignin and xylan to the cellulose-based hydrogel improved its mechanical strength and stiffness, despite still having less elastic strength than the known cellulose chemical crosslinker method, and lignin did provide retained antimicrobial properties. Conversely, utilizing raw biomass provided increased mechanical strength (poplar), similar water retention abilities (poplar and sorghum), and retained antimicrobial properties (poplar) when compared to the cellulose chemical crosslinker method. Collectively, results from this study demonstrated the potential of using ionic liquids to make physically crosslinked hydrogels directly from lignocellulosic biomass with increased mechanical and antimicrobial properties. Developing biodegradable and

antimicrobial hydrogels from lignocellulosic biomass may lead to potential applications in biomedicine and agriculture. Future research will improve properties of the and hydrogels and include a wider variety of feedstocks used for biomass-based hydrogels formation.

CHAPTER 3. ANTIMICROBIAL PROPERTIES OF CORN STOVER LIGNIN

FRACTIONS DERIVED FROM CATALYTIC TRANSFER HYDROGENOLYSIS IN

SUPERCRITICAL ETHANOL WITH A RU/C CATALYST

Abstract

Converting lignin to value added products at high yields provides an avenue for making ethanol biorefineries more profitable while reducing the carbon footprint of products generally derived from petroleum. In this study, corn stover lignin was depolymerized by catalytic transfer hydrogenolysis (CTH) in supercritical ethanol with a Ru/C catalyst. The lignin derived bio-oil was then sequentially extracted utilizing hexane, petroleum ether, chloroform and ethyl acetate as solvents in order of less polar to polar and the subsequent bio-oils were characterized using GPC, GC/MS and HSQC NMR. Results show lignin derived compounds were sequentially extracted into groups depending on the solvent polarity. Antimicrobial properties of the bio-oils were screened against gram-positive (*Bacillus subtilis*, *Lactobacillus amylovorus,* and *Staphylococcus epidermidis)*, gram-negative (*Escherichia coli*) bacteria and yeast (*Saccharomyces cerevisiae*) by examining microbial growth inhibition. Results show that CTH derived bio-oils inhibited all tested organisms at concentrations less than 3 mg/mL. Total monomer concentration and the presence of specific monomers (i.e. syringyl propane) showed correlations to antimicrobial activity, likely due to cell death or membrane damage. This study provides insights into using sequential extraction to fractionate lignin-derived compounds and correlations between the properties of the extracted compounds and their antimicrobial activity.

Keywords: Antimicrobial, Catalytic Transfer Hydrogenolysis, Depolymerization, Lignin, Liquid-Liquid Extraction, Bio-oil

Introduction

Lignocellulosic biomass has become a promising feedstock to circumvent issues associated with the exploitation of fossil fuels for energy and chemical production. In a biorefinery, lignocellulose (which is mainly comprised of cellulose, hemicellulose, and lignin) is fractionated using thermochemical or biochemical methods. Cellulose and hemicelluloses can be used to create bio-fuels like ethanol or bio-products like plastics, while lignin is considered a waste product and is most often combusted for heat generation [205]. While this helps reduce the need for fossil fuels in heat generation at a biorefinery, it does not significantly increase the profitability of a biorefinery. Recently, a considerable amount of research has shown the variety of applications for using waste lignin's natural phenolic structure to produce polymers, cement additives, resins, battery components, fuels and chemicals [44]. The socio-economic impact of lignin valorization cannot be understated as creating value from lignin by utilizing it as a source of natural phenolics will generate extra profit for a biorefinery, making biofuels less expensive, increasing their marketability as an alternative to fossil fuels.

Currently, overuse of antibiotic agents has become a growing problem facing our society. Because of this, there has been a recent spike in the evolution of antibiotic resistant organisms and a need for researchers to develop new antimicrobials. There are increasing research efforts in examining lignin derived phenolic compounds for their antimicrobial properties [206]. The native lignin in plants has been considered to play a notable role in the plant defense by providing antimicrobial, antifungal, antiviral, antioxidant, insecticidal and antifeeding properties [12]. Lignin's source of antimicrobial properties are due to the phenolic subunits that comprise lignin's polyphenolic structure [170]. These polyphenolic

62

compounds are thought to act as ionophores that increase ion permeability in the cell causing cell death or damage the cell membranes of both gram positive and negative bacteria causing cell lysis [122, 171, 172]. Based on literature, lignin concentration, the structure of lignin phenolic subunits, and origin of the extracted lignin are drivers affecting its antimicrobial properties, that also depend on the microorganism being tested [12, 122, 207]. While a variety of technical lignins (e.g. from the Kraft and organosolv processes) exhibited notable antimicrobial properties, lignin model monomers have been shown to have a greater antimicrobial affect compared to the larger and not well defined polyphenolic structures comprising technical lignins [123]. Thus, to increase the effectiveness and selectivity of antimicrobial properties, it is necessary to depolymerize polyphenolic structure in the extracted lignin into smaller units.

Since lignin is a randomly linked polyphenolic polymer containing ether linkages such as β-O-4, α-O-4, and 4-O-5, as well as condensed linkages (i.e. 5-5, β-β, β-5 and β-1), lignin is highly recalcitrant toward selective depolymerization making it difficult for effective valorization into low molecular weight phenolics [9, 45]. A variety of thermochemical methods have been employed to depolymerize lignin into fractions containing high amounts of monomeric phenolics, including pyrolysis [49, 208], hydrolysis [209, 210], and hydrogenolysis [9, 211]. However, pyrolysis and hydrolysis methods lead to increased lignin condensation and repolymerization due to reactive phenolic monomers and free-radical reactions that reduces bio-oil and monomer yields [49, 210]. Hydrogenolysis, on the other hand, operates via reductive bond cleavage of lignin linkages, generating hydrogenated and less reactive compounds [58, 59]. While more traditional hydrogenolysis methods utilize H_2 gas as a hydrogen donor to cleave ether

linkages, catalytic transfer hydrogenolysis (CTH) uses inexpensive organic alcohols to generate hydrogen molecules on the surface of catalysts while also serving as a solvent for the depolymerization products [58, 61]. While a variety of hydrogen donating agents have been utilized (i.e. formic acid, methanol, ethanol, water, isopropyl alcohol, acetonitrile, and acetone) to depolymerize lignin substrates [62], ethanol at its supercritical state has been found to produce less solid residues, facilitate higher biomass conversion, and act as a capping agent that reduces phenolic monomer repolymerization [212, 213].

To find a better use of the key lignin derived compounds, it is necessary to investigate a separation method that can selectively recover an individual compound or a group of specific compounds at low cost. Liquid-liquid extraction (LLE) is a method of separating compounds based on their solubilities in two immiscible liquids. Due to its relatively low material cost and easy operation, LLE has become an attractive option for separating aromatic/phenolic compounds from lignocellulosic derived bio-oils; especially when compared to chromatography or membrane filtration [68-73]. Previous work has shown good performance in extracting phenolic compounds from bio-oil utilizing solvents like chloroform, hexane and ethyl acetate individually and sequentially [66, 67]. They found that by using chloroform and ethyl acetate sequentially to extract compounds from pyrolytic oils created improved phenolic extraction yields compared to utilizing the solvents individually or using only non-polar solvents [66]. However, because chloroform and ethyl acetate are both polar, the use of additional non-polar solvents in the sequential extraction process could further improve specificity.

In this study, we aim to use CTH to depolymerize alkali enzymatic corn stover (AEL, a representative lignin stream from the biorefinery) producing a bio-oil with high

monomeric phenol content and then test its antimicrobial activity against a variety of microorganisms. The objectives of this study are to: 1) characterize the raw bio-oil from CTH of AEL; 2) sequentially extract compounds from the raw bio-oil in aqueous phase by utilizing solvents from low to high polarity; 3) characterize the sequentially extracted fractions; and 4) quantify the antimicrobial activity of the raw bio-oil and sequential extraction fractions against gram-positive (*Bacillus subtilis*, *Lactobacillus amylovorus*, and *Staphylococcus epidermidis)*, gram-negative bacteria (*Escherichia coli*) and yeast (*Saccharomyces cerevisiae*). The results from this study provide insights into the types of lignin derived compounds that confer antimicrobial activity and compounds that can be preferentially extracted from lignin bio-oil using a simple LLE method.

Experimental Section

Alkali-enzymatic lignin purification and analysis

Corn stover was pretreated at the National Renewable Energy Laboratory (NREL) using 70 kg NaOH/ ton of corn stover with 1:12 solid: liquid ratio loading at 92°C for 2 h. The lignin residue was produced after disk refining (200 kwh/ODMT) using a 36 inch disk refiner (Sprout Waldon) at Andritz pilot plant (Springfield, OH) and enzymatic hydrolysis (48 mg CTec2 and 12 mg HTec2 per gram of cellulose for 36 hour) [214]. The enzymatic hydrolysis residue (namely alkali enzymatic lignin, AEL) was then centrifuged to reduce the water content to approximately 20% solids. The received residual lignin was stored at -40°C until use. Following a previously reported lignin precipitation method [32], we further purified the AEL to remove the residual carbohydrates. In short, the aqueous AEL slurry was brought to pH 12.5 (~5:1 AEL to 2M NaOH), then the solution was centrifuged

at 4000 rpm for 10 min to remove the solids containing undissolved carbohydrates. Then the lignin was precipitated from the filtrate by decreasing the pH to 3.0 with 2 M H_2SO_4, centrifuged at 4000 rpm for 10 min to remove filtrate, and washed three times with 70°C DI water. The resulting lignin was then freeze-dried using FreeZone 6-liter console freeze dry system (Labconco, Kansas City, MO) at -50°C under 0.1-0.2 mBar vacuum for 72 hr.

Structural carbohydrates and lignin content of the received AEL and purified lignin samples were determined by compositional analysis according to an NREL laboratory analytical procedure [180]. The sugar concentration was determined by HPLC (Ultimate 3000, Dionex Corporation, Sunnyvale, CA, US) equipped with a refractive index detector and using a Bio-Rad Aminex HPX-87H column and guard cartridge assembly.

Catalytic transfer hydrogenolysis (CTH)

CTH was performed using a Parr Reactor (Parr Instruments, Series 4560 Mini Reactor, Moline, IL) at a set temperature of 270±5 °C for 1 h under a N_2 atmosphere and stirring speed set at 600 rpm. Purified AEL was loaded at a lignin-to-catalyst-to-solvent mass ratio of 2:1:30 [59], utilizing ethanol as solvent and 5% Ru/C as the catalyst. In a typical reaction, 1 g of lignin was loaded with 0.5 g of Ru/C (dry weight) and 30 g of ethanol during each hydrogenolysis reaction. When the reaction completed, forced air was used to cool the reactor to about 100 °C and followed by an ice bath to further cool the reactor to room temperature. All contents in the reactor were transferred out by rinsing with ethanol, and the liquid and solids were separated by centrifuging at 4000 rpm for 10 min. Solid fraction and a subsample of liquid fraction were dried in a vacuum oven at 60°C for 24-48 h to remove the solvent for mass balance and recovery of lignin derived oil-like

compounds (named bio-oil hereafter). The percent bio-oil, solid, and gas yields by weight were calculated as percentage of the loaded lignin sample[58].

Sequential extraction

The liquid fraction collected from CTH was diluted with water to a water-ethanol ratio of 80:20 v/v, before sequential extraction procedures. Four solvents in order of least polar to most polar (hexane, petroleum ether, chloroform, and ethyl acetate) were used for sequential extraction based on previous studies [66, 215]. Each solvent was added to the bio-oil ethanol and water mixture (BOEW) at a 1:1 ratio and vigorously shaken for 15 min. The immiscible layers were separated via centrifugation at 4000 rpm for 5 min, then the solvent was removed and replaced by the next solvent in the order described above and in **Figure 3.1**. As a result, a total of five fractions were obtained from the sequential extraction process, one from each solvent and an additional fraction that consisted of the water-ethanol mixture with residuals bio-oil not extracted by the solvents. All five fractions were then evaporated under vacuum at 60°C for 48 h to obtain the dry bio-oil. The percent bio-oil recovery from each fraction was calculated by dividing the mass of dried bio-oil from each individual fraction by the total mass of all fractions combined, which would equate to the total raw bio-oil mass.
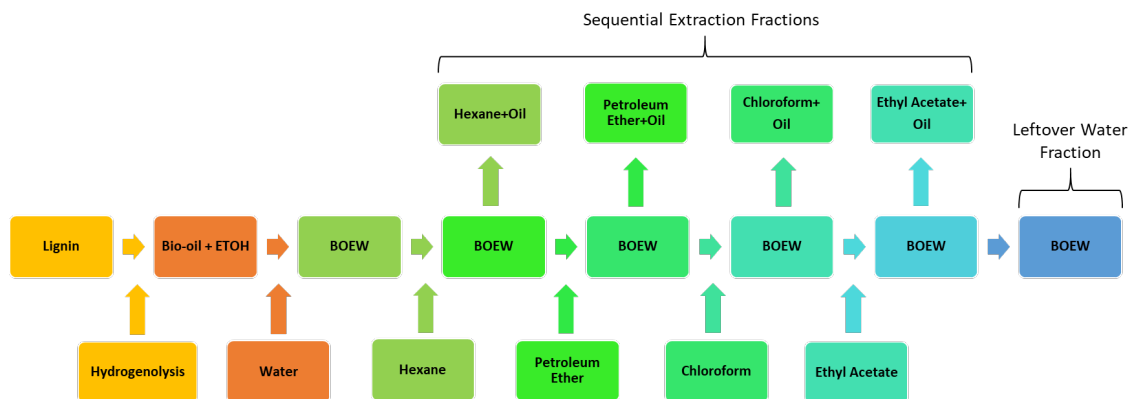
Figure 3.1: Sequential extraction flow chart (BOEW is bio-oil ethanol and water mixture).

Characterization of lignin derived bio-oil fractions

The weight-average molecular weight (M_w) and the number-average molecular weight (M_n) of the lignin sample, raw bio-oil, and sequential extraction fractions were determined using gel permeation chromatography (GPC) [216]. An Ultimate3000 HPLC system equipped with an Ultraviolet (UV) detector and Mixed-D PLgel column (5 μm particle size, 300 mm x 7.5 mm i.d., linear molecular weight range of 200 to 400,000 Da, Polymer Laboratories, Amherst, MA) was utilized. Separation was accomplished in a mobile phase of tetrahydrofuran (THF) at a flow rate of 0.5 mL min^{-1}, at 50°C. Elution profiles of materials were monitored at UV absorbance of 280 nm and calibrated using low molecular weight polystyrene standards (Product No. 48937, Sigma-Aldrich). Polydispersity Index (PDI) was calculated using the equation: PDI= Mw/Mn [216]. The molecular weight distributions of the larger oligomers in each bio-oil fractions were further elucidated utilizing matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) method based on a previously published protocol [217].

The raw and sequentially-extracted bio-oils were dissolved in ethanol and identified and quantified by GC/MS using an Agilent 7890B GC coupled 5977B MS with an Frontier

Lab Ultra Alloy-5, (60 m × 0.32 mm) capillary column. The GC was equipped with a two-way splitter which directed the gas stream separated from column into both MS and flame ionization detector (FID). The MS detector was used for compound identification and peak identification was performed via NIST MS spectra library matching, while The FID detector was used for compound quantification. Quantification of monomers was conducted based on FID peak area by using guaiacol (C_6), vanillin (C_6C_1), syringaldehyde (C_6C_2) and 4-propylphenol (C_6C_3) (Sigma Aldrich, St. Louis, MO, USA) as standards to convert peak area into concentration. These compounds were chosen as representative compounds where each response factor was used according to the origin and/or the number of carbons in the phenolic monomers identified [58, 218]. The temperature program started at 40 °C with a holding time of 6 min and increased to 240 °C at 4 °C min^{-1} with a holding time of 7 min, finally the temperature was raised to 280 °C at 20 °C min^{-1} with a holding time of 8 min. Helium was used as a carrier gas with a flow rate of 1.2 mL min^{-1}.

Approximately 100 mg lignin sample was dissolved in DMSO-d_6/pyridine-d_5 (4:1) or DMSO-d_6 under mild heat and sonication in an NMR tube until a homogeneous mixture was obtained. NMR spectra were acquired on a 500 MHz JEOL ECZR (Peabody, MA, USA) NMR spectrometer equipped with a 5-mm Royal Probe. The central DMSO solvent peak was used as an internal reference (δ_C 39.5, δ_H 2.5 ppm). The ^{1}H–^{13}C correlation experiment was an HSQC experiment (JEOL pulse sequence 'hsqc_edit_dec_en') with 25% non-uniform sampling (NUS). HSQC experiments were carried out using the following parameters: acquired from 11 to -1 ppm in f2 (^{1}H) with 1024 data points (acquisition time 136 ms), 220 to 0 ppm in f1 (^{13}C) with 64 increments (rebuilt to 256 with NUS) and 192 scans with a 1.5 second interscan delay. In all cases, processing used typical

sine bell (90º) in f2 and squared sine-bell (90º) in f1 (first point 0.5). Mestrelab MestreNova 14.0 (Mac version) software was used for volume integration of contours in HSQC plots, and quantification of lignin linkages using methods described by previous work [219]. Spectra are displayed in absolute value mode and color coded (in Adobe Illustrator CC 2019) using literature reference standards[220].

Microbial cultivation

USDA Agricultural Research Service Culture Collection (NRRL) provided the *Escherichia coli* (NRRL B-409), *Lactobacillus amylovorus* (B-4540), *Saccharomyces cerevisiae* (NRRL Y-567), *Staphylococcus epidermidis* (NRRL B-4268), and *Bacillus subtilis* (B-354) strains. Each microbe was grown on the recommended liquid media by NRRL with *E. coli* using TGY media (tryptone 5 g/L, yeast extract 5 g/L, glucose 1 g/L, dipotassium phosphate 1 g/L), *L. amylovorus* using M.R.S broth (Oxoid, CM0359), *S. cerevisiae* using YPD media (Fisher BioReagents™, BP2469), *S. epidermidis* using nutrient broth (BD Difco™, 234000), and *B. subtilis* using LB broth (Fisher BioReagents™, BP9723). Frozen cultures were prepared by first growing each microbe in liquid culture at 180 rpm shaking speed for 12 h at 37°C, besides *S. cerevisiae* which was grown at 32°C. These cultures were pelletized via centrifugation and washed with sterile media, then 500 μL of the washed cultures were added to 500 μL of sterilized 50% glycerol in a 2 mL cryovial and frozen at -80°C until use.

Antimicrobial assay

Frozen cultures of each microbe were first revived by adding cryovial contents to liquid media and allowing them to grow for 12 h at 180 rpm shaking speed and the respective incubation temperature above. Afterwards the cells were pelletized, washed, and

resuspended in fresh liquid media. To test for the bio-oil and sequential extraction fractions' antimicrobial properties, each microbe culture was cultivated in 48-well plates (flat-bottom polystyrene with clear bottom and sides, Corning Inc.) at preset bio-oil loadings and the optical density at 600 nm (OD_{600}) was monitored for 30 h with time points taken at 0, 6, 10, 18, and 30 h. These time points were found to represent key points of microbial growth curves based on our preliminary tests. All wells were brought to an OD_{600} of 0.2 prior to growth, and the lignin derived bio-oils were tested at 0.5, 1.0, 1.5, 2.0, 2.5, 3, and 4 mg/mL concentrations. The bio-oil fractions were first dissolved in ethanol as stock solutions and then added to the culture media such that all cultures had a final ethanol concentration of 5% (v/v). Two controls were used, one having the 5% ethanol concentration, and the other having just microbes and media. All samples and controls were performed in triplicate, so OD_{600} values for each time point represent the average of three replicates. To determine how the bio-oils affected microbial growth, the percent change in OD_{600} of the ethanol control during the exponential phase of growth was compared to the growth of the bio-oils at their different concentrations. This resulted in the average percent decrease in growth (degree of inhibition) for each bio-oil at each concentration, with the formula described in Eq. 1:

$$\text{Degree of Inhibition (\%)} = \left(1 - \frac{\text{Avg Max OD}_{600} - \text{Avg Min OD}_{600} \text{ of Growth with Biooil}}{\text{Avg Max OD}_{600} - \text{Avg Min OD}_{600} \text{ of Ethanol Control}}\right) * 100$$

(1)

Data was reported as the maximum concentration of each oil to have a degree of inhibition value of $\geq 90\%$, which represents little to no growth compared to the control.

Cell membrane integrity assay and microscopy

B. subtilis and *E. coli* were used as representative Gram-positive and negative bacteria to determine if cell membrane damage occurred in the presence of lignin derived bio-oils. The cells were incubated with 5% ethanol as control or with 4 mg/mL of raw bio-oil for 5 h prior to staining with LIVE/DEAD Bac Light Bacterial Viability Kit L7012 (Invitrogen, CA) in a 48 well plate (flat-bottom polystyrene with clear bottom and sides, Corning Inc.) at 37°C. The kit used a combination of green (SYTO9) and red (propidium iodide) fluorescent nuclear stains. SYTO9 is a green fluorescent dye that can penetrate cell membranes freely and bind to nucleic acids, while propidium iodide (PI) is a red fluorescence dye that can only penetrate damaged membranes. The penetration of PI causes displacement of SYTO9 due to its higher affinity for nucleic acids and the resulting damaged cell will fluoresce red instead of green [221]. Green fluorescence, indicating live cells, was measured at Excitation/Emission (Ex/Em) wavelengths of 485 nm/530 nm while red fluorescence, indicting dead membrane-damaged cells, was measured at 485 nm/630 nm using a SpectraMax M2 plate reader (Molecular Devices, Sunnyvale, CA). For fluorescence imaging, the stained cells were wet mounted and imaged immediately after staining using a Leica SD6000 spinning disk confocal microscope (Leica Microsystems, Wetzlar, Germany) equipped with 488 and 561 nm laser sources.

Results and Discussion

Mass balance

A mass balance was conducted to determine the percentage of lignin derived bio-oil, residual solids, and gas products produced during CTH of the AEL. **Table 3.1** shows

the mass percentages of each fraction after CTH, and while the oil and solid yields were directly measured, the gas fraction was estimated by percent difference from the total bio-oil and solid yields. The bio-oil yield after CTH of AEL was found to be 49.21 ± 1.70 wt% of starting lignin, the solid yield was $28.84 \pm 1.20\%$, and the gas yield was $21.95 \pm 2.90\%$. The raw bio-oil yield seen here is higher than the yield ($39.4 \pm 3.5\%$) reported by Zhou, Sharma [59]. Even though the same hydrogen donor solvent, catalyst, and temperature were used by Zhou and coworkers, the chemical composition, purity and structure of the lignin source (a commercial alkali lignin) were not fully characterized [59]. The purity of the AEL used in this study (after utilizing precipitation methods for purification) was estimated to be $95.11 \pm 0.18\%$ with $3.62 \pm 0.16\%$ glucan and $1.27 \pm 0.03\%$ xylan. The presence of glucan and xylan can suppress metal catalysts and inhibit lignin depolymerization during hydrodeoxygenation and CTH [215, 222]. Therefore, AEL's high purity and/or less condensed nature as confirmed by NMR likely support the high bio-oil yield obtained in this study.

The mass balance for bio-oils in the sequentially extracted fractions (SEF) are also shown in **Table 3.1**. Chloroform and hexane were found to extract the greatest amount of the bio-oils at 50.70 ± 6.01 wt% and $25.98 \pm 6.62\%$, respectively; while petroleum ether ($8.56 \pm 2.88\%$), ethyl acetate ($5.81 \pm 3.17\%$), and the leftover water fraction ($8.95 \pm 0.31\%$) extracted considerably less products, based on total extracted bio-oil weight. Even though hexane and petroleum ether, both non-polar, have similar polarity [223], since hexane was used first for extraction, less products with an affinity for non-polar solvents were available for extraction using petroleum ether. Similarly, since chloroform was the first polar solvent used there would be less products left to be extracted by ethyl acetate, even though ethyl

73

acetate has been found to be a superior solvent for lignin based bio-oils containing similar depolymerization products [63, 66]. Therefore, the first solvent used in each change in polarity during sequential extraction had the highest percentage of products extracted, which is similar to what Dodge, et al. [215] found during sequential extraction of bio-oil derived from CTH of unpurified alkali lignin.

Table 3.1: Mass balance of raw CTH products and bio-oil yields from each SEF

	Fraction	Average wt%
Raw CTH Products	Bio-oil	49.21 ± 1.70
	Solid	28.84 ± 1.20
	Gas	21.95 ± 2.90
Sequential Extraction Bio-oil Yields	Hexane	25.98 ± 6.62
	Petroleum Ether	8.56 ± 2.88
	Chloroform	50.70 ± 6.01
	Ethyl Acetate	5.81 ± 3.17
	Water	8.95 ± 0.31

Molecular weight distributions

By examining the weight-average (Mw) and number-average (Mn) molecular weight, as well as the polydispersity index (PDI) of the AEL, the raw bio-oil, and the SEFs, we can gain insight on changes in the molecular weight distributions (MWD) of lignin during CTH and sequential extraction. **Figure 3.2** shows the GPC chromatograms of the AEL and bio-oil fractions and **Table 3.2** summarizes the Mw, Mn, and PDI values. It is evident from the GPC chromatograms that MWD curves of all bio-oil fractions shifted to the right (i.e., lower Mw) compared to that of the unreacted AEL (**Figure 3.2**). The Mw of AEL was 3745±344 g/mol; whereas the Mw of the raw bio-oil after CTH was 755±51 g/mol, which indicates a high degree of depolymerization of the lignin into lower MW compounds. When examining the SEFs, the Mw of the hexane and petroleum ether fractions were lower than that of the raw bio-oil, while both the chloroform and ethyl

acetate fractions were not significantly different from the raw bio-oil. This indicates the more polar solvents extracted compounds with higher molecular weights, which correlates to the polar solvents extracting the greatest number of compounds based on weight combined (i.e., chloroform). Even though the Mw and Mn of the SEFs increase with polarity, the PDI's of the raw and SEF's remain low in the 1.8-1.9 range, suggesting compounds with similar MW's are extracted in each fraction. The leftover water fraction had a low Mw of 573±28 g/mol, which correlate to a higher percentage of low molecular weight compounds that may have a higher affinity and distribution coefficient for water than the other solvents (i.e., residual carbohydrates and/or phenolics with carboxylic acids) [66].

Table 3.2: The molecular weight distribution of AEL and lignin bio-oils derived from CTH and each SEF. Letters indicate differences at 95% confidence where values are mean±SE (n=3), using unpaired T-tests

Source	Mw (g/mol)	Mn (g/mol)	Polydispersity index (PDI)
Corn Stover Lignin (AEL)	3745±344 [A]	910±113 [A]	4.2±0.3 [A]
Raw bio-oil	755±51 [B]	426±56 [B]	1.8±0.12 [B]
Hexane	544±34 [C]	322±24 [C]	1.7±0.1 [B]
Petroleum Ether	629±43 [C]	382±23 [BC]	1.7±0.1 [B]
Chloroform	829±63 [B]	446±35 [BC]	1.9±0.1 [B]
Ethyl Acetate	924±70 [B]	480±38 [B]	1.9±0.1 [B]
Water	573±28 [C]	328±22 [C]	1.7±0.04 [B]
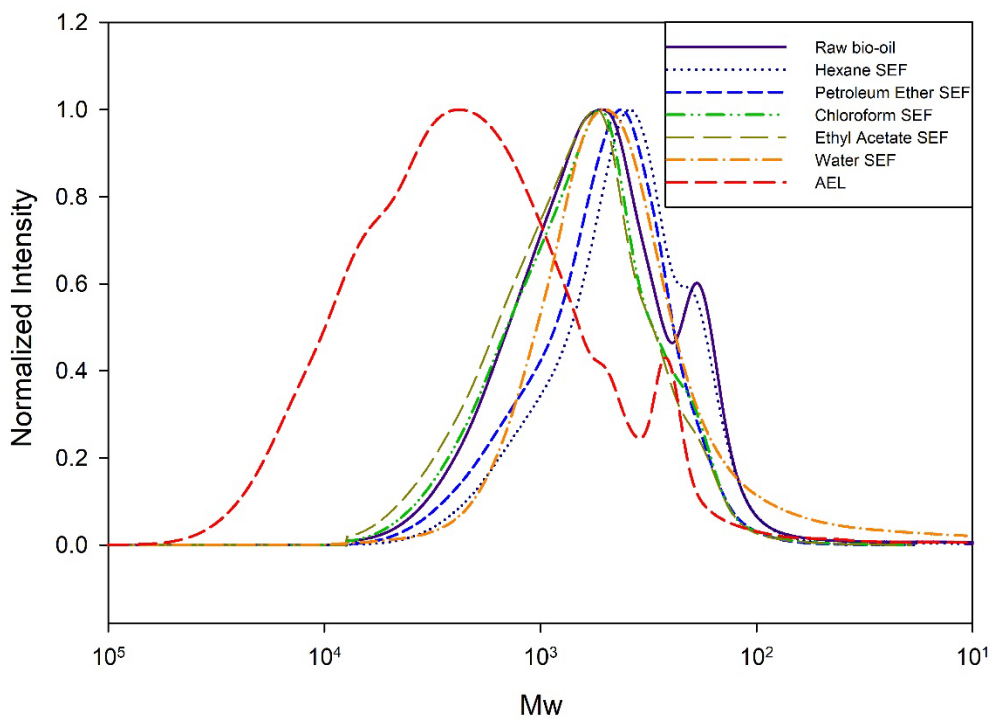
Figure 3.2: GPC spectra of purified AEL and bio-oils derived from CTH and each SEF.

MALDI-TOF has a higher sensitivity than GPC, especially for lower molecular weight compounds, since the mass to charge ratio of the ions correspond directly to the molar mass of the analyte [224]. Therefore, it is more typically used to identity the mass distribution of lignin depolymerization products, where GPC is used for larger technical lignins [225]. To this end, MALDI-TOF experiments were performed to better identify the distribution of lignin-based monomers, dimers, and trimers in the raw bio-oil and SEFs. MALDI spectra are provided in **Supplemental Figure 3.S1**, where monomers are identified below 300 m/z, dimers from 300-450 m/z, and trimers around 500 m/z [217]. The raw bio-oil, hexane, petroleum ether, and leftover water fractions have a larger amount of high intensity peaks within the monomer and dimer region compared to the other SEFs, which supports their lower Mw distributions in the GPC results (**Table 3.2**). Comparatively, the chloroform and ethyl acetate fractions have more peaks in the dimer to

trimer region and above, which supports their higher molecular weights. Furthermore, since the MALDI-TOF results corroborate the GPC results and our previous observation that polar solvents tend to extract higher molecular weight compounds. While the GPC and MALDI-TOF MS results provide a general trend on the size distribution of compounds in the raw bio-oil and SEFs, GC/MS and HSQC NMR were performed to identify the specific monomers and structures formed after CTH of AEL and better understand the specific reductive depolymerization that took place.

GC/MS and ^1H-^{13}C HSQC NMR characterization

The GC/MS analyses for the raw and SEF bio-oils are summarized in **Figure 3.3**. Since some SEFs had greater than 30 identifiable compounds, **Figure 3.3** only shows compounds that accounted for greater than 0.5 wt% of the bio-oil in question, while Supplemental **Table 3.S1** shows all identified compounds and their yields. The raw bio-oil and hexane fractions have the highest percentage of detectable monomers (32.44 and 36.57 wt%, respectively), in terms of the total bio-oil fraction's weight. This implies that bio-oil derived from CTH of AEL is mostly comprised of higher MW nonvolatile molecules (other than monomers), as evidenced by the GPC and MALDI results shown in **Table 3.2** and **Supplemental Figure 3.S1**.
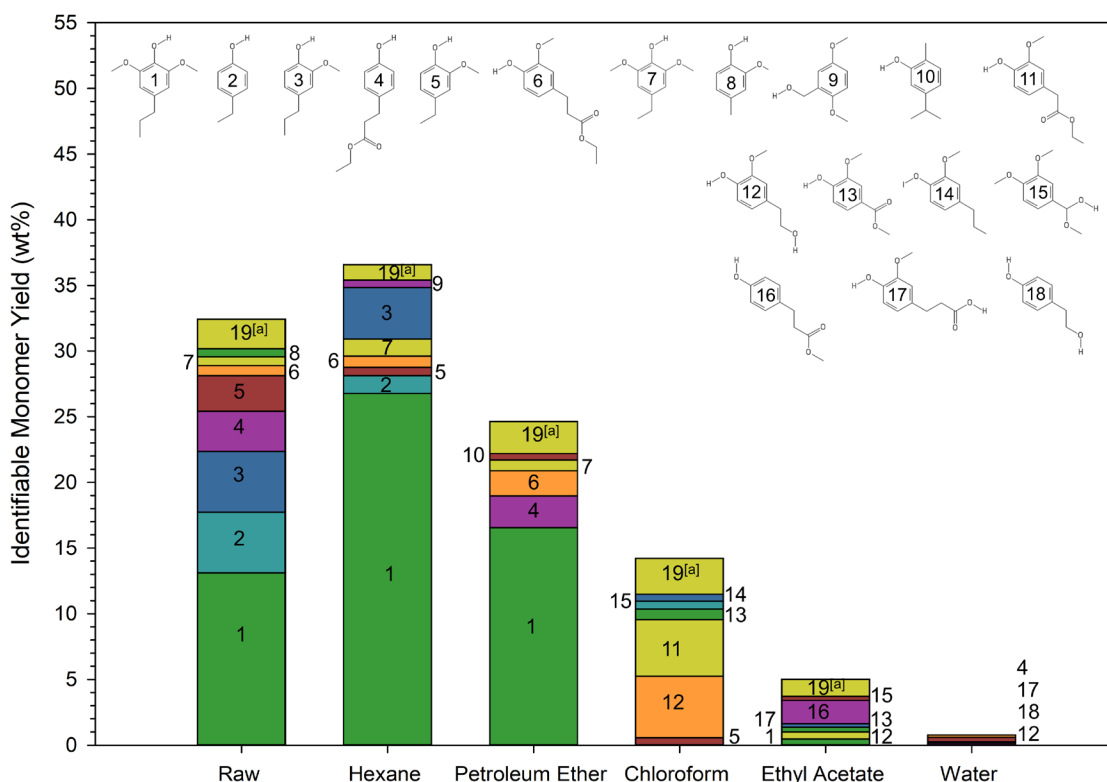
Figure 3.3: GC/MS identifiable monomers for raw bio-oil and SEFs, with total wt% of monomers in each bio-oil (only monomers representing > 0.5 wt% of the bio-oil fraction were included unless the total weight of other compounds was greater than 50% of total monomers).

[a] 19 represents the proportion of compounds not shown, see **Supplemental Table 3.S1**.

The GC/MS analysis revealed that the major monomeric phenolics in the raw bio-oil were phenolics with alkyl side chains (e.g. syringyl propane, 4-ethyl-phenol, 4-propylguaiacol, etc.) that comprised more than 25 wt% of the bio-oil. Those phenolic compounds are likely products from reductive cleavage of β-O-4 linkages and C_α/C_β or C_β/C_γ bonds during CTH of lignin [57, 62, 212]. Other major monomers were hydroxycinnamic acid derivatives such as methyl-4-hydroxyhydrocinnamate (3.04 wt%), and ethyl-β-(4-hydroxy-3-methoxy-phenyl)-propionate (0.78 wt%). Hydroxycinnamic

acids (i.e., ferulic and *p*-coumaric acid) are bound by ether, ester, C-C and β-*O*-4 linkages [226, 227], and their cleavage results in the liberation of hydroxycinnamic acid derivatives. After the hydroxycinnamic acids are liberated during CTH they can be hydrogenated to remove the double bond in the α, β positions of the side chain, or further decarboxylated to produce the alkylated phenolics and hydroxycinnamic acid derivates as identified in the raw bio-oil [226].

Additionally, since there was a presence of carboxylated phenolics like methyl-4-hydroxyhydrocinnamate and vanillin derivatives this indicates a lack of complete decarboxylation occurring, which is seen more often in reducing atmospheres (hydrogen) compared to the inert (nitrogen) atmosphere used in this study [212]. While other works have identified phenolics that maintained the C-C double bond in the α, β, or γ positions [226, 228], we only identified two compounds (2,6-dimethoxy-4-(2-propenyl)-phenol and eugenol) that combined only accounted for 0.3% of total bio-oil weight (**Supplemental Table 3.S1**). Because unsaturated C-C double bonds on side chain are highly reactive, promote lignin repolymerization, and are prone to hydrogenation [229], these compounds were only present at very low concentrations. To further elucidate the reductive chemistry performed by CTH of AEL, HSQC NMR was performed on the AEL and subsequent raw bio-oil after CTH.

HSQC NMR of the AEL displayed very few types of linkages present, however a small amount of β-*O*-4 can be seen (**Figure 3.4**). Despite its smaller presence, the integration ratio of β-O-4 to aromatics (S, G, H subunits) was found to be 27%, using the quantitative methods described by Wen et al [219]. This value represents a semiquantitative measure of the amount of monomeric phenolics that can be released from AEL after

depolymerization of β-O-4 linkages. Additionally, this assumes complete release of monomers via breaking β-O-4 bonds which, of course, may not be attainable given that the inter-unit linkages may not be an end group or part of two consecutive β-O-4 bonds. Comparatively, our results indicate that CTH of AEL produced monomer yields of 15.96 wt% based on the weight of lignin. Therefore, CTH released a little more than half of the available monomers from β-O-4 cleavage, since we do not expect to cleave 100% of linkages during CTH or produce all available monomers, these results are as anticipated. More prevalent in the AEL were cross-peaks that correspond to p-coumarate (pCA) and ferulic acid (FA). These cross-peaks, along with those corresponding to the β-O-4 linkage, were noticeably absent after CTH. This is unsurprising given that CTH should cleave the $C_β$-O bond in the lignin β-O-4 linkage [57, 62, 212]. Moreover, alkyl phenols are known to be produced from CTH of lignin with Ru/C [218]. This is reflected in the HSQC by the significantly increased number of CH_2 (red) and CH_3 (gold) cross-peaks in the CTH lignin and the absence of double bonds at the $C_α/C_β$ or $C_β/C_γ$ positions. Even more so, there were direct cross peaks associated with syringyl propane, which was found to account for over 13 wt% of the raw bio-oil. The aromatic region of the NMR for the raw bio-oil compared to AEL, also displays cross peaks consistent with reduced structures of pCA (H-pCA$_{2/6}$ and H-pCA$_{3/5}$), however, the exact functionality of the alkyl chain is not easily revealed with HSQC. Overall, HSQC NMR data supports the monomers found in the GC/MS data and the reductive chemistry that occurs during CTH, as described above.
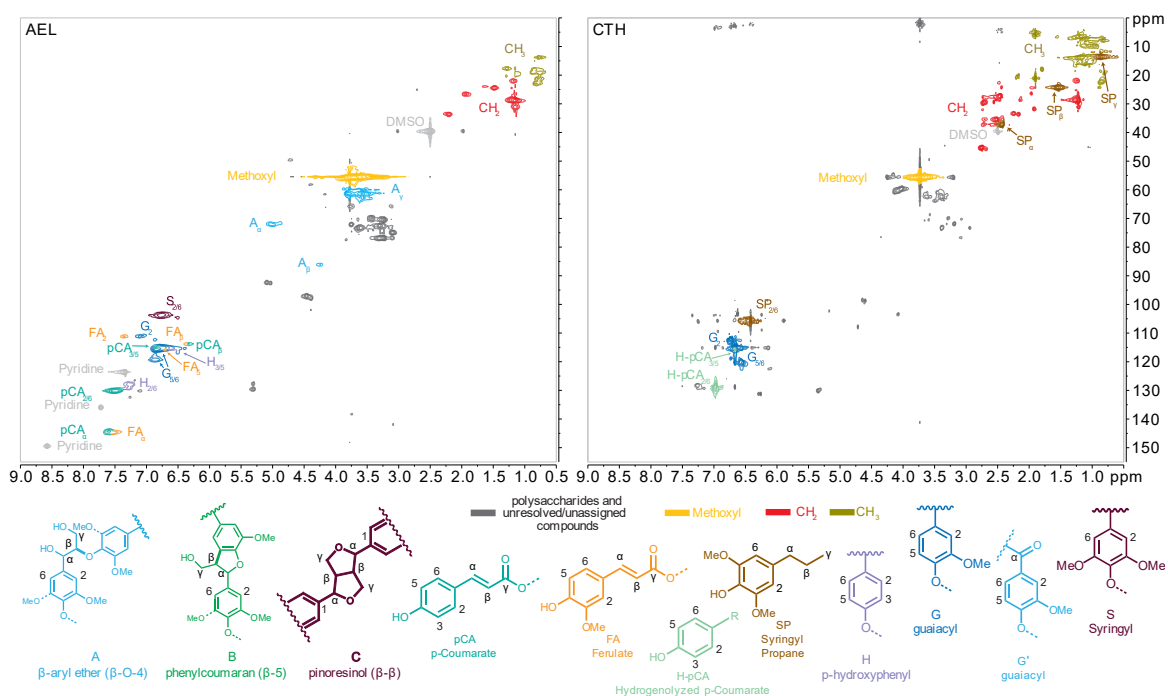
Figure 3.4: ^1H-^{13}C HSQC NMR of AEL and raw bio-oil derived from CTH. The structures of lignin compositional units and side-chain linkages were coded with colors corresponding to the cross peaks in the spectra.

Many of the same compounds found in the raw bio-oil were found in the SEFs. However, because the SEFs concentrated specific fractions of the raw bio-oil, there were a few new compounds found that were at too low of a concentration to be detected in the raw bio-oil. The SEFs had decreasing extraction efficiency of identifiable monomers with increasing polarity and order of extraction with the following values: hexane 36.57 wt%, petroleum ether 24.65 wt%, chloroform 14.24 wt%, ethyl acetate 5.04 wt%, and the leftover water fraction retaining 0.81 wt% (**Figure 3.3**). This also coincides with the GPC data, which indicated a gradual increase in the average MW in the order of extraction. It is surprising that chloroform extracted only 14.24 wt% of identifiable monomers, despite extracting over 50 wt% of the total raw bio-oil and having 30 identifiable monomers (**Supplemental Table 3.S1**). It is possible that there were no residual monomers present at

high enough concentrations to extract since chloroform was the third solvent used during sequential extraction. This is further supported by the fact that most of the chloroform's extracted monomers were vanillin derivatives (i.e., ethyl homovanillate and homovanillyl alcohol), which were found to account for less than 0.5% of total raw bio-oil's weight (**Supplemental Table 3.S1**). Additionally, the MALDI-TOF spectra for the chloroform fraction showed a much larger presence of peaks associated with dimers, trimers, and even larger oligomers compared to monomers, suggesting that the oligomers from the raw bio-oil were more easily extracted by polar solvents. Even though other works have shown chloroform and ethyl acetate have the greatest total extraction efficiency of phenolics in lignin based bio-oils when used individually [66, 67, 230], the use of sequential extraction here limits the concentration of monomeric compounds after each step.

The hexane and petroleum ether SEF's were primarily composed of alkylated phenolics (i.e., syringyl propane, 4-ethylphenol, and 4-propylguiacol), which can be attributed to the alkylated side chains with increasing affinity to the nonpolar solvents. Furthermore, results show that syringyl propane accounts for 26.77% and 16.57% of hexane and petroleum ethers total bio-oil weight, respectively, suggesting that compounds with alkylated or non-polar side chains can be preferentially extracted with the non-polar solvents used in this study. Similarly, chloroform and ethyl acetate SEF's contained large amounts of phenolics with increased oxygenated functional groups (i.e., ethyl homovanillate and homovanillyl alcohol) that would have a higher affinity for polar solvents. The leftover water fraction had four identifiable phenolic monomers that accounted for only 0.81 wt% of its weight (**Figure 3.3**), indicating that sequential liquid-liquid extraction can effectively remove almost all the bio-oil from the aqueous phase.

While the data is not reported here, low concentrations (small peaks) of sugar derivatives (e.g. glucose, xylose, mannose etc.) were identified in GC/MS spectra of the water fraction. Those sugar derivatives likely come from the residual cellulose and hemicellulose present in the AEL, and the lack of identifiable phenolics further suggests that the water fraction may contain mostly residual carbohydrate derivatives. Collectively, results from this study support that lignin based phenolics can be preferentially separated based on polarity during sequential extraction.

When examining the mass balance and the monomer yields together, our data indicates a net loss of total monomers after extraction. Given the mass balance percentages in **Table 3.1** and total monomer percentages in **Figure 3.3**, when summing the total monomer content in the SEFs there should only be 19.19 wt% of monomers in the raw bio-oil. This is different from the 32.44 wt% of monomers reported for the raw bio-oil (**Figure 3.3**), which represents a ~40% reduction. The apparent loss of monomers can be attributed to phenolic degradation, repolymerization, or evaporation as a result of removing the solvent during the drying step. Previous studies comparing drying methods of plant extracts/oils have reported losses in total phenolics when drying under vacuum or at temperatures above 40°C [231, 232]. Since the bio-oil fractions were dried in a vacuum oven at 60°C, this could explain the loss of monomers. Furthermore, due to the highly reactive nature of the lignin monomers after depolymerization, lignin condensation could be another reason causing the loss of monomeric phenolics [233-236]. Despite the recent work on preventing lignin condensation during thermochemical lignin decomposition and bio-oil aging [233, 234], these efforts are out of the scope of this study. Future investigation

is warranted to look at the stability of the recovered bio-oil fractions and methods to prevent lignin condensation.

Similar level of mass losses was seen in individual monomers probably caused by the drying and condensation reactions as well. For example, the raw bio-oil was comprised of 13.12% w/w of syringyl propane and when adding up all the SEF's syringyl propane content, in terms of their extraction efficiency, there would be 8.39% w/w of syringyl propane in the raw bio-oil. This equates to a 35% reduction, which is comparable to the 40% w/w reduction in total monomers described above. However, other monomers such as 4-ethylphenol, 4-propylguaciol, and methyl 4-hydroxyhydrocinnamate experienced a much larger reduction in weight at around 75-85%. Since they accounted for a much lower percentage of the raw bio-oils weight compared to syringyl propane (<5% individually), weight loss of those compounds may reflect into greater % reductions. Therefore, further investigations will need to optimize the drying stages after bio-oil recovery to maximize and create more consistent monomer yields.

Antimicrobial activity

The raw bio-oil and SEFs were tested for antimicrobial properties against Gram-positive bacteria (*B. subtilis*, *L. amylovorus*, and *S. epidermidis*), Gram-negative bacterium (*E. coli*), and yeast (*S. cerevisiae*) by examining differences in growth measured as OD_{600}. These microorganisms were chosen because they represent important production strains (*S. cerevisiae*) or contamination organisms involved in the corn ethanol biorefinery (*L. amylovorus*), medical (*S. epidermidis*) and food processing (*B. subtilis*) environments where antimicrobials are commonly needed [237-239].

Figure 3.5 shows a heat map of the degree of inhibition of the raw bio-oil at different concentrations against all tested organisms. Except for the lactobacillus species, the raw bio-oil was effective at inhibiting all microbes for >90% reduction in growth at concentrations ≤ 3 mg/mL. *L. amylovorus* did show ~70% reduction of growth at 2-2.5 mg/mL. Since the authors noticed a slight emulsion being formed at higher bio-oil concentrations (3-4 mg/ml) in the MRS broth culturing the lactobacillus, they attribute the apparent decrease in activity to the emulsion being read in the OD_{600} values compared to actual microbial growth. Nonetheless, the raw bio-oil was effective at inhibiting the other microorganisms, especially against the yeast and other Gram-positive bacteria (*B. subtilis* and *S. epidermidis*) at lower concentrations compared to the Gram-negative bacterium (*E. coli*). Other studies have also determined that Gram-positive bacteria are more susceptible to phenolics derived from plant extracts and wood vinegars, due to the absence of an outer membrane of lipoprotein and lipopolysaccharides that regulate access of antimicrobials into the underlying cell structure [127, 240]. Based on the GC/MS data, the raw bio-oil at concentrations of 1.5-3 mg/mL would contain ~0.4-0.9 mg/mL of monomeric phenolics; while previous work on wood vinegars have shown that phenolic content as low as 0.06-0.32 mg/mL can inhibit growth of Gram-positives and even some fungi [127]. This clearly illustrates the importance that monomers may have on the antimicrobial properties of the raw bio-oil.

	Raw Bio-Oil mg/ml							
	0.5	**1**	**1.5**	**2**	**2.5**	**3**	**4**	
B. subtilis	47.14±3.04[A]	71.36±3.90[AB]	84.97±15.03[BC]	100[C]	100[C]	100[C]	100[C]	
E. coli	49.65±4.95[A]	50.13±3.52[A]	46.93±0.59[A]	72.60±1.27[B]	53.46±1.94[A]	100[C]	100[C]	
S. epidermidis	41.75±3.86[A]	38.89±7.42[A]	47.77±10.32[AB]	65.85±2.35[BC]	79.86±4.84[C]	100[D]	100[D]	
S. cerevisiae	41.18±5.03[A]	69.27±4.18[B]	68.99±12.06[B]	100[C]	100[C]	100[C]	100[C]	
L. amylovorus	5.59±1.88[A]	15.65±2.57[A]	41.27±5.32[B]	64.70±3.49[C]	69.90±1.95[C]	59.43±5.45[BC]	57.33±6.31[BC]	

Percent Reduction:
- 80-100%
- 60-79%
- 40-59%
- 20-39%
- 1-19%

Figure 3.5: Heat map showing the percent reduction in growth for all microorganisms tested against different concentrations of raw bio-oil. Letters indicate differences at 95% confidence where values are mean±SE (n=3), using one-way ANOVAs and Tukey's pairwise comparisons. Percent reduction in growth for all other SEFs can be found in **Supplemental Table 3.S2**.

Figure 3.6 summarizes the raw bio-oil and SEF's antimicrobial activity by illustrating the concentrations of bio-oil that inhibited microbial growth by $\geq$ 90%. In general, the SEFs show a decrease in antimicrobial activity with a decrease in total identifiable monomers. For example, hexane has the highest percentage of monomers (36.57 wt%) compared to all other SEFs and it showed complete inhibition of all organisms at concentrations $\leq$ 3 mg/mL (**Figure 3.6**). However, the chloroform fraction, which had significantly less monomers, shows inhibition at lower concentrations than the hexane fraction against *L. amylovorus* and *B. subtilis*. Since chloroform extracted the largest amount of the bio-oils weight and contains a larger percentage of oligomers, its antimicrobial activity may be driven by oligomers compared to monomers. Similarly, the ethyl acetate fraction showed a $\geq$ 90% reduction in *S. epidermidis* growth at 0.5 mg/ml, which is lower than hexane, but the other organisms appear more tolerant to the ethyl acetate fraction with only 40%-75% reduction in growth even at the highest bio-oil concentration of 4 mg/ml. This could indicate the unidentifiable compounds are causing a species-specific interaction against *S. epidermidis*. Meanwhile, the water fraction showed no inhibition at 4 mg/mL against *B. subtilis*, *S. cerevisiae*, and *L. amylovorus*, but did have

a 57% reduction in growth for *E. coli* and 60% for *S. epidermidis*. The loss of antimicrobial activity for the water fraction in most of the organisms could be due to its extremely low phenolic content (0.81%) and presence of residual sugars.

In a previous study, we tested monensin as a model antibiotic. Monensin has been shown to have selectivity inhibition against *L. amylovorus* but does not inhibit the growth of *S. cerevisiae* (Oliva Neto, Lima et al. 2014). Our test revealed that monensin at the lowest concentrations of 0.0004 mg/mL completely inhibits the growth of *L. amylovorus* but has minimal effect on *S. cerevisiae* and *E. coli* (**Supplemental Figure 3.S2**). Since the traditional antibiotics are used at a much lower concentration for microbial inhibition, we do not believe that it is directly comparable to bio-oils tested in this study at the mg/ml level.
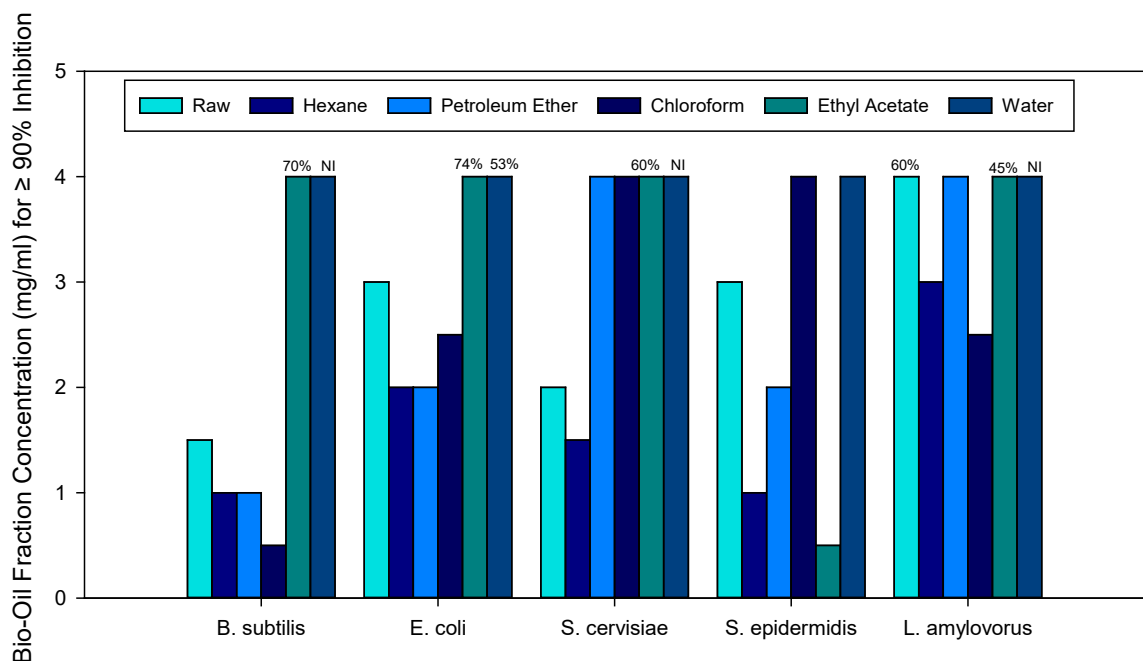
Figure 3.6: Raw bio-oil and SEF concentrations causing greater than 90% inhibition compared to the control (NI = no inhibition, and percent values are degree of inhibition at highest concentration tested). All percent reduction in growth values at every bio-oil concentration tested for the raw bio-oil and SEFs can be found in **Supplemental Table 3.S2**.

There is a general trend that monomer yields seem to be a driver of antimicrobial properties in the bio-oils. For example, when comparing the raw bio-oil with the hexane and petroleum ether fractions there is a direct correlation to syringyl propane content and antimicrobial activity against the bacteria. The hexane fraction has a syringyl propane content of 26.77 wt% of bio-oil, petroleum ether 16.57%, and the raw bio-oil 13.12% (**Figure 3.3**), so syringyl propane content is in the following order: hexane>petroleum ether>raw bio-oil. Against each bacterium, the minimum bio-oil concentration causing ≥ 90% reduction in growth is lowest in hexane and highest in the raw bio-oil, which follows the same trend as syringyl propane content. Based on this observation, syringyl propane was synthesized and tested against each microorganism for antimicrobial properties at the

same concentrations as the bio-oils. **Supplemental Table 3.S3** shows the percent inhibition values for syringyl propane with a methods section describing its synthesis. It can be seen that syringyl propane was effective at inhibiting 100% of growth for each microorganism at concentrations less than 4 mg/ml. Specifically, *S. epidermidis*, *S. cerevisiae*, and *L. amylovorus* were affected the greatest as they were completely inhibited at 1-1.5 mg/ml, whereas *B. subtilis* and *E.coli* were completely inhibited at 3 and 4 mg/ml, respectively. This data illustrates that syringyl propane has significant antimicrobial properties, but the concentration of syringyl propane in each bio-oil fraction at an overall bio-oil concentration that completely inhibited each microorganism is still much lower than the concentration when syringyl propane was tested alone. For example, *B. subtilis* was inhibited at a concentration of 1 mg/ml by the hexane fraction that contains ~ 0.26 mg/ml of syringyl propane based on GC/MS results, but pure syringyl propane needs to be at a concentration of 3 mg/ml to completely inhibit *B. subtilis*. Thus, it is believed that the synergism between the mixture of compounds identified and unidentified could still be a major driver for the bio-oils antimicrobial properties.

On the other hand, the chloroform and ethyl acetate fractions did not contain a large concentration of syringyl propane, but they did clearly show a larger distribution of dimers, trimers, and larger oligomers in their MALDI spectra compared to the other SEFs, albeit unquantified or identified, so these larger oligomers could also be a driver and source of these bio-oils antimicrobial activity. Furthermore, despite the fact that many of the monomers present were individually at lower concentrations than previously reported to having antimicrobial activity [125, 127, 128, 137, 241], previous work has shown that when mixing phenolics at low concentrations their activity was higher than the individual

89

monomers, which suggests the existence of positive synergism when multiple phenolics are present [242]. Therefore, these data illustrate that each bio-oil is a complex and diverse cocktail of monomers and unidentified or quantified oligomers, so there may be unknown synergisms or reactions occurring with the microorganisms that drive the antimicrobial activity seen here than just individual monomers. To this end, future work should examine various mixtures of the identifiable monomers found here, at low concentrations, to elucidate their antimicrobial interactions. With the absence of larger oligomers, these types of experiments could determine the true importance of the unidentified oligomers and interactions of model monomers in these bio-oil's antimicrobial properties.

In order to infer the mode of action of the raw bio-oil, representative Gram-negative (*E. coli*) and Gram-positive (*B. subtilis*) bacteria were grown in the presence of raw bio-oil and assessed for potential membrane damage by staining with SYTO9 and propidium iodide (PI). SYTO9 is a green fluorescent dye that penetrates cell membranes freely and has moderate affinity to their nucleic acids, while propidium iodide is a red fluorescence dye that can only penetrate damaged membranes and due to its higher affinity for nucleic acids it can displace the weaker bound SYTO9 dye causing the damaged cell to show a strong red fluorescence instead [221]. Therefore, cells that fluoresce green (SYTO9) represent live cells without membrane damage and cells that fluoresce red (PI) can be considered membrane damaged or dead. **Figure 3.7 (A-D)** shows images of the bacterial cells stained with PI and SYTO9 after growing for 5 h at 37ºC with and without raw bio-oil (4 mg/ml). Additionally, **Figure 3.7 (E-F)** shows the ratio of SYTO9/PI fluorescence, representing the ratio of live cells to membrane damaged/dead cells, for both *E. coli* and *B. subtills* incubated with varying concentrations of raw bio-oil for 5 h at 37ºC.
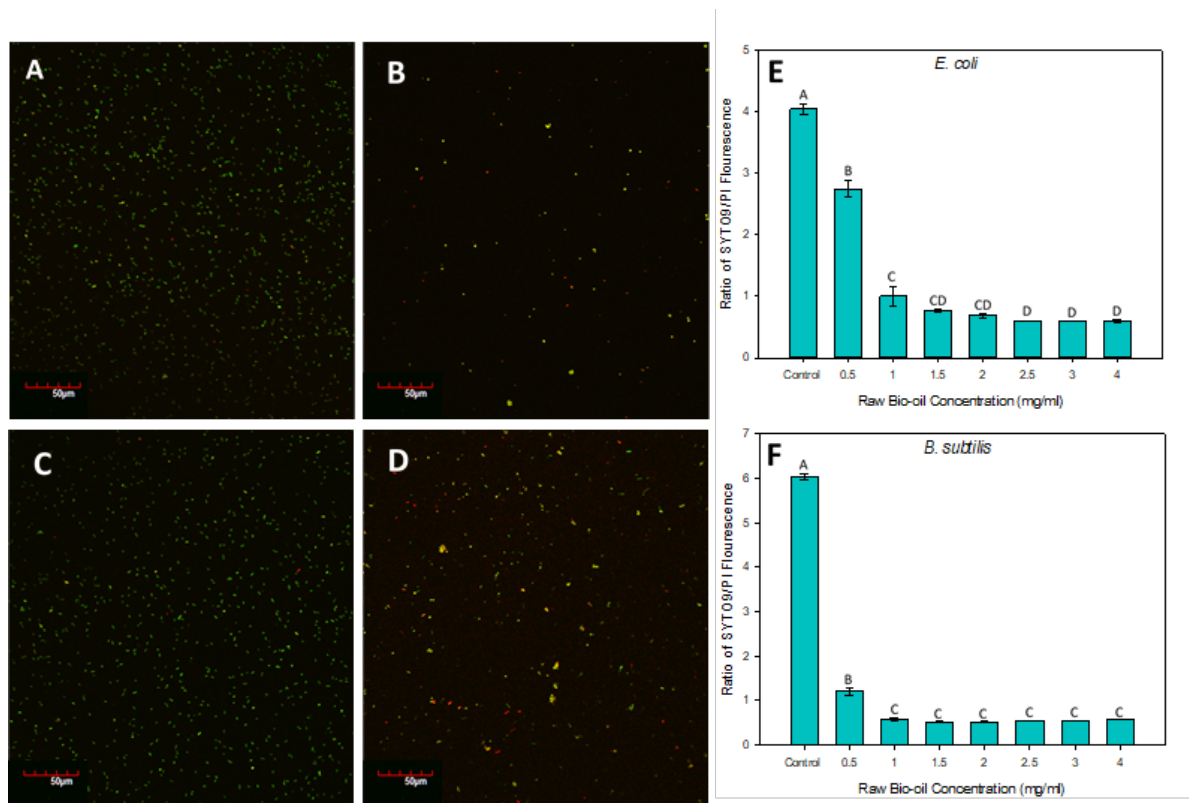
Figure 3.7: Fluorescence (red and green) images of *E. coli* (A,B) and *B. subtilis* (C,D) grown without bio-oil (A,C) and with 4 mg/mL of raw bio-oil (B,D) for 5h at 37ºC stained using SYTO9 (green) and propidium iodide (PI, red). (E,F) SYTO9/PI fluorescence ratios for *E. coli* and *B. subtilis* treated with varying concentration of raw bio-oil which indicates the ratio of live/dead or undamaged/membrane-damaged cells. Letters on the bars indicate differences at 95% confidence where values are mean±SE (n=3), using one-way ANOVAs and Tukey's pairwise comparisons.

The data show a statistically significant decrease in SYTO9/PI ratio comparing the control with cells in the presence of raw bio-oil, where the ratio decreased from ~4.5 and 7.4 (control) to 0.6 and 0.5 (at bio-oil concentration of 4 mg/ml) for *E. coli* and *B. subtills*, respectively. This along with the microscopic imaging clearly shows a significant increase in the proportion of cells that fluoresce red and are assumed to have PI-permeable membranes primarily due to death or damaged membranes [221]. Lignin derivatives have been thought to directly cause cell membrane damage or have ionophoric activity that ultimately results in cell lysis and death [122, 243]. However, since both Gram-negative

and positive bacterial cells have increases in PI stained cells when exposed to the raw bio-oil, this suggests that the outer membranes of Gram-negatives that can confer insensitivity to ionophores were not sufficient in providing defense to the bio-oils [171, 244]. Subsequently, this indicates that the raw bio-oil does not have an ionophoric mode of action but could have a non-specific physical interaction with the bacterial membranes that results in membrane permeability (morphological), physical damage, or direct cell death [245, 246]. Furthermore, the cells of both Gram-negative and positive bacteria showed a greater amount of cells stained with PI compared to SYTO9 at high bio-oil concentrations, suggesting that the bio-oil displays more bactericidal activity than bacteriostatic activity [247], due to the direct cell death caused by cell membrane damage compared to just inhibiting cell growth. Therefore, in summary, these results suggest total monomer concentration and the presence of specific monomers (i.e., syringyl propane) show correlations to antimicrobial activity, but the exact mode of action remains unclear and the antimicrobial activity of unidentified or quantified oligomers/compounds remains to be further investigated.

From an applications standpoint, the bio-oil and SEFs show a very general antimicrobial action that is not specific to either Gram-negative or positive bacteria, nor fungi (yeast). This would exclude them from being used in highly specific antimicrobial roles, such as preventing contamination in ethanol fermentation. Compared to traditional antibiotics like ampicillin, kanamycin, monensin, and virginiamycin, which are needed at the ppm level to completely inhibit microbial growth, the bio-oils tested here may not be suitable for commercial applications based on the high concentrations needed for microbial inhibition. Additionally, to be considered for medical or food related uses, further work in

using non-toxic extraction solvents and catalysts must be done, and the bio-oils must be examined for damage and mutagenic effects on mammalian cells. Despite these shortcomings, this study still provides insights into the general antimicrobial properties of lignin derived compounds that could guide future developments.

Conclusions

Purified alkali-enzymatic corn stover lignin (AEL) was depolymerized by catalytic transfer hydrogenolysis using supercritical ethanol and a Ru/C catalyst, generating a bio-oil stream at high yields. Sequential extraction using hexane, petroleum ether, chloroform, and ethyl acetate extracted the raw bio-oil into 5 fractions at 50.7-5.8 wt% yields of total bio-oil in the order of chloroform > hexane > petroleum ether ≈ water > ethyl acetate. Extraction efficiency followed the trend that the first solvent used in each change in polarity during sequential extraction had the highest percentage of products extracted. Molecular weights of the raw bio-oil and sequential extraction fractions (SEF) were much lower than the purified AEL, demonstrating depolymerization of lignin into low molecular weight products. The monomers in the bio-oil fractions contained primarily alkylated phenols, hydrogenated hydroxycinnamic acid derivatives, syringol and guaiacol-type lignins. Results suggest that the total monomer concentration and the presence of specific monomers (i.e., syringyl propane) may correlate to the antimicrobial activity, but the exact mode of action or antimicrobial activity caused by the complex mixtures of monomers and unidentified oligomers remains unclear. This study provides insights into the types of lignin derived compounds that confer antimicrobial activity and that compounds can be preferentially extracted from lignin bio-oils using LLE method.

Associated Content

Acknowledgements

Figure 3.S1: DHAP/Li+ MALDI spectrum (0–2000 m/z) of the raw bio-oil (B) and SEFs (C-G), compared to the background spectra of the matrix (A), DHAP/Li+ MALDI spectrum (0–2000 m/z) of the raw bio-oil (B) and SEFs (C-G), compared to the background spectra of the matrix (A).

Table 3.S1: List of compounds in raw bio-oil and sequentially extracted fractions identified by GC-MS and their yields based on wt% of the bio-oil fraction

Fraction	Numeric ID from Figure 3	Compound	Yield (wt%)
Raw Bio-Oil	1	Syringyl propane	13.12
	2	4-ethyl-phenol	4.63
	3	4-propylguaiacol	4.62
	4	Methyl 4-hydroxyhydrocinnamate	3.04
	5	4-ethylguaiacol	2.70

	6	Ethyl-β-(4-hydroxy-3-methoxy-phenyl)-propionate	0.78
	7	2,6-dimethoxy-4-ethylphenol	0.66
	8	Creosol	0.61
	19	p-Cresol	0.31
	19	Homovanillyl alcohol	0.30
	19	Syringol	0.28
	19	1-Propanone, 1-(2,4-dimethoxyphenyl)-	0.22
	19	Guaicol	0.22
	19	Ethyl vanillate	0.21
	19	4-propyl-phenol	0.18
	19	4-Allylsyringol	0.16
	19	Eugenol	0.14
	19	Phenol	0.12
	19	Ethyl homovanillate	0.11
		Total	**32.44**
Hexane	1	Syringyl propane	26.77
	3	4-propylguaiacol	3.93
	2	4-ethyl-phenol	1.36
	7	2,6-dimethoxy-4-ethylphenol	1.29
	6	Ethyl-β-(4-hydroxy-3-methoxy-phenyl)-propionate	0.87
	5	4-ethylguaiacol	0.63
	9	2,5-Dimethoxybenzyl alcohol	0.57
	19	Methyl 4-hydroxyhydrocinnamate	0.39
	19	4-methoxy-propylbenzene	0.27
	19	4-((1E)-3-Hydroxy-1-propenyl)-2-methoxyphenol	0.15
	19	1-Propanone, 1-(2,4-dimethoxyphenyl)-	0.13
	19	4-propyl-phenol	0.10
	19	Syringol	0.06
	19	3-(1,1-dimethylethyl)-4-methoxy-phenol	0.05
		Total	**24.65**
Petroleum Ether	1	Syringyl propane	16.57
	4	Methyl 4-hydroxyhydrocinnamate	2.43
	6	Ethyl-β-(4-hydroxy-3-methoxy-phenyl)-propionate	1.91
	7	2,6-dimethoxy-4-ethylphenol	0.81
	10	2-methyl-5-(1-methylethyl)-phenol	0.49
	19	Homovanillyl alcohol	0.45
	19	3-ethoxy-2-hydroxy-benzaldehyde	0.31
	19	1,2,4-Trimethoxybenzene	0.31
	19	3,4,5-Trimethoxyphenylacetic acid	0.23
	19	Ethyl vanillate	0.23
	19	Ethyl homovanillate	0.22

	19	4-propylguaiacol	0.22
	19	4-Allylsyringol	0.19
	19	3-(3-Methoxyphenyl)propionic acid ethyl ester	0.12
	19	4-ethyl-phenol	0.06
	19	4-propylguaiacol	0.06
	19	Syringol	0.02
	19	4-Ethylguaiacol	0.01
	19	4-propyl-phenol	0.01
	19	4-butyl-phenol	0.01
	19	2-ethyl-4,5-dimethyl-phenol	0.01
		Total	**24.65**
Chloroform	11	Ethyl homovanillate	4.68
	12	Homovanillyl alcohol	4.34
	13	Ethyl vanillate	0.80
	14	Dihydromethyleugenol	0.59
	5	4-ethyl-phenol	0.58
	15	(3,4-Dimethoxyphenyl)-methoxymethanol	0.53
	19	Syringol	0.47
	19	1,2,4-Trimethoxybenzene	0.40
	19	4-Allylsyringol	0.28
	19	2,6-dimethoxy-4-ethylphenol	0.24
	19	Benzeneacetic acid, 4-hydroxy-, ethyl ester	0.17
	19	Ethyl-β-(4-hydroxy-3-methoxy-phenyl)-propionate	0.11
	19	3-hydroxy-4-methoxy-benzaldehyde	0.10
	19	3,4-Diethoxyphenylacetic acid	0.10
	19	3,4,5-Trimethoxyphenylacetic acid	0.09
	19	Syringyl propane	0.08
	19	Methyl (3,4-dimethoxyphenyl)(hydroxy)acetate	0.08
	19	Apocynin	0.08
	19	3-ethoxy-2-hydroxy-benzaldehyde	0.07
	19	5-methoxy-2,3-dimethyl-phenol	0.07
	19	4-hydroxy-3-methoxy-Phenylacetylformic acid	0.07
	19	Methyl-(2-hydoxy-3-ethoxy-benzyl)ether	0.07
	19	4-Ethylguaiacol	0.06
	19	Dihydroeugenol	0.05
	19	p-Cresol	0.04
	19	Methyl 4-hydroxyhydrocinnamate	0.03
	19	1-(4-hydroxy-3-methoxyphenyl)-2-Propanone	0.02
	19	3-Methoxy-5-methylphenol	0.02
	19	3-Methylguaiacol	0.01
	19	2-propyl-phenol	0.01
		Total	**14.24**
Ethyl Acetate	16	Methyl 4-hydroxyhydrocinnamate	1.78

	12	Homovanillyl alcohol	0.52
	1	Syringyl propane	0.50
	13	Ethyl vanillate	0.36
	17	β-(4-Hydroxy-3-methoxyphenyl)propionic acid	0.30
	15	(3,4-Dimethoxyphenyl)-methoxymethanol	0.27
	19	3,5-Dimethoxy-4-hydroxyphenylacetic acid	0.22
	19	Ethyl-β-(4-hydroxy-3-methoxy-phenyl)-propionate	0.15
	19	1,2-Dimethoxy-4-n-propylbenzene	0.14
	19	Dihydroeugenol	0.09
	19	1,2,4-Trimethoxybenzene	0.08
	19	Vanillin	0.08
	19	2,6-dimethoxy-4-ethylphenol	0.07
	19	Homovanillyl alcohol	0.07
	19	4-Allylsyringol	0.06
	19	Apocynin	0.05
	19	4-ethyl-phenol	0.05
	19	3-(p-Hydroxyphenyl)-1-propanol	0.05
	19	Syringol	0.05
	19	Ethyl homovanillate	0.04
	19	Homovanillic acid	0.03
	19	Phenyl Vinyl Ether	0.02
	19	1-(4-hydroxy-3-methoxyphenyl)-2-Propanone	0.02
	19	4-Ethylguaiacol	0.02
	19	Tyrosol	0.01
	19	1-ethoxy-2-methoxy-4-methylbenzene	0.01
		Total	**5.04**
Water	17	β-(4-Hydroxy-3-methoxyphenyl)propionic acid	0.35
	4	Methyl 4-hydroxyhydrocinnamate	0.20
	18	Tyrosol	0.18
	14	Homovanillyl alcohol	0.08
		Total	**0.81**

Table 3.S2: Percent reduction of growth for microorganisms tested against different concentrations of the raw bio-oil and SEFs. Letters indicate differences at 95% confidence where values are mean±SE (n=3), using one-way ANOVAs and Tukey's pairwise comparisons across the different concentrations for each microorganism.

		Percent Reduction in Growth						
		Bio-oil Concentration (mg/ml)						
Fraction	Organism	0.5	1	1.5	2	2.5	3	4
Raw	*B. subtilis*	47.14±3.04^A	71.36±3.90^{AB}	84.97±15.03^{BC}	100^C	100^C	100^C	100^C
	E. coli	49.65±4.95^A	50.13±3.52^A	46.93±0.59^A	72.60±1.27^B	53.46±1.94^A	100^C	100^C
	S. epidermidis	41.75±3.86^A	38.89±7.42^A	47.77±10.32^{AB}	65.85±2.35^{BC}	79.86±4.84^C	100^D	100^D
	S. cerevisiae	41.18±5.03^A	69.27±4.18^B	68.99±12.06^B	100^C	100^C	100^C	100^C
	L. amylovorus	5.59±1.88^A	15.65±2.57^A	41.27±5.32^B	64.70±3.49^C	69.90±1.95^C	59.43±5.45^{BC}	57.33±6.31^{BC}
Hexane	*B. subtilis*	87.47±3.30^A	98.23±1.77^B	100^B	100^B	100^B	100^B	100^B
	E. coli	19.77±2.98^A	62.55±3.35^B	80.92±4.97^C	88.15±3.20^{CD}	100^D	100^D	100^D
	S. epidermidis	37.20±12.60^A	85.99±1.02^B	100^B	100^B	100^B	100^B	100^B
	S. cerevisiae	11.42±5.96^A	72.60±3.64^B	90.91±3.41^{CD}	85.04±1.71^D	98.75±1.25^C	100^C	100^C
	L. amylovorus	NI	5.44±1.62^A	27.60±3.92^B	55.32±2.57^C	57.13±7.93^C	100^D	100^D
Petroleum Ether	*B. subtilis*	24.50±6.50^A	85.38±3.10^B	100^B	100^B	100^B	100^B	100^B
	E. coli	23.35±10.90^A	21.84±8.52^A	28.71±5.11^A	65.74±9.08^B	100^B	100^B	100^B
	S. cerevisiae	17.52±3.93^A	32.69±3.92^{AB}	50.53±3.49^B	74.02±0.65^C	82.89±2.76^C	85.03±5.26^C	100^D
	S. epidermidis	67.48±7.56^A	54.25±0.34^A	52.58±4.48^A	94.59±5.41^B	100^B	100^B	100^B
	L. amylovorus	NI	24.04±3.54^{AB}	54.64±2.21^{AB}	55.49±6.38^{AB}	65.91±11.68^{AB}	62.54±11.29^{AB}	89.70±6.54^B
Chloroform	*B. subtilis*	89.47±2.72^A	86.65±4.12^A	92.06±1.63^A	89.09±2.57^A	93.84±1.36^A	88.32±2.34^A	93.69±2.26^A
	E. coli	42.07±7.94^A	54.36±3.51^{AB}	57.21±3.67^{AB}	64.94±4.07^B	97.11±2.50^C	100^C	100^C
	S. cerevisiae	20.62±6.92^A	38.35±9.77^{AB}	32.11±5.96^{AB}	28.69±6.82^A	41.00±7.31^{AB}	62.02±6.82^B	100^C

	S. epidermidis	46.53±4.53[A]	66.57±3.50[AB]	57.61±4.33[AB]	68.23±2.94[AB]	69.33±0.69[AB]	78.49±8.94[BC]	100[C]
	L. amylovorus	NI	24.69±5.76[B]	58.35±4.31[C]	77.94±3.29[D]	88.61±3.75[DE]	93.50±4.47[DE]	100[E]
Ethyl Acetate	*B. subtilis*	26.04±2.12[A]	22.80±4.14[A]	40.06±3.98[AB]	39.11±9.27[A]	26.22±7.86[A]	33.14±3.14[A]	70.28±3.18[B]
	E. coli	12.12±7.90[A]	12.52±6.43[A]	46.21±3.30[AB]	68.88±19.81[B]	63.35±5.96[B]	72.68±3.71[B]	73.98±0.83[B]
	S. cerevisiae	NI	5.16±0.40[A]	33.53±1.00[B]	41.02±0.12[BC]	37.73±1.47[B]	40.67±6.23[BC]	59.21±6.08[C]
	S. epidermidis	78.24±8.73[A]	89.33±3.04[A]	89.77±8.27[A]	100[A]	100[A]	100[A]	100[A]
	L. amylovorus	6.27±2.44[A]	27.63±2.83[B]	53.16±2.61[CD]	40.95±3.51[BC]	57.64±3.46[D]	43.68±3.84[BCD]	44.26±1.21[CD]
Water	*B. subtilis*	NI	NI	NI	NI	NI	NI	NI
	E. coli	11.00±4.26[A]	7.06±1.40[A]	35.31±5.59[B]	18.23±2.34[AC]	34.84±1.08[BC]	53.86±0.90[D]	42.46±1.79[BD]
	S. cerevisiae	NI	NI	NI	NI	NI	NI	NI
	S. epidermidis	NI	NI	NI	16.16±2.12[A]	42.77±2.17[B]	83.79±1.61[C]	100.00±0.00[D]
	L. amylovorus	NI	NI	NI	NI	NI	NI	NI

Table 3.S3.3: Percent reduction of growth for microorganisms tested against different concentrations of pure syringyl propane. Letters indicate differences at 95% confidence where values are mean±SE (n=3), using one-way ANOVAs and Tukey's pairwise comparisons across the different concentrations for each microorganism.

	Percent Reduction in Growth						
	Syringyl Propane Concentration (mg/ml)						
Organism	**0.5**	**1**	**1.5**	**2**	**2.5**	**3**	**4**
B. subtilis	18.41±0.81 [A]	48.07±1.24 [B]	62.59±3.95 [C]	60.82±4.07 [C]	74.78±1.23 [D]	100 [E]	100 [E]
E. coli	22.79±1.71 [A]	52.14±3.71 [B]	69.38±2.26 [C]	70.33±1.43 [C]	71.05±6.65 [C]	79.76±3.95 [C]	100 [E]
S. epidermidis	19.06±4.16 [A]	100 [B]	100 [B]	100 [B]	100 [B]	100 [B]	100 [B]
S. cerevisiae	18.42±4.19 [A]	82.61±4.08 [B]	100 [C]	100 [C]	100 [C]	100 [C]	100 [C]
L. amylovorus	25.58±2.08 [A]	100 [B]	100 [B]	100 [B]	100 [B]	100 [B]	100 [B]

Synthesis of syringyl propane:

The synthesis method was adopted from Lundevall et. al. [248]with minor changes:

4-Allyl-2,6-dimethoxyphenol (1mmol) was dissolved in 5mL methanol in a 10 mL round bottom flask. Cobalt sulfate heptahydrate (1mmol, 281mg) was dissolved in 0.7 mL water and added to the solution. The reaction mixture was cooled to 0° C using an ice bath and $NaBH_4$ (4mmol, 151mg) was gradually added to the solution. Post addition, the reaction mixture was removed from the ice bath and was stirred for an additional 5 more minutes at room temperature. The solution was filtered and washed twice with 10 mL ethyl acetate. The filtrate was extracted with water and ethyl acetate (three times) and the organic phase was collected, combined and dried over sodium sulfate. After evaporation of the solvent, 128 mg syringyl propane was obtained (95 % pure by GC-MS) as a clear liquid. A QExactive mass spectrometer was used to obtain high resolution accuratemass electrospray mass spectrum of the product as a lithum cation adduct that showed an observed m/z 203.1242 consistent with $[C_{11}H_{16}O_3+Li^+]^+$ (expected m/z 203.1254, Δppm -5.7247).
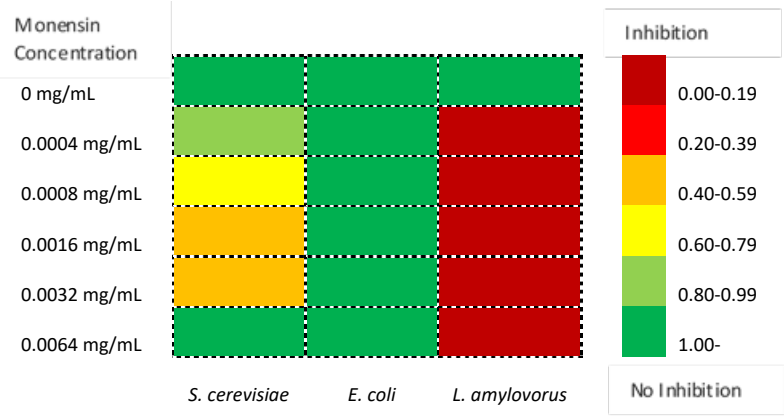
Figure 3.S2: Heat map showing the percent difference in growth for microorganisms tested against different concentrations of monensin, adapted from Dodge, et al. [215]

CHAPTER 4. CONTROLLING BACTERIAL CONTAMINATION DURING FUEL

ETHANOL FERMENTATION BY UTILIZING THERMOCHEMICALLY

DEPOLYMERIZED LIGNIN BIO-OILS

Abstract

Lactic acid bacteria (LAB) are the most problematic contamination source during fuel ethanol fermentation, which can contribute to significant losses in ethanol yields. While fermentation facilities attempt to prevent contamination through extensive sanitation practices, the most effective contamination preventatives are antibiotics that selectivity inhibits the LAB. However, due to the overuse of these antibiotics, there is an increased incidence of antibiotic-resistant bacterial strains. Since lignin is a polyphenolic complex derived of phenolic subunits that can confer antimicrobial properties, with potential ionophoric modes of action (i.e. selective inhibition of Gram-positives), this study aims to depolymerize lignin oxidatively into a bio-oil with selective antimicrobial properties that can replace antibiotics in fuel ethanol fermentation. Oxidative depolymerization of corn stover lignin with peracetic acid saw over 35% conversion to bio-oil yields and demonstrated up to 90% inhibition of commercially sampled LAB at 4 mg/ml with no inhibition against fermentation yeast. The highly selective antimicrobial properties of the bio-oil are attributed to an ionophoric or membrane damaging mode of action that results in cell death, based on fluorescent staining. No inhibition of enzymatic activity for α-amylase or glucoamylase was also observed. Using the bio-oil (4 mg/ml) as an alternative antibiotic treatment during SSF of raw corn starch showed an increase in ethanol production as bacterial contamination increased. At the highest contamination ratio of 1:100 yeast to lactobacillus, the bio-oil increased ethanol production by 8% compared to the contamination control. This study illustrates the efficacy of using a lignin bio-oil as an antibiotic replacement during fuel ethanol fermentation to reduce contamination by LAB.

Introduction

The U.S. alone produces over 16 billion gallons of fuel ethanol each year and because fuel ethanol is primarily produced from the fermentation of corn, anything that limits yeast viability will cause significant economic losses (Energy Information Administration). Since fuel ethanol fermentations are not produced under completely aseptic conditions, chronic and acute bacterial contaminations can occur [249, 250]. Lactic acid bacteria (LAB) are considered to be the most problematic due to their production of by-products such lactic acids, polysaccharides and gummy biofilms that reduce yeast viability [251, 252]. Furthermore, LAB proliferation in fermentation reactors [250], consumes essential micronutrients and sugar required for optimal yeast growth and ethanol production. Therefore, these bacterial contaminations reduce ethanol yields and can result in "stuck" fermentations that cause costly shutdowns of facilities for cleaning [253]. While fermentation facilities attempt to prevent contamination through extensive sanitation practices, there are so many reservoirs of bacterial contamination that one of the most effective contamination preventatives is antibiotics [251, 253].

Virginiamycin is one of the most commonly used antibiotics in fuel ethanol production in the U.S. [254]. However, due to the overuse of antibiotics, there is an increased incidence of antibiotic-resistant LAB strains isolated from dry-grind ethanol plants [255]. These antibiotics have been shown to persist in downstream coproducts like distillers' grains [255, 256], which is becoming a major concern for consumers of livestock that are fed the dietary supplement. Therefore, efforts in the development of new antimicrobial agents with good biodegradability and high selectivity against LAB are needed to circumvent these issues.

Lignin is one of the most abundant naturally occurring sources of phenolic polymers on earth and is currently considered a major waste product in the paper and pulp industries and industrial lignocellulosic biorefineries [44]. Since lignin is a polyphenolic complex, much research has shown that its phenolic subunits can confer antimicrobial properties [122, 207]. Lignin's antimicrobial properties are dictated by the source of the lignin, its extraction methods, and chemical structure (i.e. monomers, oligomers and functional groups) [12, 122]. Nonetheless, it is believed that lignin phenolics can increase the ion permeability of cell membranes in microorganisms through ionophoric activity, causing cell lysis [171, 172]. Since ionophores are highly selective against Gram-positive bacteria compared to eukaryotes or Gram-negatives that have outer membranes that confer insensitivity to ionophores, lignin phenolics with similar selective antimicrobial properties would be ideal for selectively inhibiting LAB in fermentation systems. Additionally, while a variety of technical lignins (i.e. Kraft lignin and organosolv lignin) have had notable antimicrobial properties, smaller depolymerized lignin oligomers and phenolic monomers are noted for increased antimicrobial activity [123]. Therefore, if a lignin depolymerization product can show highly selective antimicrobial activity, then that product can serve as an alternative to traditional antibiotics while simultaneously valorizing lignin waste streams.

Some of the most popularly studied depolymerization methods are pyrolysis, acid/base/metal catalyzed hydrolysis, hydrogenolysis and oxidation [9, 45, 46]. However, pyrolysis and hydrolysis are characterized by increased condensation and repolymerization reactions that significantly reduce bio-oil yields [49, 210], and while catalytic transfer hydrogenolysis provides increased bio-oil yields and more stable compounds it has energy intensive reactions that occur at high pressure and temperatures ranging from 200-300°C

[9, 59, 62]. On the other hand, oxidative procedures utilizing oxygen, hydrogen peroxide, or peroxyacids can be performed at significantly lower reaction temperatures (24-100°C) while still producing relatively high monomer yields [64]. More recent literature has focused on peracetic acid as an oxidizer due to its ability to cleave C-C and ether bonds, its higher monomer selectivity, high oil yields (18-22% w/w), and the fact that it is considered an environmentally benign oxidant [94, 95]. To that end, peracetic acid represents a viable lignin depolymerization strategy that could be low cost due to mild reaction conditions while maintaining high product yields.

Therefore, the main objective of this study is to utilize peracetic acid to depolymerize lignin and examine the resulting bio-oils antimicrobial activity for use in a fuel ethanol fermentation environment. Specifically, the goals are to 1) depolymerize alkali-treated corn stover lignin from an ethanol biorefinery into a low molecular weight bio-oil by utilizing mild oxidative procedures with peracetic acid, 2) test the antimicrobial properties of the lignin bio-oil on yeast and LAB, 3) examine the effects of the lignin bio-oil on enzyme function for both α-amylase and glucoamylase, and 4) determine the efficacy of using the lignin bio-oil as an antibiotic during the simultaneous saccharification and fermentation (SSF) of corn starch to reduce contamination by LAB. The results from this study provide insights into using depolymerized lignin derivatives as an antibiotic replacement for reducing contamination during fuel ethanol fermentation.

Experimental Methods

Lignin Purification

Corn stover was pretreated at the National Renewable Energy Laboratory (NREL) using 0.1g NaOH/ g biomass with 15% solid loading at 80°C for 2 hrs. The lignin residue (alkaline enzymatic lignin - AEL) was collected after enzymatic hydrolysis of the pretreated corn stover and fermentation and further purified to remove remaining carbohydrates by a precipitation method [32]. The resulting lignin was then freeze-dried using FreeZone 6-liter console freeze dry system (Labconco, Kansas City, MO), at -50°C under 0.1-0.2 mBar vacuum for 72 hrs.

Structural carbohydrates and lignin composition of the resulting purified AEL samples were determined by compositional analysis according to an NREL laboratory analytical procedure [180]. The sugar concentration was determined by HPLC (Ultimate 3000, Dionex Corporation, Sunnyvale, CA, US) equipped with a refractive index detector and using a Bio-Rad Aminex HPX-87H column and guard assembly.

Oxidative Depolymerization of Lignin

Oxidative depolymerization was carried out by following the procedures in an earlier study by Ma, Guo [63]. In short, the purified AEL was treated with peracetic acid (PAA) at a PAA dosage of 0.8g PAA/g lignin, with acetic acid used to dilute the reaction mixture to 5% solid loading. The reaction occurred at 60 °C for 1 hour while being mixed every 10 min. Once the reaction was completed, the reaction mixture was centrifuged at 4000 rpm to remove unreacted solids and the supernatant was mixed with water at a 1:4 ratio to create an aqueous phase prior to liquid-liquid extraction. The lignin

depolymerization compounds were extracted from the aqueous phase using ethyl acetate at a 1:4 ratio for three times. The ethyl acetate fractions were combined and dried under vacuum at 60 °C for 24 hrs to obtain the extracted lignin depolymerization compounds (namely bio-oil thereafter) that were then dissolved in ethanol and centrifuged to remove any undissolved solids. Bio-oil yield was determined by weighing the total bio-oil content dissolved in ethanol and dividing by the starting lignin weight.

Bio-Oil Characterization

The weight-average molecular weight (M_w) and the number-average molecular weight (M_n) of the purified AEL and PAA derived lignin bio-oils were determined using gel permeation chromatography (GPC) [216]. An Ultimate3000 HPLC system equipped with an Ultra Violet (UV) detector and Mixed-D PLgel column (5 μm particle size, 300 mm x 7.5 mm i.d., linear molecular weight range of 200 to 400,000 u, Polymer Laboratories, Amherst, MA) were utilized. Separation was accomplished in a mobile phase of tetrahydrofuran (THF) at a flow rate of 0.5 ml min^{-1}, at 50°C. Elution profiles were monitored at UV absorbance of 280 nm and calibrated using low molecular weight polystyrene standards (Product No. 48937, Sigma-Aldrich). Polydispersity Index (PDI) was calculated using the equation: PDI= Mw/Mn [216].

GC/MS was performed on the bio-oil to quantify monomer yields. The bio-oil was derivatized by first dissolving it in 0.5ml of pyridine then adding 0.5 ml of BSTFA and incubating at 50°C for 30 min. Monomers were identified and quantified by GC/MS using an Agilent 7890B GC coupled 5977B MS with an HP-5ms (60 m × 0.32 mm) capillary column. The temperature program started at 40 °C with a holding time of 6 min and

increased to 240 °C at 4 °C min^{-1} with a holding time of 7 min, finally the temperature was raised to 280 °C at 20 °C min^{-1} with a holding time of 8 min. Helium was used as a carrier gas with a flow rate of 1.2 mL min^{-1}. Helium was used as a carrier gas at a flow rate of 1.2 mL min^{-1}. Calibration curves were created using commercially available pure compounds: guaiacol, syringaldehyde, vanillin, and 4-propylphenol (Sigma Aldrich, St. Louis, MO, USA).

NMR was performed on the purified lignin and the bio-oil. Approximately 100 mg lignin sample was dissolved in DMSO-d$_6$/pyridine-d$_5$ (4:1) or DMSO-d$_6$ under mild heat and sonication in an NMR tube until a homogeneous mixture was obtained. NMR spectra were acquired on a 500 MHz JEOL ECZR (Peabody, MA, USA) NMR spectrometer equipped with a 5-mm Royal Probe. The central DMSO solvent peak was used as an internal reference (δ_C 39.5, δ_H 2.5 ppm). The ^{1}H–^{13}C correlation experiment was an HSQC experiment (JEOL pulse sequence 'hsqc_edit_dec_en') with 25% non-uniform sampling (NUS). HSQC experiments were carried out using the following parameters: acquired from 11 to -1 ppm in f2 (^{1}H) with 1024 data points (acquisition time 136 ms), 220 to 0 ppm in f1 (^{13}C) with 64 increments (rebuilt to 256 with NUS) and 192 scans (384 scans for NREL lignin) with a 1.5 s interscan delay. In all cases, processing used typical sine bell (90º) in f2 and squared sine-bell (90º) in f1 (first point 0.5). Volume integration of contours in HSQC plots used Maetrelabs MestReNova 14.0 (Mac version) software, and quantification of lignin linkages using methods described by previous work [219]. Spectra are displayed in absolute value mode and color coded (in Adobe Illustrator CC 2019) using literature reference standards [220].

Additionally, the total amount of phenolic compounds present in the bio-oil was estimated *via* microtiter-plated *Folin–Ciocalteu* assay [257]. In short, reactions took place in 96-well microtiter plates and in each well 150 uL of water, 10 uL of *Folin–Ciocalteu* (F-C) reagent, and 2 μL of the proper dilution of test compound were added. After this, the wells were mixed for 5 min and 30 uL of a 20% aqueous sodium carbonate solution was added to each well. The contents of the were then incubated at 45 °C for 30 min in a dry bath. The absorbance of the aliquots at 765 nm after the reaction with F-C reagent was measured against a blank using deionized water. The amount of total phenolics were quantified by correlating absorbances to standard curve generated from phenol standards at different concentrations.

Microbial Cultivation

Lactobacillus fermentum (0315-1) was provided by Dr. Chris Skory (Renewable Product Technology Research Unit, USDA-Agricultural Research Service, Peoria, IL). The other lactic acid producing bacteria used were directly sampled from commercial ethanol refiners and they were provided by Dr. Patrick Heist from Ferm-Solutions™ (Danville, KY): *Pediococcus pentosaceus, Enterococcus faecalis, Bacillus amyloliquefaciens*, *Lactobacillus fermentum*, and *Acetobacter pasteurianus*. The yeast strain used in this study was a commercially available high-performance fuel ethanol yeast strain (*Saccharomyces cerevisiae*) from Ferm-Solutions™ called Fermpro S ®. Each microbe was grown on the recommended liquid media by Agricultural Research Service Culture Collection (NRRL) with all LAB using M.R.S broth (Oxoid, CM0359) and *S. cerevisiae* using YPD media (Fisher BioReagents™, BP2469). All LAB had frozen cultures prepared by first growing each microbe in liquid culture at 180 rpm shaking speed

for 12 hrs at 37°C. These culture's cells were pelletized via centrifugation and washed with sterile media, then 500 μL of the washed cultures were added to 500 μL of sterilized 50% glycerol in a 2 mL cryovial and frozen at -80°C until use. The yeast strain was provided as an active dried product and prior to experiments the dried product was dissolved in YPD and allowed to grow overnight at 32°C and shaking speed of 180 rpm in a flask.

Antimicrobial Assay

Frozen cultures of each microbe were first revived by adding cryovial contents to liquid media and allowed to grow for 12 hr at 180 rpm shaking speed and respective incubation temperature above. Afterwards the cells were pelletized, washed, and resuspended in fresh liquid media. To test for the bio-oil and sequential extraction fractions antimicrobial properties, each microbe was cultivated in 48-well plates and the OD_{600} was monitored for 30 hrs with time points taken at 0, 6, 10, 18, and 30 hrs. These time points were previously found to represent key points of microbial growth curves in unpublished data. All wells were brought to an OD_{600} of 0.2 prior to growth, and the lignin oils were tested at 0.5, 1.0, 1.5, 2.0, 2.5, 3, and 4 mg/ml concentrations. To facilitate the solubility of the oils in media, all cultures had a final ethanol concentration of 5% (v/v). Two controls were used, one having the 5% ethanol concentration, and one having just microbes and media. To determine how the bio-oils affected microbial growth, the percent change in OD_{600} of the ethanol control during the exponential phase of growth was compared to the growth of the oils at their different concentrations. This resulted in the percent decrease in growth (degree of inhibition) for each oil at each concentration, with the formula described in Eq. 1:

$$\text{Degree of Inhibition (\%)} = \left(1 - \frac{\text{Max OD}_{600} - \text{Min OD}_{600} \text{ of Growth with Oil}}{\text{Max OD}_{600} - \text{Min OD}_{600} \text{ of Ethanol Control}}\right) * 100$$

(1)

Cell membrane integrity

A cell membrane integrity assay was performed to elucidate the mode of action of the bio-oil against *L. fermentum* (0315-1). Bacterial staining was performed using the LIVE/DEAD Bac Light Bacterial Viability Kit L7012 (Invitrogen, CA), according to manufacturer's direction, on bacterial cells incubated with or without bio-oil (4 mg/ml) for 5 hr at 37°C in a 96 well plate (clear bottom and black sides). This kit uses a SYTO9 (green) and propidium iodide (red) nuclear stains to assess cell viability and membrane damage. SYTO9 is a fluorescent dye that can penetrate cell membranes freely and once bound to nucleic acids it will fluoresce green, while propidium iodide (PI) is a red fluorescent dye that can only bind to nucleic acids in cells with damaged membranes. Since PI has a higher affinity for nucleic acids compared to SYTO9, damaged cells will fluoresce red instead of green [221]. Green fluorescence was measured at Excitation/Emission (Ex/Em) wavelengths of 485 nm/530 nm while red fluorescence was measured at 485 nm/630 nm using a SpectraMax M2 plate reader (Molecular Devices, Sunnyvale, CA).

Enzyme Inhibition Assays

To examine the effects of the lignin bio-oil on enzyme function during enzymatic saccharification, both α-amylase and glucoamylase were screened for activity while in the presence of the bio-oil at the highest concentration of 4 mg/ml. The DNS method was used to screen α-amylase activity [258], which is described below. Prior to hydrolysis reaction, the α-amylase was suspended in phosphate buffer with or without the bio-oil, at

concentration of 4 mg/ml, and allowed to interact for 30 min at ambient temperature. During the DNS assay, hydrolysis reactions took place in 2 ml Eppendorf tubes, where 0.5% (w/v) of corn starch in phosphate buffer (20 mM Sodium Phosphate with 6.7 mM Sodium Chloride, pH 6.9) was reacted with ~1 unit of α-amylase for 10 min at 60°C. Additional bio-oil was added to the reaction mixture to ensure a constant concentration of 4 mg/ml. After the reaction DNS color reagent (5.3 M potassium sodium tartrate and 96 mM 3,5-Dinitrosalicylic acid solution) was added to the tubes and boiled for 15 min. The samples were immediately placed in an ice bath until they reached room temperature and then diluted with DI water prior to spectrophotometry. The absorbance at 540 nm was measured for the samples via spectrophotometry in 96 well plates. The difference in activity were determined by comparing the amount of sugar released in the samples with standard curves of maltose.

Glucoamylase inhibition was screened by measuring glucose content after hydrolysis using HPLC, as described below. Prior to hydrolysis reaction, the glucoamylase was suspended in acetate buffer (pH 5.6) with or without the bio-oil, at concentration of 4 mg/ml, and allowed to interact for 30 min at ambient temperature. For hydrolysis, the glucoamylase with or without bio-oil was added to a 10 mg/ml maltose solution (in acetate buffer) and allowed to react for 30 min at 60°C. Additional bio-oil was added to the reaction mixture to ensure a constant concentration of 4 mg/ml. Afterwards, the reaction mixture was boiled for 15 min prior to glucose measurement. The glucose concentration released after hydrolysis was determined by HPLC (Ultimate 3000, Dionex Corporation, Sunnyvale, CA, US) equipped with a refractive index detector and using a Bio-Rad Aminex

HPX-87H column and guard assembly. The difference in activity were determined by comparing the amount of sugar released in the samples with standard curves of glucose.

Ethanol Fermentation

To test the antimicrobial properties of the PAA bio-oil in a fermentation system contaminated with LAB, the bio-oil was used at 4 mg/ml concentration against just yeast and yeast contaminated with *L. fermentum* (0315-1). Only *L. fermentum* (0315-1) was tested in these model "stuck" fermentation experiments as it has already been shown to be a prominent strain causing contamination in the fuel ethanol industry [239, 253]. We tested different inoculation rates of yeast and *L. fermentum* at 1:100, 1:10, and 1:1 yeast to LAB ratios. Stock cultures of yeast and L. *fermentum* were prepared as previously described above. After 24 hr incubation the microbial cells were pelletized via centrifugation and resuspended in phosphate buffered saline. Using one OD_{600} of yeast corresponds to 6×10^7 CFU/mL and one OD_{600} of *L. fermentum* is 1×10^8 CFU/mL, yeast was inoculated at starting concentration of 10^6 CFU/ml and *L. fermentum* was inoculated at either 10^6, 10^7, or 10^8 CFU/ml.

Simultaneous saccharification and fermentation (SSF) was performed on raw corn starch (17% w/w) in sterile YP medium (10 g of yeast extract and 20 g of peptone per liter of water). First, 1 ml α-amylase (Sigma-Aldrich, A8220) was added to 1 L of starch solution and brought to 85 °C and held for 15 min. After enzymatic liquefaction, the starch solution was autoclaved at 121 °C for 15 min. The mixture was cooled to 85 °C, an additional 4 ml of α-amylase was added, and then it was placed in a water bath at 85 °C for 1 hr with intermittent stirring. The mixture was then brought down to 32 °C and

glucoamylase was added to yield a concentration of 0.05% (v/v) glucoamylase (Sigma-Aldrich, A7095) right before inoculation and fermentation.

SSF was performed for 72 hrs at 32 °C in 50 ml serum bottles capped with a rubber septum that had a 20-gage needle inserted for gas release. 30 ml of the starch solution was added to the serum bottles and inoculated with 0.15 ml yeast and depending on the treatment 0.5ml of *L. fermentum* and/or 0.15 ml of bio-oil dissolved in ethanol. Treatments without *L. fermentum* had 0.5 ml of PBS added to serve as negative control. SSF treatments were performed in duplicate and samples were withdrawn at 0, 12, 24, 36, 48, 60, and 72 hr time points. The samples were analyzed using the same HPLC methods described above to monitor ethanol, lactic acid, acetic acid, and glucose concentrations. Additionally, lactobacillus density was enumerated at each time point using colony counting on MRS agar media supplemented with cycloheximide (10 µg/ml) to selectively inhibit the growth of yeast. When the bio-oil was added, the fermentation solution had a 3.75 g/L starting concentration of ethanol, thus an additional control was used with the same starting concentration of ethanol for comparisons. The final reported ethanol production values for all time points were subtracted by 3.75 g/L for the bio-oil treatments and ethanol controls to compensate for the addition of ethanol to the system.

Results and Discussion

Lignin depolymerization

After the depolymerization reaction of alkali-extracted corn stover lignin (AEL) with peracetic acid, the reaction mixture was diluted with water and the resulting bio-oil was extracted with ethyl acetate. The ethyl acetate extracted bio-oil, which is the fraction

used for all further experimentation, was found to be 36.1 ± 0.4 wt% of the starting lignin, whereas the remaining water soluble and undissolved solids were found to be 23.6 ± 0.9 and 40.3 ± 1.4 wt%, respectively (**Table 4.1**). While previous reports using similar reaction and extraction conditions with diluted acid corn stover lignin and kraft lignin found bio-oil yields of 58 and 16-45 wt% of starting lignin [63, 259, 260], respectively, the difference in lignin origin and purity can greatly affect depolymerization bio-oil yields. For example, Ma et al. [63] did not fully characterize their lignin source and without further purification there are likely large amounts of carbohydrates still present, which would inflate the conversion. The purity of the AEL used in this study (after utilizing precipitation methods for purification) was found to be $95.11 \pm 0.18\%$ with $3.62 \pm 0.16\%$ glucan and $1.27 \pm 0.03\%$ xylan, so the comparison of yields may not be truly applicable. Furthermore, Ma et al. [63] also found that during their depolymerization reactions the starting lignin was completely dissolved. This was not true during the reactions seen here and could indicate that the AEL is more resistant to oxidative depolymerization than the other lignin sources.

Table 4.1: Mass balance of lignin depolymerization into bio-oil as a percentage of starting lignin weight.

Fraction	Average Wt%
Ethyl Acetate Extracted Bio-oil	36.1 ± 0.4
Remaining Water Soluble	23.6 ± 0.9
Solids	40.3 ± 1.4

Bio-oil characterization

The weight-average (Mw) and number-average (Mn) molecular weight, as well as the polydispersity index (PDI) of unreacted AEL were compared with the PAA depolymerized bio-oil, as shown in **Table 4.2**. After treatment with PAA the weight average molecular weight (Mw) of AEL was reduced from approximately 4095 to 2277 Da (g/mol) in the bio-oil. The PDI was also reduced in the bio-oil, showing there was a significant reduction in the molecular weight and a narrower distribution of molecular weight products compared to untreated AEL. This is also evident from the GPC chromatograms where MWD curves of the bio-oil shifted to the right (i.e., lower MW) compared to that of the untreated AEL (**Figure 4.1**). Previous studies depolymerizing kraft lignin with PAA at varying concentrations of PAA and temperature found that untreated kraft lignin was reduced from 2813 Da to bio-oils with Mw ranging from ~750-1500 Da [259, 260]. Even though their data also show an approximate two-fold reduction in Mw after depolymerization with PAA, the bio-oil created here is more similar in Mw to that of unreacted kraft lignin than the previously reported bio-oils. Thus, the GPC results provide a general trend on the size distribution of compounds in the PAA derived bio-oil and indicate that some depolymerization did occur. In order to identify the specific compounds formed after oxidative depolymerization, GC/MS analysis and the *Folin–Ciocalteu* assay were performed.

Table 4.2: The molecular weight distribution of raw corn Stover lignin and PAA bio-oil.

Sample	Mw (g/mol)	Mn (g/mol)	Polydispersity index (PDI)
Corn Stover Lignin (AEL)	4095	1112	3.6
PAA Bio-oil	2277	785	2.9
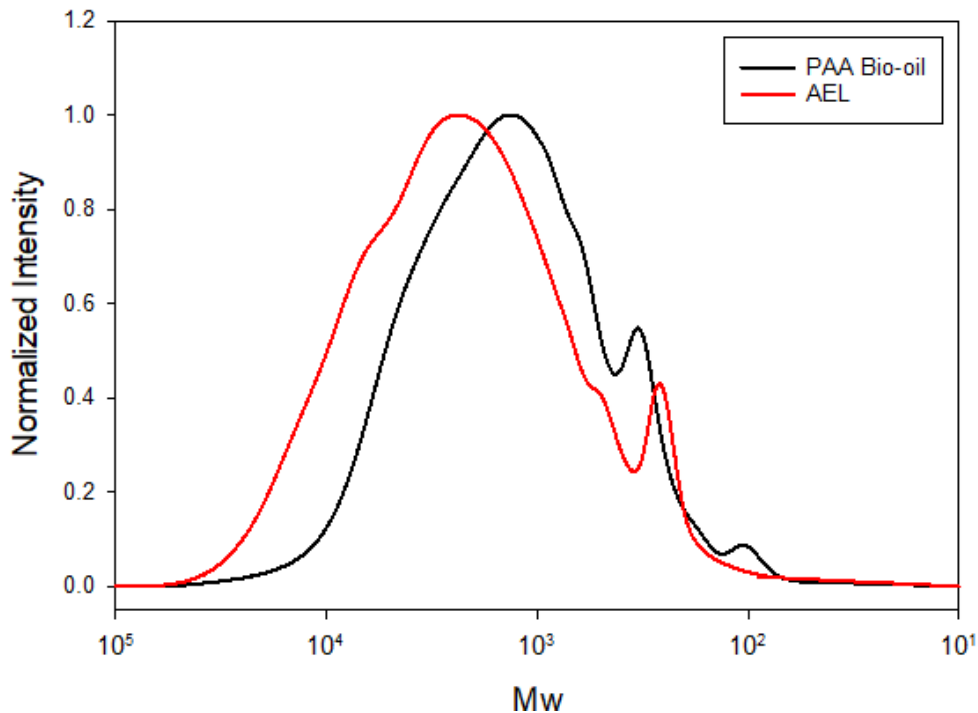
Figure 4.1: GPC chromatogram spectra of purified AEL and PAA derived bio-oil.

GC/MS analysis found 10 lignin-derived phenolic monomers in the bio-oil that only accounted for 1.77 wt% of the bio-oil (**Table 4.3**). Hydroxylated phenolics (i.e. hydroquinone) represented 46.9 wt% of the total monomers detected. In terms of the depolymerization reactants acetic acid and peracetic acid, GC/MS analysis found that the bio-oil was comprised of 2.2% acetic acid, while no residual peracetic acid was found using peracetic acid test strips (MilliporeSigma^{TM}, MQuant^{TM}) that can detect 5-50 ppm of peracetic acid. The concentration of the bio-oil used for the test strips was around 500 mg/ml, so the amount of peracetic acid was below the detection limit of 5 ppm at this bio-oil concentration. This would mean the bio-oil contains less than 0.001% peracetic acid. Considering the final bio-oil concentrations in the antimicrobial tests were less than 4 mg/ml, the PAA concentration in the antimicrobial tests would be far below the 150-200

ppm inhibitory loading of PAA that is found to inhibit microorganisms [261, 262]. Neither Ma et al [63] or Park et al [259] found any residual peracetic acid in the resulting lignin depolymerization products, which makes sense as peracetic acid will quickly degrade to hydrogen peroxide and acetic acid during the dilution step with water and subsequent drying. This confirms that the oxidizers used during the depolymerization reaction were removed and should not a play a role in the antimicrobial properties of the bio-oil.

While the monomeric phenolic yields seen here are very low, the *Folin–Ciocalteu* assay revealed that the bio-oil had a total phenolic content of 22.84±0.30% in terms of gallic acid equivalents. This is higher than the total monomer phenolic content found in the GC/MS results, but the *Folin–Ciocalteu* assay does not only measure monomeric phenolics, and thus larger oligomers can also be represented in this value [263]. These results indicate that the degradation compounds from AEL were primarily present as larger oligomers. Despite significant degradation occurring as evidenced by GPC results, PAA oxidation was limited in the production of monomers. Ma et al. [63] also found a total phenolic yield of 22% using the *Folin–Ciocalteu,* but they represented this value as total monomer yields and utilized total ion chromatogram (TIC) peak area instead of flame ionization detector (FID) for individual monomer quantification. On the other hand, Park et al. [259] used FID for monomer quantification and found less than 0.08% of lignin monomers after treating kraft lignin with PAA, which is similar to the results of this study. Nonetheless, the data clearly illustrates that the bio-oil created here is comprised primarily of large molecular weight oligomers that are unidentifiable in GC/MS analysis.

Table 4.3: GC/MS identifiable monomers in lignin bio-oil, with yields represented as mg/ml and wt% of total oil weight.

Compound	Yield (mg/ml)	Yield (wt %)
Hydroquinone	0.69	0.83
p-Coumaric acid	0.30	0.36
2,6-Dimethoxyhydroquinone	0.09	0.11
Syringic acid	0.09	0.11
Phloroglucinol	0.08	0.10
4-Hydroxybenzaldehyde	0.07	0.09
4-Hydroxyacetophenone	0.06	0.07
Ferulic acid	0.05	0.06
3-Ethylphenol	0.01	0.02
2-Hydroxybenzyl alcohol	0.01	0.01
Totals	**1.46**	**1.77**

^{1}H-^{13}C-HSQC NMR was also performed on the starting lignin and PAA depolymerized bio-oil (**Figure 4.2**). Notably, the AEL lignin showed only β-O-4 linkages in the linkage region. Additionally, the starting lignin had many conjugate esters as evident from the presence of pCA (p-coumaric acid ester) and FA (ferulic acid ester) in the HSQC. After treating the lignin with PAA, much of the conjugate esters and G-type structures remained relatively unchanged. This is unsurprising given that acid catalyzed hydrolysis of esters is kinetically slower than the analogous base catalyzed reaction, and the potential for rearrangement of these esters on the lignin polymer under the current reaction conditions. Our GC/MS results support the low amounts of bond cleavage seen in the bio-oil and further indicate the lack of depolymerization into monomeric fragments occurring after PAA treatment. Moreover, HSQC of the PAA lignin revealed the complete loss of S-type structures from the bio-oil. We contend that this may be due the increased lability

of the S-lignin under acidic oxidizing conditions from the electron donation of the methoxy groups to the β-O-4 Cα-OH, however more studies are still needed to confirm this hypothesis [63].
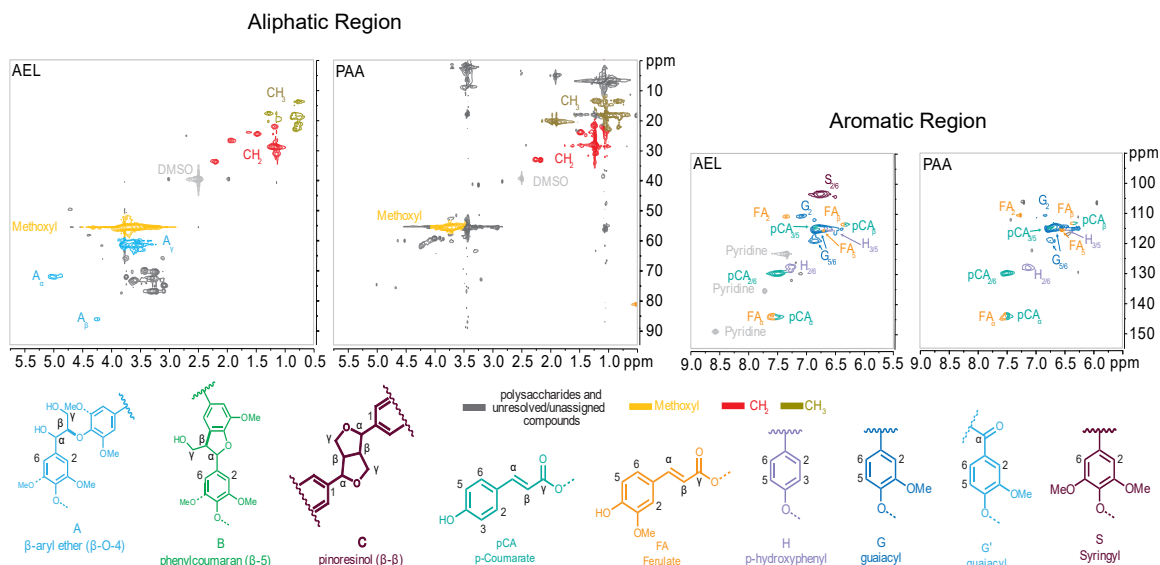
Figure 4.2: ^{1}H-^{13}C HSQC NMR of AEL and raw bio-oil derived from PAA oxidation. The structures of lignin compositional units and side-chain linkages were coded with colors corresponding to the cross peaks in the spectra.

Antimicrobial Assay

The PAA derived lignin bio-oil was tested against several LAB sampled directly from commercial facilities and a commercially available high-performance fuel ethanol yeast strain (Fermpro S ®) for antimicrobial properties by measuring growth differences utilizing spectrophotometry. Since the bio-oil is hydrophobic and becomes tar-like after drying, ethanol was used as a solvent. Consequently, an ethanol control was used in all further analyses to ensure ethanol's growth effects were accounted for. The results in **Table 4.4** illustrate that the bio-oil as no growth effects against yeast at any of the concentrations tested, but that the LAB showed significant growth reduction at all tested concentrations.

The bio-oil was more inhibitory to both *L. fermentum* strains tested, which experienced a growth reduction of greater than 60% at bio-oil concentrations ranging from 1-2.5 mg/ml and then over 80% at 3 mg/ml. This is important as the *L. fermentum* (0315-1) strain used here is found to be one of the most prolific strains causing stuck fermentation in the fuel ethanol industry [239, 253]. Therefore, the bio-oil was effective at reducing LAB growth while showing no effects on yeast growth. This provides evidence for a selective mode of action that targets Gram-positive bacteria compared to eukaryotic yeast cells. As stated previously, the bio-oil contained 2.2% acetic acid and less than 0.001% peracetic acid, if at all. At the highest bio-oil concentration of 4 mg/ml this would represent a maximum of 0.088 mg/ml of acetic acid and 0.00004 mg/ml of peracetic acid, and at these concentrations neither would have an impact on either yeast or LAB growth [261, 264]. Our data supports this assertion, as yeast would be more susceptible to acetic acid compared to LAB [265], but the data clearly show the bio-oil has no inhibition to yeast and only the LAB. Furthermore, since the bio-oil was found to contain mostly unidentifiable lignin oligomers, it is these oligomers that are responsible for the highly selective antimicrobial activity seen in this bio-oil.

Table 4.4: Percent inhibition of PAA bio-oil at varying concentrations. Letters indicate differences at 95% confidence across all bio-oil concentrations for each organism, where values are mean±SE (n=3), using one-way ANOVAs with Tukey's pairwise comparisons or a T-test.

	Percent Inhibition						
	PAA Bio-Oil Concentration (mg/ml)						
Organism	0.5	1	1.5	2	2.5	3	4
S. cerevisiae (Fermpro)	-6.88 ±1.56 [a]	4.96 ±2.03 [a]	1.73 ±1.44 [a]	-6.93 ±1.16 [a]	1.01 ±6.31 [a]	-6.83 ±1.78 [a]	-7.19 ±2.28 [a]
L. Fermentum (0315-1)	23.43 ±1.89 [a]	43.58 ±1.11 [b]	58.72 ±3.98 [bc]	61.06 ±4.02 [c]	59.74 ±5.06 [bc]	66.82 ±3.66 [c]	87.45 ±2.03 [d]
L. Fermentum	12.24 ±1.50 [a]	75.17 ±1.02 [bc]	69.57 ±4.46 [b]	74.97 ±0.76 [bc]	72.20 ±2.85 [bc]	91.25 ±0.63 [d]	83.96 ±3.61 [dc]
P. pentosaceus	N/A	N/A	N/A	74.99 ±5.07 [a]	N/A	N/A	75.18 ±0.00 [a]
E. faecalis	N/A	N/A	N/A	51.71 ±15.47 [a]	N/A	N/A	79.42 ±10.35 [a]
B. amyloliquefaciens	N/A	N/A	N/A	28.57 ±13.93 [a]	N/A	N/A	65.55 ±0.10 [a]
A. pasteurianus	N/A	N/A	N/A	42.48 ±1.86 [a]	N/A	N/A	56.40 ±12.41 [a]

In order to infer the mode of action of the PAA derived lignin bio-oil, *L. Fermentum* (0315-1) was grown in the presence of the bio-oil and assessed for potential membrane damage by staining with SYTO9 and propidium iodide (PI) nuclear dyes. SYTO9 is a green fluorescent dye, and PI is a red fluorescence dye that both bind to nuclear material in the cell. However, while SYTO9 can penetrate cells freely, PI can only penetrate damaged membranes and due to its higher affinity for nucleic acids it can displace the weaker bound SYTO9 dye causing the damaged cell to show a strong red fluorescence instead of green [221]. Therefore, cells that fluoresce green (SYTO9) represent live cells without membrane damage and cells that fluoresce red (PI) can be considered membrane damaged or dead. **Figure 4.3** shows the ratio of SYTO9/PI fluorescence, representing the ratio of live cells

to membrane damaged/dead cells, after *L. Fermentum* (0315-1) was incubated with the highest tested concentration of bio-oil (4 mg/ml) for 5 h at 37ºC.

The data show a significant decrease ($p < 0.05$) in SYTO9/PI ratio when comparing the controls with cells in the presence of the bio-oil, where the ratios decreased from ~4.3 and 3.95 (control and ethanol control, respectively) to 0.7 when treated with the bio-oil. By significantly increasing the proportion of cells that fluoresce red when treated with bio-oil, it is assumed that these treated cells are PI-permeable membranes primarily due to death or damaged membranes [221]. Lignin derivatives have been thought to directly cause cell membrane damage or have ionophoric activity that ultimately results in cell lysis and death [122, 243]. However, since PI stained cells may not only indicate membrane damage, the mode of action may still be molecular in nature, which could affect protein synthesis or influence expression of genes, also resulting in cellular death [266]. Additionally, this data coupled with the percent inhibition data suggests that the bio-oil displays more bactericidal activity than bacteriostatic activity [247], due to the direct cell death compared to just inhibiting cellular growth. In summary, the bio-oil is effective at selectively inhibiting a variety of LAB due to cell death without inhibiting yeast, which supports the use of this bio-oil as an alternative to control bacterial contamination in fuel ethanol fermentation.
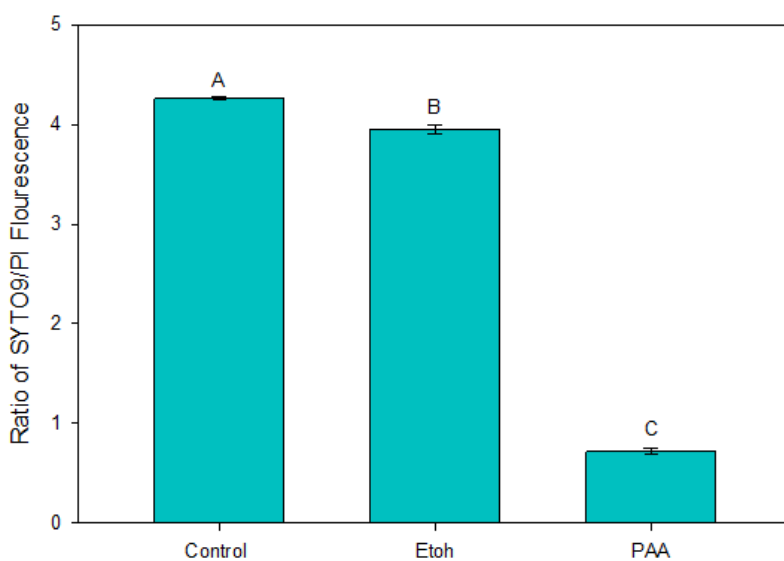
Figure 4.3: SYTO9/PI fluorescence ratios of *L. Fermentum* (0315-1) treated with or without bio-oil at a concentration of 4 mg/ml after incubating for 5 hr at 37 °C. These ratios indicate the ratio of live/dead or undamaged/membrane-damaged cells. In the **Figure**, Etoh is the control with ethanol added and PAA is the treatment with the bio-oil (4 mg/ml). Letters on the bars indicate differences at 95% confidence where values are mean±SE (n=3), using students T-tests.

Model Fermentations

Prior to conducting ethanol fermentation experiments, both α-amylase and glucoamylase were screened for inhibition when in the presence of the bio-oil at the highest concentration tested of 4 mg/ml. **Figure 4.4** shows that α-amylase had a significant increase in activity, as measured by an increase in the amount of maltose released from hydrolysis of corn starch, while glucoamylase had no significant difference in the amount glucose released from hydrolysis of maltose when in the presence of bio-oil compared to the control. Thus, enzymatic saccharification during corn ethanol fermentation will not be impacted by the bio-oil and may actually be benefited by the slight increase in α-amylase activity.
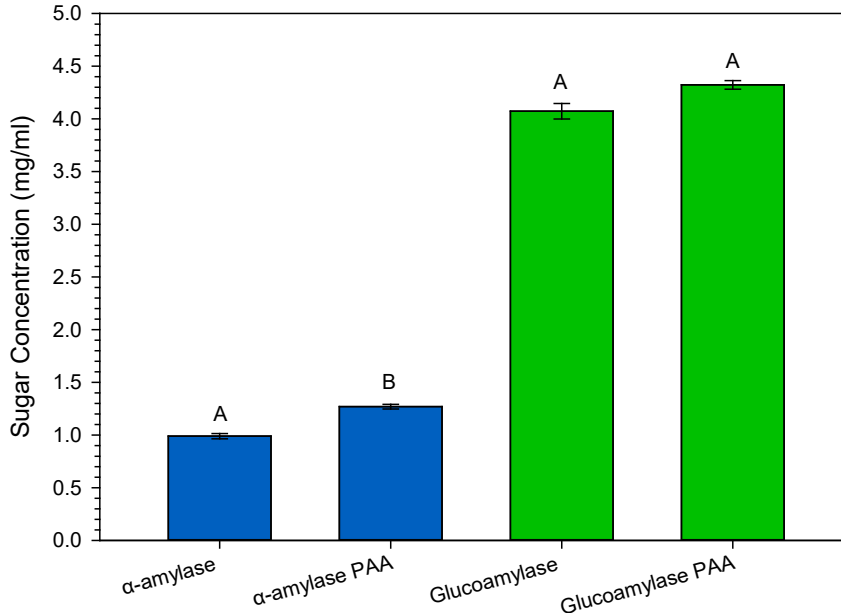
Figure 4.4: Sugar concentrations after enzymatic hydrolysis with or without the presence of PAA bio-oil at a concentration of 4 mg/ml. α-amylase bars (blue) indicate the amount of maltose released after 10 min of starch hydrolysis, while the glucoamylase bars (green) indicate amount of glucose released after 30 min of maltose hydrolysis. For the bio-oil treatments, the enzymes were pre-incubated for 30 min in the presence of bio-oil, and the same concentration of bio-oil was maintained during hydrolysis reactions. Letters on the bars indicate differences at 95% confidence where values are mean±SE (n=3), using students T-test.

In this study, we tested SSF of corn starch challenged with a previously reported bacterial strain that causes "stuck" fermentation in fuel ethanol facilities (*L. Fermentum*, 0315-1) [253, 267]. This strain is also known to be virginiamycin-resistant with a MIC value of 16 μg/ml compared to ≤2 μg/ml for susceptible strains [267], this is important as our study aims to utilize lignin bio-oil as an alternative antibiotic. Based on previous surveys of bacterial contaminants in fuel ethanol facilities that found bacterial loads can reach 10^8 CFU/ml [252], we challenged our model fermentations with yeast to lactobacillus ratios of 1:1, 1:10, and 1:100, which resulted in initial bacterial loads of 10^6, 10^7, and 10^8, respectively. Furthermore, we utilized a bio-oil concentration of 4 mg/ml for all

experiments as it was the highest tested concentration in our antimicrobial experiments and was the highest concentration that could be achieved due to bio-oil solubility in ethanol.

Figures 4.5 and **4.6** show the differences in ethanol production, glucose consumption, and lactic/acetic acid production for the uncontaminated and contaminated SSF runs, respectively. The uncontaminated fermentations showed no significant difference in ethanol production, glucose consumption, or acetic acid production after 72 hrs of fermentation for the control, ethanol control, and bio-oil treatment. Available glucose was mostly consumed after 36 hrs (**Figure 4.5**), which is also marked by no significant increase in ethanol production. Since the fermentation broth utilized isolated corn starch instead of a traditional corn mash, there was also a large amount of starch solids leftover in the fermentation broth and the ethanol production was only 76% of the theoretical yield (96.5 g/L compared to 75 g/L). Despite this, the data clearly indicate that the addition of the bio-oil had no significant effect on corn starch fermentation and yeast metabolism for the uncontaminated controls, supporting our previous antimicrobial results.
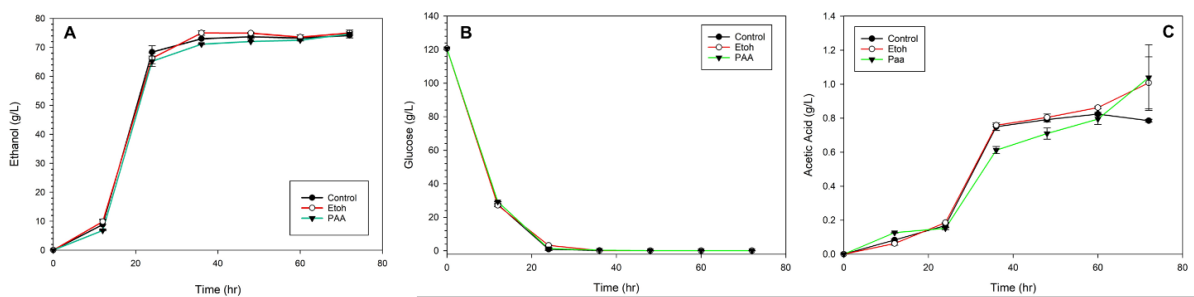
Figure 4.5: Ethanol (A), glucose (B), and acetic acid (C) concentrations during fermentation without contamination over time. In each graph Etoh is the control with ethanol added and PAA is the treatment with the bio-oil.

Figure 4.6 shows the fermentation products for the contaminated samples at yeast to LAB ratios of 1:1 (A,D,G), 1:10 (B,E,H), and 1:100 (C,F,I). In terms of ethanol production, after 72 hrs of fermentation, the 1:100 yeast to LAB ratio saw the greatest reduction in ethanol at 17% compared to the uncontaminated controls (**Figure 4.6 C**). The 1:1 ratio had no significant reduction in ethanol, and the 1:10 ratio had an 11% reduction in ethanol production. Conversely, Rich et al. [253] used yeast to LAB ratio of 1:6 and found an ethanol reduction of 23%, while Bischoff et al. [267] found a 17% reduction in ethanol at a ratio of 1:10, where both studies used the same lactobacillus strain applied here. Thus, the effect of lactobacilli contamination observed in the present study is less pronounced than that reported previously. These differences may be attributed to use of corn starch instead of corn mash and/or the difference in our yeast strain, which could be more vigorous, causing the lactobacillus to be a less potent antagonist. Moreover, the bio-oil treatment did not significantly improve the ethanol yields for either the 1:1 or 1:10 yeast to LAB ratios (**Figure 4.6 A** and **B**), but it did significantly improve ethanol yields by 8% for the 1:100 ratio (**Figure 4.6 C**). This increase was even more drastic at the 24 hr time point, where the bio-oil treatment had 212% increase in ethanol produced compared to both the contaminated control and ethanol control (**Figure 4.6 C**). Similarly, the 1:1 and 1:10 contamination ratios also saw an increase in ethanol production at the 24 hr time point, but this did not impact total ethanol production like in the 1:100 ratio. Therefore, as the amount of bacterial contamination increased and at earlier time points the bio-oil treatment did have a greater beneficial effect on ethanol production compared to untreated contaminated samples.
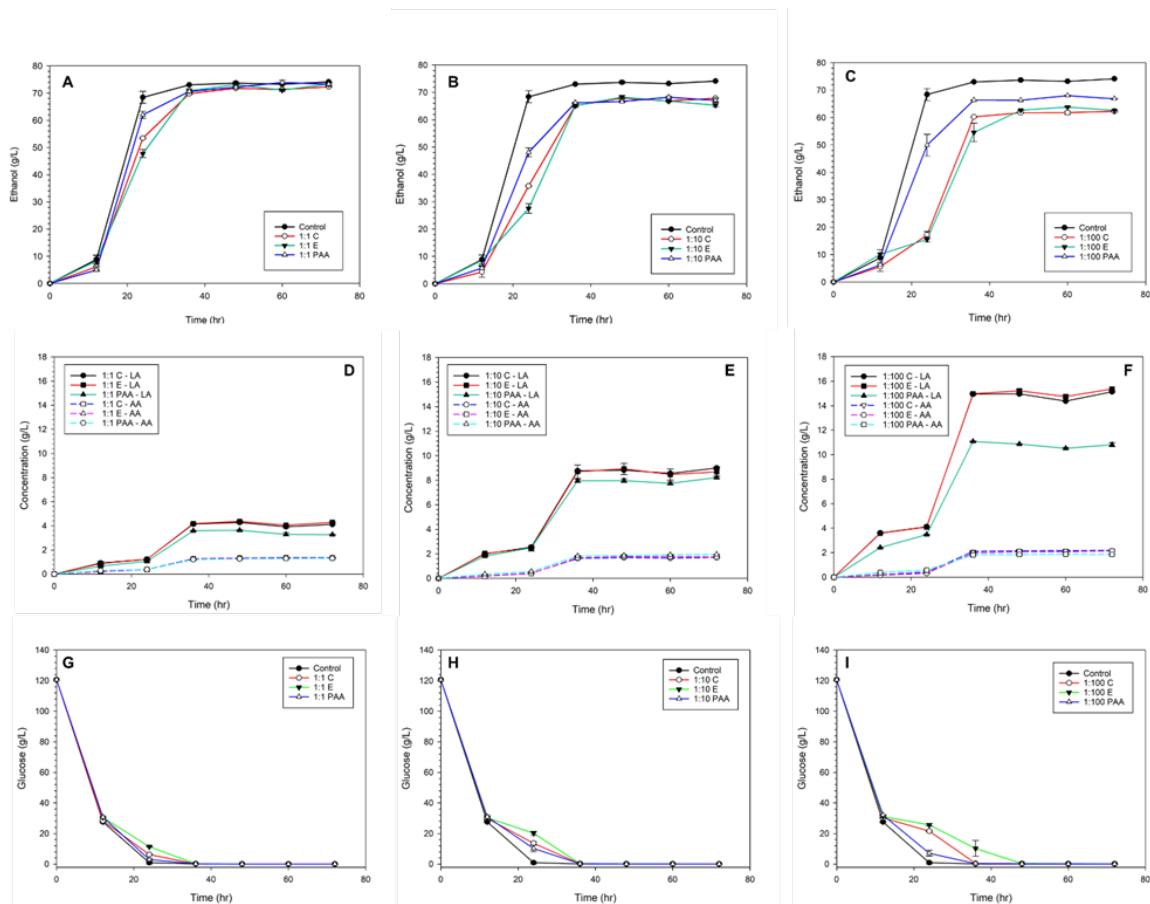
Figure 4.6: Ethanol (A-C), lactic/acetic acid (D-F), and glucose (G-I) concentrations during fermentation contaminated with *L. Fermentum* (0315-1) overtime for 72 hrs. The inoculation rates for the LAB were at yeast:LAB ratios of 1:1 (A,D,G), 1:10 (B,E,H), and 1:100 (C,F,I). For D-F the solid lines indicate lactic acid (LA) and the dotted lines indicate acetic acid (AA). The uncontaminated control from **Figure 4.4** is provided for the ethanol and glucose concentrations for comparison and is labeled "Control". In each graph, C (i.e. 1:100 C) represents the control, E is control with ethanol added, and PAA is the treatment with the bio-oil (4 mg/ml).

While the lower bacterial contamination ratios of 1:1 and 1:10 did not see improvements from the bio-oil treatment for ethanol production, there was still a 10% reduction in lactic acid production, indicating there was an effect on the lactobacillus growth/metabolism (**Figure 4.6 D and E**). This was even more pronounced in the 1:100 contamination ratios where lactic acid was reduced by 33% when treated with bio-oil (**Figure 4.6 F**). To this end, we also monitored the lactobacillus population (CFU/ml) over

131

the 72 hr fermentation period for the 1:100 ratio as it had the most detrimental effect on ethanol yields. The results in **Figure 4.7** indicate that the lactobacillus population drastically decreased overtime for the bio-oil treatment and when compared to the control and ethanol control there was an almost 100-fold reduction in the lactobacillus population. However, even with this reduction there was still a viable lactobacillus population at 3.6×10^6 CFU/ml after 72 hrs of fermentation when treated with the bio-oil, which ultimately reduced final ethanol yields. The presence of a viable lactobacillus population in the bio-treatments makes sense, as our initial antimicrobial experiments showed at 4 mg/ml there was only a 90% reduction in growth (i.e. 10% of the population was still viable). Despite the fact that the bio-oil did not completely inhibit LAB growth during SSF, the improvement in ethanol production based on the contamination controls clearly illustrates the effectiveness of using this lignin bio-oil as an antibiotic replacement to control antibiotic-resistant LAB strains.

It must be noted that during the fermentation experiments with the bio-oil treatments, the bio-oil was seen mostly in a solid phase when added to the fermentation broth. Over time as the ethanol concentration inevitably increased, more of the bio-oil was seen to go into solution (as noted by a color change); thus, future work needs to improve the bio-oil's initial solubility in order to increase its efficacy as an alternative antibiotic. The issues with concentrating the bio-oil and limiting the addition of solvent into the fermentation system must also be re-examined and improved. Furthermore, since traditional antibiotics such as virginiamycin are used at concentrations less than 2 ppm [267], the use of the bio-oil here at a maximum of 4 mg/ml is not directly comparable. As

L. Fermentum, 0315-1 is also virginiamycin-resistant the lack of a commercial antibiotic control is warranted.
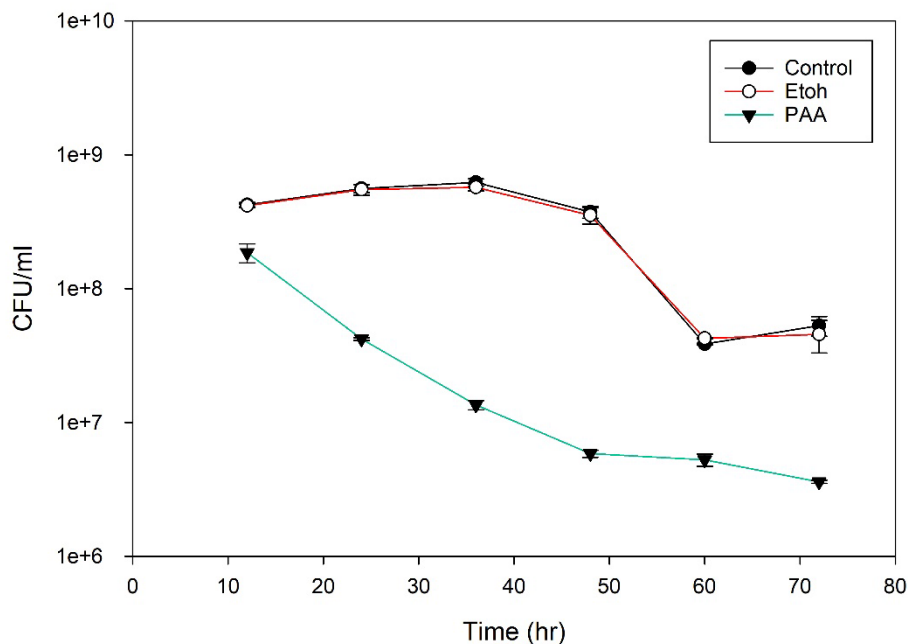
Figure 4.7: Colony forming units (CFU) per ml of *L. Fermentum* (0315-1) during contaminated fermentation at a yeast:LAB ratio of 1:100 overtime for 72 hrs. Where Etoh is the control with ethanol added and PAA is the treatment with the bio-oil (4 mg/ml).

Conclusions

In the present study, we have demonstrated that depolymerization of AEL by PAA oxidation produces mostly unidentifiable lignin oligomers with highly selective antimicrobial properties. Even though the resulting bio-oil contained less than 1.77 wt% of identifiable monomeric phenolic compounds, it demonstrated no inhibition against yeast and up to 90% inhibition of commercially sampled LAB at 4 mg/ml. The highly selective antimicrobial properties of the bio-oil are attributed to an ionophoric or membrane damaging mode of action that results in cell death, based on fluorescent staining. Using the

bio-oil (4 mg/ml) as an alternative antibiotic treatment during SSF of raw corn starch showed an increase in ethanol production as bacterial contamination increased. At the highest contamination ratio of 1:100 yeast to lactobacillus, the bio-oil increased ethanol production by 8% compared to the contamination control. While the bio-oil did not completely inhibit lactobacillus growth, which still resulted in net losses of ethanol production (9%) compared to the uncontaminated control, the ability of the bio-oil to improve ethanol yields clearly show its efficacy as an alternative antibiotic. Further research must improve the bio-oils solubility during fermentation to increase its antimicrobial action and resulting beneficial effects on ethanol production. Therefore, the results obtained from this study offer a new application in lignin valorization and a better understanding of lignin-based bio-oil's antimicrobial properties/potential.

Acknowledgements

CHAPTER 5. PREDICTING THE ANTIMICROBIAL PROPERTIES OF LIGNIN

DERIVATIVES USING TRADITIONAL AND MACHINE LEARNING BASED

QSAR MODELS

Abstract

Lignin is a waste stream from biorefineries, and its polyphenolic structure can be depolymerized into small molecules that are inherently antimicrobial. However, during depolymerization the bio-oils created can be a complex mixture of mainly unidentifiable compounds. This makes the process of identifying active compounds or extracting them expensive/time consuming using conventional methods. Therefore, methods need to be developed that can predict the antimicrobial potential of lignin derivatives so that the search for lignin depolymerization products with enhanced antimicrobial properties can be expedited. In this context, the aim of this study was to develop and compare QSAR models that can predict the antimicrobial properties of lignin derivatives against representative Gram-positive (*Bacillus subtilis*) and negative bacteria (*Escherichia coli*). The compounds used to construct the models were selected from a large public access database (ChEMBL) that were non-specific to lignin, a database created from a metanalysis of available lignin compounds with activity measurements (minimum inhibitory concentration, MIC), and an experimentally derived dataset of lignin monomers and dimers by measuring activity as relative Bacterial Load Difference (BLD). The ChEMBL dataset's QSAR models were developed using different machine learning algorithms (support vector machine, random forest, k-nearest neighbor, decision tree, and neural networks) and were found to underpredict the antimicrobial activity of actual lignin compounds. The metanalysis data used to validate the ChEMBL dataset's QSAR models for *B. subtills* and *E. coli* were used to build their own more traditional QSARs using ordinary least square (OLS) regressions. An accurate QSAR model for *E. coli* was not found, but a satisfactory model was obtained for the *B. subtilis* metanalysis dataset. MOE-type descriptors and the number of aliphatic

carboxylic acid groups were the descriptors that showed strong correlations to the MIC values (R^2 of 0.759). Specifically, as the number of aliphatic carboxylic acid groups increased, the model predicted an increase in antimicrobial activity (i.e. lower MIC). Comparatively, an additional dataset was experimentally derived by screening 25 lignin monomers and three dimers against *B. subtilis* by measuring BLD. The experimentally based QSAR found that MOE-type descriptors and the number of aromatic hydroxyl groups were better predictors of BLD (R^2 of 0.831). Thus, the smaller dataset's models show how the variability in antimicrobial measurements and the specific compounds used will impact the predictive nature of the resulting QSARs. This study is the first attempt to predict the antimicrobial properties of lignin derivatives, and the results provide insights into the type of descriptors that correlate to an increase in the antimicrobial properties of lignin.

Keywords: Lignin, Quantitative Structure−Activity Relationship, Machine Learning, Open-Source Databases

Introduction

Due to the overuse of antibiotics in our society, there has been a steady rise in highly antimicrobial-resistant bacteria in the last decade. This has created a renewed interest in natural compounds for antimicrobial discovery amongst the scientific community [81, 82]. Plant-based phenolics have a wide spectrum of antimicrobial activity and a variety of ring scaffolds and low human toxicity potential that makes them a promising source of potential antimicrobial replacements [82, 83].

To this end, lignin is one of the most abundant naturally occurring sources of phenolic polymers on earth and is currently considered a major waste product in the paper and pulp industries and industrial lignocellulosic biorefineries [44]. Lignin is already known to have antimicrobial properties against different microorganisms, which is due to the phenolic subunits that comprise lignin's polyphenolic structure [122, 170]. Lignin's antimicrobial properties are dictated by the source of the lignin, its extraction methods and chemical structure (i.e. monomers, oligomers and functional groups) [12, 122]. In general, it is believed that lignin phenolics have a mode of action that centers around their ability to increase the ion permeability of cell membranes or cause direct membrane damage resulting in cell lysis [99, 119, 124, 207]. However, lignin's inhomogeneity and complex structure greatly reduces its capacity to be used in industrial and commercial sectors.

For example, while a variety of technical lignins (i.e. Kraft lignin and organosolv lignin) with large undefined structures have had notable antimicrobial properties, there remain inconsistencies in different batches, across different lignin sources, and extraction methods [12, 122]. Conversely, when lignin is depolymerized into smaller more defined structures, these smaller oligomers and phenolic monomers have shown increased

antimicrobial activity and higher specificity [123]. Thus, to increase the effectiveness and selectivity of lignin's antimicrobial properties, it is necessary to depolymerize the polyphenolic structure of technical lignins into smaller units.

There are a plethora of lignin depolymerization techniques to produce small molecule monolignols from lignin including: pyrolysis, acid/base/metal catalyzed hydrolysis, hydrogenolysis and oxidation [9, 45, 46]. Depending on the lignin source each depolymerization method will produce a variety of different phenolic compounds with potential antimicrobial properties. After lignin depolymerization reactions, the resulting product is usually a bio-oil that is composed of a complex cocktail of monomers (<50% w/w) and larger oligomers. While lignin bio-oils have shown promising antimicrobial properties for a variety of industrial applications, as seen in Chapters 3 and 4 of this dissertation, there remain questions as to what individual compounds are responsible for their diverse antimicrobial properties.

In practice, when antimicrobials are developed, they are usually composed of a single component or compound. When considering the use of lignin-based bio-oils, it would be incredibly difficult to attribute a single compound to its antimicrobial properties, as it is too complex of a mixture. Separation technologies like chromatography, membrane filtration, or liquid-to-liquid extraction could be utilized to extract a specific highly active phenolic compound from the bio-oil, but the compound would first need to be identified, and those separation technologies can be costly or produce harmful byproducts [68-73]. While future separation techniques can be developed at lower cost and hazards, if an effective depolymerization method can produce a consistently highly active mixture, then separation would not be necessary. Therefore, methods need to be developed that can

predict the antimicrobial potential of lignin derivatives so that the search for lignin depolymerization products with enhanced antimicrobial properties can be expedited.

Quantitative structure−activity relationship (QSAR) models are an indispensable tool in drug design and discovery. They work by finding relationships between the variations in calculated molecular descriptors (properties) or fingerprints (functional groups) with the biological activity of a group of compounds so that biological activity of new chemical entities can be assessed more quickly [154]. Traditional QSAR modeling utilizes experimentally derived datasets with a limited number of compounds (<50) and selected descriptors for developing a predictive regression type model, such as multiple linear regressions (MLR) [160, 268]. While this increases the specificity of the model to predict the identified target compounds, it simultaneously limits the model's ability to predict the activity of new compounds with a wider variety of structures. One of the ways to circumvent this issue would be to increase dataset size and compound variability. However, due to the lengthy experimental procedures used to measure antimicrobial activity, and the fact that many lignin oligomers after depolymerization are currently unidentifiable, it would be difficult to drastically increase the number of compounds tested in an efficient manner. Given the recent advances in machine learning and the increase in the amount of chemical and biological activity data available in the public domain in recent years [269], QSAR models that can explore a vaster chemical space (thousands of compounds) can now be more widely applied [270].

In this context, the aim of this study was to develop and compare QSAR models that can predict the antimicrobial properties of lignin derivatives against representative Gram-positive (*Bacillus subtilis*) and negative bacteria (*Escherichia coli*). The compounds

used to construct the models were selected from a large public access database that was non-specific to lignin, a database created from a metanalysis of available lignin compounds with activity measurements, and an experimentally derived dataset of lignin monomers and dimers. ChEMBL was used as the open-access database, which contains over 1.9 million distinct bioactive molecules with drug-like properties and 16 million activity measurements [271]. Since minimum inhibitory concentration (MIC) is one the most widely used antimicrobial activity measurements [272], both the ChEMBL and metanalyses datasets used MIC to describe the activity of the compound. For both *B. subtilis* and *E. coli*, three distinct datasets from ChEMBL were obtained by first choosing all the available compounds with MIC measurements against both organisms, selecting a subset of compounds with only C, H, and O atoms (the only atoms present in lignin), and then an additional subset of compounds with at least one phenolic ring. Therefore, increasing the potential specificity of the resulting QSAR model's ability to predict the activity of phenolic lignin derivatives. Due to the large sizes of these ChEMBL datasets, five different regression-based machine learning algorithms were used to create their QSAR models: support vector machine, random forest, k-nearest neighbor, decision tree, and neural networks.

Next, a metanalysis of the available literature with MIC activity measurements for lignin derivatives against both *B. subtilis* and *E. coli* was conducted. Not only was this dataset used to develop a more traditional QSAR model using ordinary least square (OLS) regressions, but it was also used as a validation set for determining the ChEMBL-based model's performance for predicting lignin specific compounds. Finally, a variety of commercially available lignin monomers and dimers were screened for antimicrobial

properties against *B. subtilis* and a subsequent OLS regression-based QSAR was developed. The activity measurement used in the experimental set was the Bacterial Load Difference (BLD) (percent inhibition of growth) as it is more easily measured, encompasses the low antibacterial activity, absence of antibacterial activity, and potential growth-promoting effect sometimes observed with phenolics compared to MIC [160]. The results from this study will provide insights into using different types of databases (open access, metanalysis, experimentally derived, and lignin specific/non-specific) that can be used to develop QSAR models with the potential to predict the antibacterial activity of lignin derivatives.

Materials and Methods

ChEMBL Datasets

Antimicrobial data for both *Bacillus subtilis* and *Escherichia coli,* used as representative Gram-positive and negative bacteria, were obtained from the ChEMBL database (version 27) [271]. Using the ChEMBL web server, a dataset was created for each bacteria type by selecting minimum inhibitory concentration (MIC) as the biological/antimicrobial activity measurement. The datasets were then downloaded, and further filtering was performed in the Python environment.

Firstly, compounds with 'non standard unit for type' or 'outside typical range' in the data validity comments were removed. Then compounds with standard relation values of '<' or '>' were also removed, and duplicates based on compound 'Molecule ChEMBL ID' were averaged into one value. At this point the *B. subtilis* dataset had 9,828 compounds and *E. coli* had 21,657 compounds, which are hereafter referred to as 'B-All' and 'E-All',

respectively. Since lignin has a chemical composition that only contains carbon (C), hydrogen (H), and oxygen (O) atoms, the datasets were further filtered by keeping compounds with only those atoms. This was performed by searching for compounds with a canonical simplified molecular-input line-entry system (SMILES) string with only C, H, and O atoms [273]. The resulting filtering produced a *B. subtilis* dataset with 768 compounds and an *E. coli* dataset with 703 compounds, which are hereafter referred to as 'B-Sorted' and 'E-Sorted,' respectively. Finally, to increase the datasets specificity for predicting lignin phenolics, the previously SMILE sorted dataset was filtered for compounds with at least one phenolic ring. This resulted in a *B. subtilis* dataset with 309 compounds and an *E. coli* dataset with 278 compounds, which are hereafter referred to as 'B-Phenolic' and 'E-Phenolic,' respectively. Therefore, three datasets for both *B. subtilis* and *E. coli* were created with MIC data. Furthermore, MIC values originally determined in μg/ml were converted to micromolar values (μM/ml) and then converted to pMIC (i.e. -logMIC, in molar) for all datasets [268].

Lignin Monomers Metanalysis Dataset

A new dataset of MIC biological activity measurements for lignin monomers against both *B. subtilis* and *E. coli* were compiled from published sources. Multidisciplinary databases such as Academic OneFile, Academic Search Complete, EBSCO, and Google Scholar for terms including combinations such as "lignin," "antimicrobial," "phenolic," "MIC," "monomer," "antibacterial," as well as authors with previous work containing appropriate data, were used to find journal articles that contained MIC antimicrobial data for phenolics that can be derived from lignin. In total, 16 compounds were found with MIC data for *B. subtilis* and 27 compounds for *E. coli* (**Table**

5.5). MIC values originally determined in μg/ml were converted to micromolar values (μM/ml) and then converted to pMIC (i.e. -logMIC, in molar) prior to modeling [268]. The resulting datasets for *B. subtilis* and *E. coli* are hereafter referred to as 'B-Meta' and 'E-Meta', respectively.

Experimental Dataset

The antibacterial activity of 25 lignin derived monomers and three dimers were assessed by monitoring the cell growth (as represented by the optical density at 600 nm, OD_{600}) of *B. subtilis* (NRRL B-354) using a spectrophotometry. The full list of compounds and subsequent antimicrobial activity measurements are listed in **Table 5.7**. The monomers were of analytical quality and purchased from either Sigma Aldrich (St. Louis, MO, USA) or TCI America. The guaiacylglycerol-beta-guaiacyl ether dimer was purchased from TCI America, while 2-(2-methoxyphenoxy)-1-(4-methoxyphenyl)ethanol and 3-hydroxy-2-(2-methoxyphenoxy)-1-(4-methoxyphenyl)-1-propanone dimers were kindly provided by Dr. Mark Crocker at the Center for Applied Energy, University of Kentucky [274].

Briefly, frozen cultures were first revived in liquid growth media (LB broth, Fisher BioReagents™, BP9723) and allowed to grow at 180 rpm shaking speed for 12 h at 37°C. Afterwards the cells were pelletized, washed, and resuspended in fresh liquid media. To test for the antimicrobial properties, each microbe was cultivated in 96-well plates and the OD_{600} was monitored for 24 h with time points taken every 10 min. All wells were brought to an OD_{600} of 0.2 prior to growth, and the phenolics were added to treatment wells to create a final concentration of 1 g/L. To facilitate the solubility of the phenolics in media, all cultures had a final ethanol concentration of 5% (v/v). Two controls were used, one having the 5% ethanol concentration, and one having just microbes and the media. To

determine how the phenolics affected microbial growth, the percent change in OD_{600} of the ethanol control during the exponential phase of growth was compared to the growth of the phenolic treatments. This resulted in the percent decrease in growth or Bacterial Load Difference (BLD) for each phenolic treatment [160], with the formula described in Eq. 1:

$$BLD\ (\%) = \left(1 - \frac{\text{Max OD}_{600} - \text{Min OD}_{600} \text{ with phenolic}}{\text{Max OD}_{600} - \text{Min OD}_{600} \text{ of Ethanol Control}}\right) * 100$$

(1)

After obtaining the BLD values for each phenolic, the structures of each compound were converted to canonical SMILES strings using PubChem for use in descriptor calculations. The final experimental dataset for *B. subtilis* is here after referred to as 'B-Experimental'.

Descriptor Calculations and QSAR Modeling

To calculate the various molecular descriptors, all the compound's structures in each dataset were converted into canonical SMILES strings [273], if not already provided. These SMILES were then entered into an open-access molecular descriptor calculator software package for Python, RDKit (http://www.rdkit.org). RDKit has a variety of calculatable descriptors that describe a molecule's lipophilicity (i.e. LogP, LogD), topological indices (i.e. fragment complexity, size, polarity), connectivity indices and different molecular fingerprints (i.e. number of hydroxyl groups, phenolic rings, carboxylic acids etc.). A full list of descriptors and their description is provided in **Supplementary Table 5.S1**. While it is possible to create models with all the calculatable descriptors, a variety of descriptor selection methods were utilized to improve model accuracy by reducing dimensionality of input space without losing important information.

For the B-All, E-ALL, B-Sorted, E-Sorted, B-Phenolic, and E-Phenolic datasets 200 of RDKit's available descriptors were calculated. Highly correlated ($|r| \geq 0.8$) and constant descriptors were eliminated from the list for each individual dataset. To further reduce the dimensionality of the predictors (descriptors) a principle component analysis (PCA) was performed using scikit-learn [275]. The number of new principle components to be used was assessed by plotting the number of components vs the percent explained variance, and the number of components that explained 99% of the variance were chosen for each dataset. After the optimal number of principle components were chosen and calculated these values were used as the independent variables for predicting the pMIC values in the subsequent QSAR models. Before modeling, each of the above dataset's with pMIC and PCA data were randomly split into training (80%) and test (20%) sets three times for cross-validation. We compared and utilized five machine learning algorithms to build the QSAR models for the B-All, E-ALL, B-Sorted, E-Sorted, B-Phenolic, and E-Phenolic datasets. They were the support vector machine (Epsilon-Support Vector Regression), random forest regressor, k-nearest neighbors regressor, decision tree regressor, and neural network regressor (Multi-layer Perceptron regressor) algorithms provided by scikit-learn. The specific settings and parameters used to build each machine learning algorithm are provided below. QSAR models were assessed based on their average coefficient of determination (R^2) and root mean squared error (RMSE) based on the predictions made for the three training and test sets.

The best QSAR models constructed from the ChEMBL datasets were further tested for prediction accuracy, by using the metanalysis datasets as a test set for predicting lignin-specific compounds. Kernel density estimate (KDE) plots using the Seaborn plugin for

python were constructed to determine the distribution of each dataset's pMIC values. Furthermore, the applicability domain (AD) for estimating the reliability in the prediction of new compounds from the ChEMBL datasets were evaluated against the metanalysis datasets, according to previous work [276].

For the B-Meta, E-Meta, and B-Experimental datasets all 200 of RDKit's available descriptors were calculated. Pearson's correlation coefficient $(|r| \geq 0.5)$ was used to select a fixed subset of predictors (descriptors) best able to predict the antimicrobial activities (either pMIC or BLD) using the ordinary least squares (OLS) regression analysis [277]. The OLS regressions were performed using Statsmodels [278]. As the number of compounds for each of these datasets were very low (less than 30 compounds), the datasets were not separated into training and test sets due to higher risks of chance correlation and overfitting [268]. For each dataset, the selected descriptors were fed into an OLS regression and backwards elimination was used until the significance of each descriptor coefficient in the model (*p*-value) was less than 0.05, which identified the best fitting model.

Machine Learning Algorithms

All machine learning models were created using scikit-learn and either the default hyper parameters were used or a number of different parameters through a grid search based exploration of model parameter space was utilized [270]. The final parameters used for the machine learning algorithms that used grid search for QSAR model development are reported in **Table 5.2**.

The support vector machine (SVM) or Epsilon-Support Vector regression is a non-linear regression that calculates an optimal hyper-plane where the distance and error

between each data points is minimized [279]. The SVM performed here used the default parameters provided by scikit-learn. These included a radial basis function kernel, gamma of 1/number of descriptors, parameter cost of 1, and epsilon of 0.1.

Decision tree regressors (DT) are a non-parametric learning method that works by creating a set of binary rules to calculate the target value by dividing the data into subsets that contain data with similar values [280]. The DT used a grid search to select the optimal maximum depth from 1-21 and minimum number of sample leaves from 1-100 for each dataset, by fitting the training set and using five cross-fold validations and RMSE to choose the best values. All other parameters utilized the scikit-learn default settings.

The random forest regressor (RF) is an ensemble learning method for non-linear regression analysis, that operates by constructing a multitude of decision trees and outputting the mean prediction of the individual trees [281]. We used all the default parameters provided by scikit-learn, but the number of estimators was increased from the default 100 to 500.

K-nearest neighbor regressions (KNN) are a non-parametric method that stores all available cases and predicts a continuous target based on the similarity measure (distance function) between different features in the same neighborhood [282]. The KNN used a grid search to select the optimal number of neighbors from 2-15 for each dataset, by fitting the training set using five cross-fold validations and RMSE to choose the best number of neighbors. The rest of the parameters including the weight function and leaf size utilized scikit-learns default settings.

Neural networks (NN) are brain-inspired algorithms where input features are fed into an input layer, and after a number of nonlinear transformations are performed in a hidden layer, the predictions are generated in an output layer to produce a regression [269, 270, 283]. The NN relied on a Multi-Perceptron regressor along with a grid search technique to select the optimal hidden layer sizes [(50,50,50), (50,100,50), (100,)], activation (rectified linear unit function 'relu' or hyperbolic tan function 'tanh'), and learning rate (constant or adaptive) by using five cross-fold validations and RMSE to choose the best values. All other parameters utilized the scikit-learn default settings.

Software Used

Python (version 3.7.7) was used with the following libraries: RDKit (version 2020.03.6) for the calculation of fingerprints and descriptors, Scikit-learn (version 0.23.2) for all machine learning algorithms and descriptor selection techniques, seaborn (version 0.11.0) with Matplotlib (version 3.3.2) for all figure visualizations, and Pandas (version 1.1.2) for all dataset analysis and manipulation.

Results and Discussion

ChEMBL Dataset Models

The open access database, ChEMBL, was used to develop datasets of compounds with antimicrobial activity (MIC) against both *B. subtilis* and *E. coli*. These datasets were used alongside machine learning algorithms to develop QSAR models with the potential to predict the antimicrobial activity of lignin-derived phenolics from compounds that are not lignin specific.

The initial ChEMBL datasets created for *B. subtlis* (B-All) and *E. coli* (E-All) contained 9,628 and 21,657 compounds, respectively. These datasets were filtered into two additional subsets, that contained compounds having more similar structures to that of lignin derivatives. The first subset was created by selecting compounds with only C, H, and O atoms, resulting in a *B. subtilis* dataset with 768 compounds (B-Sort) and an *E. coli* dataset with 703 compounds (E-Sort). By removing compounds with nitrogenous, chlorine, or fluorine-based functional groups, the remaining compounds could have more similar chemical characteristics to that of lignin derivatives. Then those subsets were further filtered by selecting compounds with at least one phenolic ring, resulting in a *B. subtilis* dataset with 309 compounds (B-Phenolic) and an *E. coli* dataset with 278 compounds (E-Phenolic). Lignin's antimicrobial properties have been reported to attribute to its phenolic structures, so it was important to include a subset of compounds that contained only phenolic-based structures. These were the final six datasets used for QSAR model development from the ChEMBL database.

The QSAR models for these datasets used antimicrobial activities measured in pMIC (- log MIC, in μM/mL) values as the dependent variable, and molecular descriptors calculated from RDKit were used to develop the independent variables. 200 of RDKit's calculatable descriptors and molecular fingerprints that describe the lipophilicity of a compound (i.e. LogP, LogD), topological indices (i.e. fragment complexity, size, polarity), connectivity indices, and functional groups were chosen based on previous work [160, 268, 284, 285]. The specific descriptors used for each dataset after pre-processing are found in **Supplementary Table 5.S2**, and the number of descriptors is summarized in **Table 5.1**.

Principle component analysis (PCA) was used to reduce the number of descriptors used as the independent variables and the dimensionality of the feature space.

PCA reconstructs features of a dataset into a new set of uncorrelated features called principal components (PCs). The optimal number of new PCs for each dataset was selected by the number of components that explained 99% of the variance in the dependent variable. **Figure 5.1** shows the number of PCs vs. the percent explained variance, and **Table 5.1** summarizes the number that explained 99% of the variance. Since this feature extraction technique creates new independent variables that are less interpretable, the ability to examine how each descriptor influences pMIC is no longer easily obtainable. This is actually beneficial when using the ChEMBL datasets, as we are attempting to predict the antimicrobial properties of lignin with non-lignin based compounds from a data-driven perspective and do not need to understand the exact relationship between these compound's descriptors and pMIC values. Therefore, the QSAR models were developed from the pMIC and PC values from each dataset using five popular regression-based machine learning algorithms: support vector machine (SVM), random forest (RF), k-nearest neighbor (KNN), decision tree (DT), and neural networks (NN).

Table 5.1: Each dataset's final number of compounds, descriptors, and hyperparameters for machine learning algorithms that used grid search parameterization. The datasets denoted with 'B' and 'E' represent the data utilized from ChEMBL for *Bacillus subtilis* and *Escherichia coli*, respectively.

	Data Processing		PCA	k-Nearest Neighbor	Decision Tree		Neural Network		
Dataset	Compounds	Descriptors	Components	Neighbors	Depth	Sample Leaves	Hidden Layers	Activation	Learning Rate
B-All	9828	118	80	3	14	50	(50, 100, 50)	tanh	constant
B-Sort	768	62	40	3	4	50	(50, 50, 50)	relu	constant
B-Phenol	309	61	40	3	3	10	(50, 100, 50)	tanh	constant
E-All	21657	114	80	3	13	20	(100,)	tanh	constant
E-Sort	703	67	40	2	4	1	(100,)	tanh	constant
E-Phenol	278	67	40	5	2	20	(50, 50, 50)	relu	constant
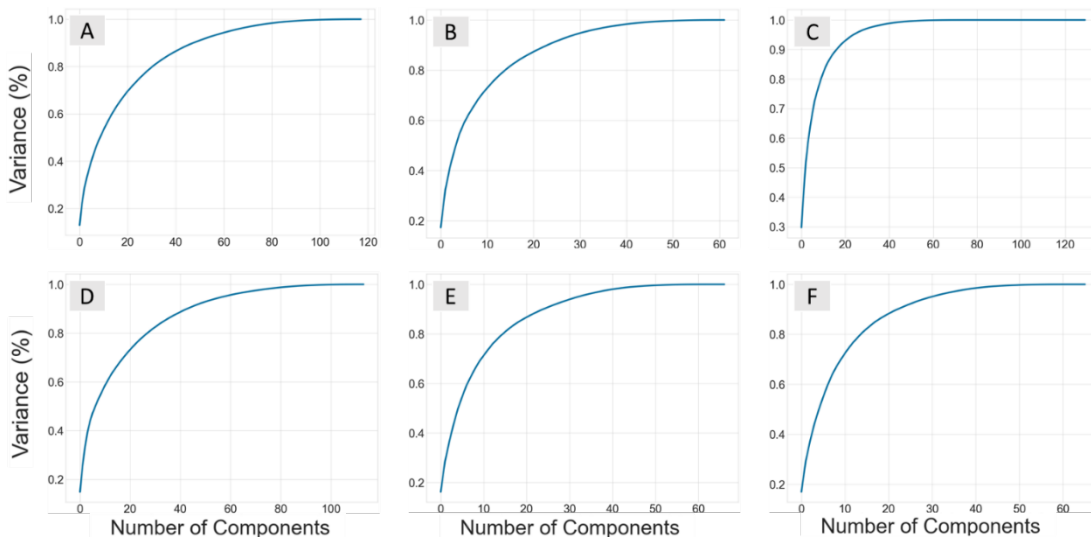
Figure 5.1: Plots showing the number of components from the principle component analysis performed on each datasets descriptor set against the explained variance (%). The ChEMBL datasets for *B. subtilis* are B-All (A), B-Sort (B), and B-Phenol, while the *E. coli* sets are E-All (D), E-Sort (E), and E-Phenol (F).

The performance summary of all five-machine learning QSAR models for each ChEMBL dataset is provided in **Tables 5.2-4**. Each dataset was split randomly into three

different training (80%) and test (20%) sets for cross-validation. The training sets were used to build each machine learning model and the test sets were used for model validation. The metrics used for measuring model performance was the average coefficient of determination (R^2) and root mean square error (RMSE) for the three training and test sets. The better performing model is identified as having a high R^2 and low RMSE value for the average test scores and training scores, and in **Tables 5.2-4** they are highlighted in bold. When comparing models, if one model had a higher R^2 and lower RMSE for the test sets, but not the training sets, the model with better performance for the test set was chosen, as it is ultimately the more important metric [286]. For example, B-All's best performing QSAR model was the KNN algorithm (**Table 5.2**), as it had the highest R^2 of 0.69 for the test set, despite a slightly lower R^2 for the training sets (0.86) compared to the RF algorithm (0.95). Accordingly, the E-All, B-Sort, E-Sort, B-Phenol, and E-phenol datasets had the most robust QSAR models using the RF, NN, KNN, RF, and KNN algorithms, respectively (**Table**s **5.2-4**).

Table 5.2: QSAR model performance for the B-All (*B. subtilis*) and E-All (*E. coli*) ChEMBL datasets using the different machine learning algorithms. Measured by average coefficient of determination (R^2) and root mean square error (RMSE) for both the training and test sets, where values are mean±SE (n=3). Each dataset was split into random test and train sets three different times to obtain the average performance score. The number of compounds, selected descriptors, and number of principle components used to develop models can be found in **Table 5**.**2** and **Supplementary Table 5.S2**.

Model	B-All			
	Test		Train	
	R^2	RMSE	R^2	RMSE
SVM	0.64 ± 0.005	0.66 ± 0.005	0.68 ± 0.001	0.50 ± 0.001
RF	0.65 ± 0.002	0.61 ± 0.006	0.95 ± 0.000	0.24 ± 0.001
KNN	**0.69 ± 0.008**	**0.58 ± 0.009**	**0.86 ± 0.001**	**0.39 ± 0.002**
DT	0.39 ± 0.027	0.81 ± 0.028	0.47 ± 0.008	0.75 ± 0.008
NN	0.63 ± 0.009	0.63 ± 0.004	0.83 ± 0.002	0.42 ± 0.002
Model	**E-All**			
SVM	0.63 ± 0.004	0.65 ± 0.005	0.72 ± 0.002	0.52 ± 0.002
RF	**0.69 ± 0.004**	**0.62 ± 0.002**	**0.95 ± 0.000**	**0.24 ± 0.001**
KNN	0.68 ± 0.005	0.68 ± 0.006	0.71 ± 0.000	0.41 ± 0.000
DT	0.46 ± 0.004	0.82 ± 0.004	0.65 ± 0.004	0.65 ± 0.003
NN	0.63 ± 0.008	0.68 ± 0.003	0.87 ± 0.014	0.40 ± 0.021

Table 5.3: QSAR model performance for the B-Sort (*B. subtilis*) and E-Sort (*E. coli*) ChEMBL datasets using the different machine learning algorithms. Measured by average coefficient of determination (R^2) and root mean square error (RMSE) for both the training and test sets, where values are mean±SE (n=3). Each dataset was split into random test and train sets three different times to obtain the average performance score. The number of compounds, selected descriptors, and number of principle components used to develop models can be found in **Table 5**.**2** and **Supplementary Table 5.S2**..

	B-Sort			
	Test		Train	
Model	R^2	RMSE	R^2	RMSE
SVM	0.49 ± 0.015	0.63 ± 0.015	0.79 ± 0.013	0.42 ± 0.005
RF	0.45 ± 0.038	0.65 ± 0.032	0.92 ± 0.002	0.26 ± 0.003
KNN	0.35 ± 0.028	0.71 ± 0.014	0.78 ± 0.006	0.44 ± 0.006
DT	0.21 ± 0.026	0.78 ± 0.008	0.29 ± 0.004	0.78 ± 0.001
NN	**0.71 ± 0.014**	**0.49 ± 0.005**	**0.79 ± 0.032**	**0.41 ± 0.036**
Model	E-Sort			
SVM	0.41 ± 0.022	0.59 ± 0.024	0.75 ± 0.004	0.53 ± 0.001
RF	0.38 ± 0.023	0.73 ± 0.025	0.92 ± 0.000	0.28 ± 0.001
KNN	**0.49 ± 0.067**	**0.79 ± 0.017**	**0.69 ± 0.007**	**0.42 ± 0.007**
DT	0.00 ± 0.073	0.94 ± 0.040	0.40 ± 0.009	0.78 ± 0.011
NN	0.18 ± 0.041	0.85 ± 0.043	0.88 ± 0.051	0.32 ± 0.095

Table 5.4: : QSAR model performance for the B-Phenol (*B. subtilis*) and E-Phenol (*E. coli*) ChEMBL datasets using the different machine learning algorithms. Measured by average coefficient of determination (R^2) and root mean square error (RMSE) for both the training and test sets, where values are mean±SE (n=3). Each dataset was split into random test and train sets three different times to obtain the average performance score. The number of compounds, selected descriptors, and number of principle components used to develop models can be found in **Table 5.2** and **Supplementary Table 5.S2**.

Model	B-Sort			
	Test		Train	
	R^2	RMSE	R^2	RMSE
SVM	0.55 ± 0.023	0.65 ± 0.113	0.67 ± 0.187	0.31 ± 0.001
RF	0.50 ± 0.023	0.57 ± 0.033	0.93 ± 0.001	0.22 ± 0.002
KNN	**0.57 ± 0.007**	**0.59 ± 0.036**	**0.63 ± 0.005**	**0.42 ± 0.007**
DT	0.49 ± 0.025	0.63 ± 0.015	0.43 ± 0.010	0.62 ± 0.004
NN	0.51 ± 0.052	0.58 ± 0.072	0.80 ± 0.046	0.37 ± 0.049
Model	E-Sort			
SVM	0.22 ± 0.011	0.87 ± 0.040	0.69 ± 0.005	0.56 ± 0.003
RF	0.17 ± 0.049	0.89 ± 0.064	0.89 ± 0.002	0.33 ± 0.007
KNN	**0.38 ± 0.019**	**0.75 ± 0.041**	**0.53 ± 0.015**	**0.76 ± 0.002**
DT	0.02 ± 0.065	1.02 ± 0.053	0.17 ± 0.016	0.90 ± 0.015
NN	0.00 ± 0.085	1.04 ± 0.094	0.88 ± 0.061	0.30 ± 0.117

A common theme with all the models in each dataset, was that the R^2 for the test set was always lower than the training set. This could be a sign of model overfitting or unrepresentative data between the training and test sets [286]. However, all the models had very low SE values when averaging the R^2 values of the three different test/train splits for cross-validation, which would suggest compounds are not being underrepresented. The number of independent variables (PCs) used for each dataset were also rather large (80 or 40), which could contribute to overfitting, but they explained 99% of the dependent variable's variation and when smaller numbers of PCs were used the model's performance drastically decreased (data not shown). Coupled with the fact that most models used a grid

search parametrization technique to fine-tune the hyperparameters, these discrepancies may just be a function of the data itself and not with how the models were evaluated or fit. Furthermore, the E-Sort, B-Phenol, and E-Phenol datasets did not have any QSAR models with a $R^2 > 0.6$, which is usually needed to describe a truly predictive model [154]. Yet, since these datasets are not lignin-specific, the true measure of these model's performance needs to be evaluated with an additional test set of actual lignin-derived compounds.

To this end, the available literature was searched for lignin-derived monomers that had reported MIC values against *B. subtilis* and *E. coli*. The results from this metanalysis are reported in **Table 5.5**, where 16 compounds were found with MIC data for *B. subtilis* (B-Meta) and 27 compounds for *E. coli* (E-Meta). These two datasets were then evaluated as an additional test set for each of the best performing QSAR models found for each ChEMBL dataset described above. The data is summarized in **Figure 5.2**, where the predicted vs. actual pMIC values of the lignin monomers are plotted. It can immediately be seen that none of the ChEMBL QSAR models could accurately predict the lignin monomers. All the models predicted the lignin compounds as having pMIC values roughly less than 2.5, when they are reported as actually having pMIC values greater than 2.5. This suggests these models are grossly underpredicting the pMIC values for the lignin compounds, which would correlate to them having a lower MIC and subsequently greater antimicrobial activity. To understand this, a kernel density estimate (KDE) plot for the ChEMBL and metanalysis datasets were constructed to visualize the distribution of their pMIC values (**Figure 5.3**), and their applicability domains evaluated (**Table 5.6** and **5.7**).

The KDE plots show that the metanalysis datasets for both *E. coli* and *B. subtilis* have pMIC distributions centered around 3-4, while the ChEMBL datasets are centered

between 0-2.5. Even though the ChEMBL datasets clearly contain compounds with pMIC values within the distribution of the metanalysis datasets, they did not lie within the applicability domains (AD) of the ChEMBL datasets. The AD is a useful measure for determining the reliability of a model's prediction for a new set of compounds. Based on the PCA for each ChEMBL dataset, their ADs were calculated based on the Euclidean distances among all their compounds and a final threshold value is determined [276]. Then, the same measure is calculated for each of the compounds in the metanalysis dataset to test if they lie within the threshold of the ChEMBL dataset's AD. We can see in **Table**s **5.6** and 7 that none of the B-Meta or E-Meta compounds fall within the AD of the ChEMBL datasets. Therefore, the use of these ChEMBL datasets, with compounds that are not lignin-specific, to develop a QSAR model that could accurately predict lignin monomer antimicrobial activity was not realized. Even though these results are not what the author had hoped, this data just creates a more concrete conclusion that a comprehensive dataset of lignin derivatives with antimicrobial measurements need to be developed. Therefore, more traditional QSAR models using actual lignin compounds from the metanalysis datasets and an experimentally derived dataset were developed and are discussed in the further sections.

Table 5.5: Source articles that reported antimicrobial data (MIC) and converted pMIC values for phenolics that can be derived from lignin against both *B. subtilis* (B-Meta dataset) and *E. coli* (E-Meta dataset).

Compound	Organism	MIC (ug/ml)	pMIC	Source
caffeic acid	B. subtilis	720.64	3.6	[124]
caffeic acid	B. subtilis	600	3.5	[287]
coniferaldehyde	B. subtilis	712.72	3.6	[124]
coniferyl alcohol	B. subtilis	1441	3.9	[124]
eugenol	B. subtilis	656.8	3.6	[124]
ferulic acid	B. subtilis	388.36	3.3	[124]
p-coumaraldehyde	B. subtilis	296.32	3.3	[124]

p-coumaric acid	B. subtilis	328.0875	3.3	[124]
p-coumaryl alcohol	B. subtilis	1201.396	3.9	[124]
pyrocatechol	B. subtilis	533	3.7	[287]
sinapaldehyde	B. subtilis	1664.589	3.9	[124]
sinapic acid	B. subtilis	448.42	3.3	[124]
sinapyl alcohol	B. subtilis	1682	3.9	[124]
gallic acid	B. subtilis	1600	3.97	[287]
protocatechuic acid	B. subtilis	2667	4.2	[287]
pyrogallol	B. subtilis	267	3.3	[287]
resveratrol	B. subtilis	2667	4.07	[287]
1-(4-hydroxy-3-methoxyphenyl)-2-propanone	E. coli	375	3.3	[123]
2,3-bis(a-hydroxyvanillyl)-l,4-butanediol	E. coli	375	2.99	[123]
4-hydroxy-3-methoxy-b-hydroxy-propiophenonoe	E. coli	375	3.3	[123]
caffeic acid	E. coli	1441.28	3.9	[124]
caffeic acid	E. coli	2667	4.	[287]
coniferaldehyde	E. coli	356.36	3.3	[124]
coniferylalcohol	E. coli	375	3.3	[123]
dehydrodiferulic acid	E. coli	150	2.6	[123]
dehydrodiisoeugenol	E. coli	180	2.7	[123]
di-o-acetylpinoresinol	E. coli	150	2.6	[123]
eugenol	E. coli	656.8	3.6	[124]
eugenol	E. coli	3000	4.26	[123]
ferulic acid	E. coli	388.36	3.30	[124]
ferulic acid	E. coli	375	3.3	[123]
isoeugenol	E. coli	100	2.8	[123]
p-coumaraldehyde	E. coli	296.32	3.3	[124]
p-coumaric acid	E. coli	328.0875	3.3	[124]
p-coumaryl alcohol	E. coli	1201	3.9	[124]
pyrocatechol	E. coli	533	3.7	[287]
sinapaldehyde	E. coli	832.2944	3.6	[124]
sinapic acid	E. coli	448.42	3.3	[124]
sinapyl alcohol	E. coli	1682	3.9	[124]
syringaldehyde	E. coli	375	3.3	[123]
gallic acid	E. coli	600	3.55	[287]
protocatechuic acid	E. coli	2667	4.2	[287]
pyrogallol	E. coli	256	3.3	[288]
pyrogallol	E. coli	83	2.8	[287]
resveratrol	E. coli	3200	4.1	[287]
p-hydroxy benzoic acid	E. coli	400	3.46	[289]
vanillic acid	E. coli	400	3.4	[289]
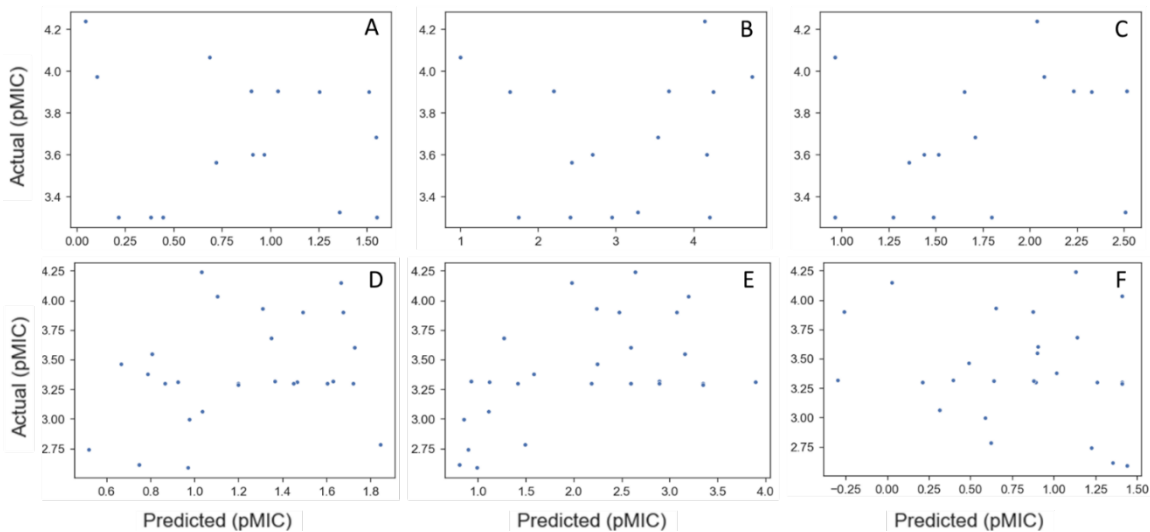
Figure 5.2: Plots of predicted vs actual pMIC values for the B-Meta (A-C) and E-Meta (D-F) datasets by utilizing the best QSAR models developed from the ChEMBL datasets. The ChEMBL datasets used to predict pMIC of the metanalysis datasets for *B. subtilis* are B-All (A), B-Sort (B), and B-Phenol (C), while the *E. coli* sets are E-All (D), E-Sort (E), and E-Phenol (F). The best QSAR models used in each prediction are as follows: RF (A), NN (B), RF (C), RF (D), SVM (E), and KNN (F).
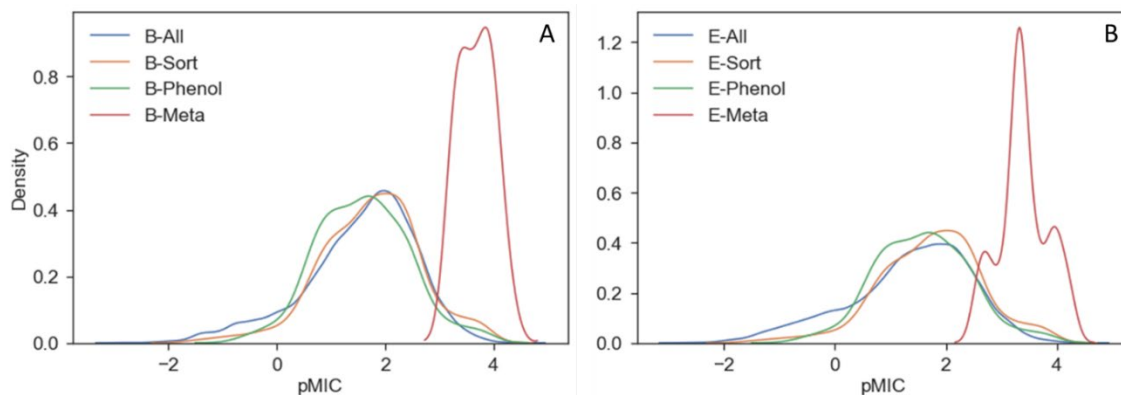
Figure 5.3: Kernel density estimates describing the distribution of pMIC values for the *B. subtilis* (A) and *E. coli* (B) ChEMBL/metanalysis datasets.

Table 5.6: The applicability domain (AD) limit value for each of the *B. subtilis* ChEMBL datasets and the respective Euclidian distances for each compound in the B-Meta dataset. Where True (T) or False (F) indicates if each compound lies within the AD of the each ChEMBL dataset.

Compound	B-All: Limit 13.95		B-Sort: Limit 9.73		B-Phenol: Limit 9.75	
	Distance	T/F	Distance	T/F	Distance	T/F
Caffeic acid	37.48	F	20.16	F	17.65	F
Coniferaldehyde	35.31	F	20.22	F	22.17	F
Coniferyl alcohol	30.53	F	16.28	F	16.40	F
Eugenol	35.56	F	18.04	F	18.82	F
Ferulic acid	36.93	F	21.64	F	20.23	F
Gallic Acid	39.74	F	18.74	F	16.02	F
Protocatechuic acid	32.41	F	16.97	F	14.53	F
Pyrocatechol	26.42	F	15.33	F	14.15	F
Pyrogallol	32.13	F	15.33	F	14.08	F
Resveratrol	40.74	F	22.65	F	21.93	F
Sinapaldehyde	41.58	F	22.93	F	21.52	F
Sinapic acid	43.76	F	27.10	F	20.81	F
Sinapyl alcohol	36.09	F	20.85	F	18.25	F
p-Coumaraldehyde	35.08	F	20.85	F	22.74	F
p-Coumaric acid	38.10	F	20.33	F	18.98	F
p-Coumaryl alcohol	28.69	F	16.93	F	15.97	F

Table 5.7: The applicability domain (AD) limit value for each of the *E. coli* ChEMBL datasets and the respective Euclidian distances for each compound in the B-Meta dataset. Where True (T) or False (F) indicates if each compound lies within the AD of the each ChEMBL dataset.

Compound	E-All: Limit 13.95		E-Sort: Limit 9.77		E-Phenol: Limit 9.71	
	Distance	T/F	Distance	T/F	Distance	T/F
1-(4-hydroxy-3-methoxyphenyl)-2-propanone	24.47	F	10.50	F	11.11	F
2,3-Bis(a-hydroxyvanillyl)-l,4-butanediol	47.50	F	16.79	F	16.38	F
4-Hydroxy-3-methoxy-B-hydroxy-propiophenonoe	27.64	F	10.76	F	10.77	F
Caffeic acid	34.22	F	17.03	F	14.54	F
Coniferaldehyde	44.11	F	17.57	F	17.47	F
Coniferyl alcohol	30.11	F	13.71	F	12.88	F
Dehydrodiferulic acid	47.47	F	24.42	F	21.26	F
Dehydrodiisoeugenol	47.74	F	22.51	F	20.95	F
Di-O-acetylpinoresinol	52.39	F	18.86	F	18.87	F
Eugenol	37.00	F	13.11	F	13.22	F
Ferulic acid	35.01	F	16.36	F	14.47	F
Ferulic acid	35.01	F	16.36	F	14.47	F
Gallic Acid	33.16	F	18.92	F	16.70	F
Protocatechuic acid	27.79	F	15.74	F	14.01	F
Pyrocatechol	26.27	F	15.69	F	15.04	F
Pyrogallol	28.18	F	14.92	F	14.43	F
Resveratrol	36.45	F	16.71	F	16.77	F
Sinapaldehyde	46.28	F	17.81	F	17.07	F
Sinapic acid	40.47	F	18.58	F	15.58	F
Sinapyl alcohol	32.90	F	14.10	F	12.15	F
Vanillic acid	27.31	F	15.64	F	13.01	F
isoeugenol	34.04	F	13.21	F	12.52	F
p-Coumaraldehyde	44.91	F	17.97	F	18.20	F
p-Coumaric acid	34.51	F	16.83	F	15.33	F
p-Coumaryl alcohol	29.86	F	14.31	F	14.06	F
p-hydroxy benzoic acid	25.91	F	15.22	F	13.49	F
syringaldehyde	40.89	F	15.61	F	14.19	F

Metanalysis Dataset Models

The metanalysis datasets, used for validating the ChEMBL QSAR models, were applied to develop their own traditional QSARs using ordinary least square (OLS) regressions. Instead of using PCA as a feature extraction technique, univariate feature selection relying on Pearson's correlation coefficient (r) was employed. Since OLS regressions rely on linear relationships, it made more sense to utilize Pearson's correlation as it measures the strength of the linear correlation between the independent (descriptors) and dependent variables (pMIC). Therefore, the same 200 molecular descriptors from RDKIt were calculated for the B-Meta and E-Meta datasets, and the descriptors with a $r > 0.5$ were selected to develop the OLS regressions. Subsequently, the selected descriptors were fed into an OLS regression and backwards elimination was used until the significance of each descriptor coefficient in the model (p-value) was less than 0.05, which indicated the best fitting model.

No reliable QSARs using OLS was obtained for the E-Meta dataset (results not shown). This result was not that surprising considering the E-Meta datasets pMIC distribution had three different centers, as shown in the KDE plot (**Figure 5.3B**). Its variable distribution and small sample size could prevent the QSAR model from capturing any relevant relationships in the feature space [160]. Conversely, even though the B-Meta (16 compounds) dataset was smaller than E-Meta (27 compounds), a more successful QSAR model was developed.

The model for best predicting the antibacterial activity (pMIC) of the lignin monomers in the B-Meta dataset is summarized in **Table 5.8** and **Figure 5.4**. As observed, the selected OLS model showed good predictive power with a R^2 of 0.759. Three

descriptors, SLogP_VSA3, SLogP_VSA5, and fr_AL_COO were used to develop the best fitting OLS regression model. The SLogP_VSA3 and SLogP_VSA5 descriptors are Molecular Operation Environment (MOE)-type descriptors that bin the output from other descriptor types (i.e SLogP) and calculate the van der Waals (VDWs) surface area (VSA) of atoms contributing to any specified bin of that output. Thus, SLogP_VSA3 and SLogP_VSA5 calculate the sum of VSA contributions to the lipophilicity measurement SLogP (partition coefficient of a compound in two immiscible solvents) within -0.2-0 and 0.1-0.15 bin ranges, respectively. While SLogP and VSA are "primary" descriptors that have more-or-less interpretable contributions to a compound's mechanism of action, the MOE-type descriptors are intended to be used as model predictors and are not as interpretable [290]. Therefore, the negative and positive relationships SLogP_VSA3 and SLogP_VSA5 contribute to the OLS regression can only be used as a data-driven identifier for predicting the pMIC values of lignin compounds. On the other hand, the fr_AL_COO descriptor represents the number of aliphatic carboxylic acid groups in each compound and can directly be used to infer the mechanism of action.

Table 5.8: Statistical performance of the best OLS models obtained through backwards elimination of descriptors, for predicting pMIC values of lignin phenolics against *B. subtilis* in the B-Meta dataset. The compounds used and their pMIC values can be found in **Table 5.1** and the descriptor meaning can be found in **Supplementary Table 5.S1**. N: number of compounds; R^2: coefficient of determination

Dataset	N	R^2	Descriptor	Coefficient	Standard Error	*p*-value
B-Meta	16	0.759	SLogP_VSA3	-0.2951	0.108	0.041
			SLogP_VSA5	0.6025	0.129	0.003
			fr_Al_COO	-0.2588	0.164	0.047
			Intercept	3.5442	0.117	0.000
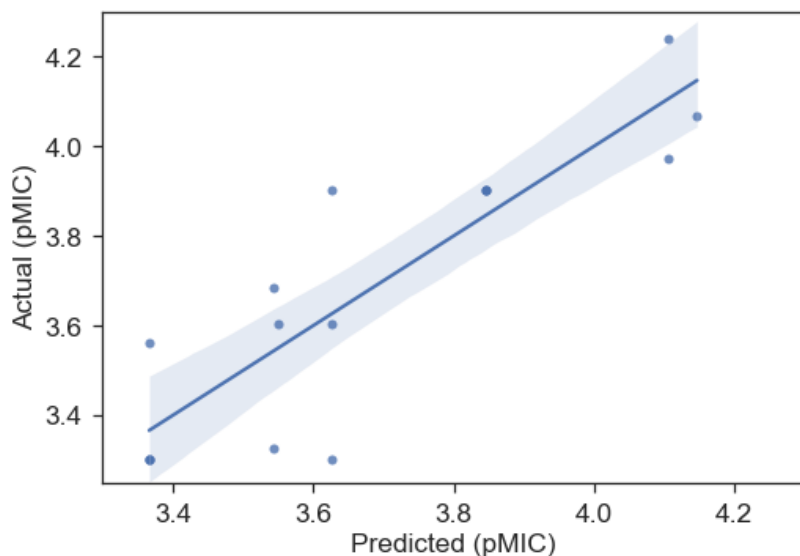
Figure 5.4: Predicted vs actual pMIC regression from the OLS QSAR model for the B-Meta dataset, whose parameters can be found in **Table 5.6**. The shaded region represents the 95% confidence interval for the regression

Caffeic, ferulic, sinapic, and p-coumaric acid were the only compounds with an aliphatic carboxylic acid group present in this dataset and they had the lowest observed pMIC values (~3.3). They also represent hydroxycinnamic acid derivatives that are known to have increased antimicrobial properties compared to their more polar hydroxybenzoic acid counterparts [140]. This is confirmed here by the fact that gallic and protocatechuic acids, with aromatic carboxylic acid groups, had higher MIC values that correspond to lower antimicrobial activity. Previous work has suggested that hydroxycinnamic acid's propenoid side chain is responsible for its increased antimicrobial properties, as it facilitates the transport of the molecule through the cell membrane of Gram-positive bacteria [140, 143, 151]. Therefore, this explains why an increase in aliphatic carboxylic acid groups correlated to an increase in antimicrobial activity (lower pMIC) for this dataset. Nonetheless, the B-Meta dataset only represents an extremely small number of lignin

monomers and more compounds need to be examined to truly understand or predict the properties that influence their antimicrobial activity.

Experimental Dataset Models

The antibacterial activity of 25 lignin-derived monomers and three relevant dimers were assessed by measuring their BLD or percent inhibition against *B. subtilis* at concentrations of 1 g/L. Their BLD values are presented in **Table 5.9** and they ranged from 10% up to 100% (B-Experimental dataset), indicating the existence of completely inhibitory effects. The 3-hydroxy-2-(2-methoxyphenoxy)-1-(4-methoxyphenyl)-1-propanone lignin dimer was the only compound to show complete inhibition against *B. subtilis*, which was ~30% higher than the next highest BLD for 4-ethyl phenol (73.3%). Interestingly, the 2-(2-methoxyphenoxy)-1-(4-methoxyphenyl)ethanol dimer only had a BLD of 66% and its chemical structure differs only by an absence of a methoxy group on β-carbon compared to 3-hydroxy-2-(2-methoxyphenoxy)-1-(4-methoxyphenyl)-1-propanone. Therefore, the presence of this one methoxy group seems to increase the molecules BLD by ~34%. Moreover, by simply examining the chemical structures of the compounds and their BLD values in the B-Experimental dataset, we can immediately see that alkyl chains on the phenolic subunit (4-ethylphenol) and lignin dimers themselves play an important role in these lignin derivatives antimicrobial properties (i.e. higher BLD values). However, the development of a QSAR model will provide a statistical relationship between these molecules BLD values and descriptors for more predictive purposes.

Table 5.9: Experimental antimicrobial activity of lignin monomers and dimers against *B. subtilis* (BLD %), where experimental values are mean±SE (n=3). The predicted BLD values obtained from the OLS QSAR model developed for the B-Experimental dataset, whose parameters can be found in **Table 5.8**.

Type	Compound	Experimental (BLD %)	Predicted (BLD %)
Monomers	2-6-dimethoxyphenol	42.44±6.05	34.49
	4-ethyl phenol	62.43±1.11	69.22
	4-propyl phenol	73.34±0.04	70.08
	acetovanillone	46.13±3.69	43.97
	coniferyl alcohol	35.74±3.22	31.90
	coniferyl aldehyde	36.89±13.35	38.04
	ethyl 3,4 hydroxy propionate	64.33±0.60	58.48
	eugenol	60.77±2.27	66.71
	ferulic acid	36.89±13.35	33.65
	gallic acid	31.13±0.54	23.27
	guiacol	23.24±2.10	44.67
	homosyringic acid	29.94±3.81	41.05
	homovanillic acid	37.73±2.09	32.43
	hydroquinone	35.06±0.73	29.24
	p-coumaric acid	46.05±3.60	43.83
	p-coumaryl alcohol	43.51±5.88	34.98
	p-creosol	64.68±3.67	57.37
	syringaldehyde	44.29±4.75	38.95
	syringic acid	26.64±1.88	31.89
	syringyl alcohol	37.86±3.41	34.77
	syringyl propane	48.07±0.43	49.46
	vanillic acid	43.82±4.09	36.49
	vanillin	16.10±3.86	34.53
	protocatechuic acid	10.08±2.36	21.03
	Catchetol	19.22±6.99	22.55
Dimers	2-(2-methoxyphenoxy)-1-(4-methoxyphenyl)ethanol	66.00±13.79	58.75
	3-hydroxy-2-(2-methoxyphenoxy)-1-(4-methoxyphenyl)-1-propanone	100.00±0.00	100
	Guaiacylglycerol-beta-guaiacyl ether	30.97±1.03	31.51

The same methods used to the develop the QSAR models for the B-Meta dataset were used for the B-Experimental dataset. Where RDKit's calculated descriptors were chosen based on univariate feature selection (r>0.5) and an OLS regression with backwards elimination was performed until all descriptors had a *p*-value less than 0.05. The best fitting

OLS regression in summarized in **Table 5.10** and the predicted vs actual BLD values are plotted in **Figure 5.5**. As observed, the selected OLS model showed better predictive power with a R^2 of 0.831 than that of the B-Meta dataset. Four descriptors were used to develop the best fitting OLS regression model: MinABSEStateIndex, PEOE_VSA13, VSA_EState8, and fr_Ar_OH. As stated previously, PEOE_VSA13 and VSA_EState8 are MOE-type descriptors that are intended to be used as model predictors and are not interpretable for describing the compound's mechanism of action [290]. The MinABSEStateIndex is the minimum absolute electrotopological state (E-state) of a skeletal atom, formulated as an intrinsic value plus a perturbation term arising from the electronic interaction and modified by the molecular topological environment of each atom in the molecule [291]. This descriptor, like the MOE-type descriptors, is used as more of a classification tool for identifying similar compounds instead of describing a feature that could relate to the compound's mode of action. Therefore, while the MinABSEStateIndex, PEOE_VSA13 and VSA_EState8 descriptors show a positive relationship to the lignin compound's BLD value against *B. subtilis*, this information can only be used to classify future compounds. Comparatively, fr_Ar_OH represents the number of aromatic hydroxyl groups in each compound and is better at elucidating their potential antibacterial mechanisms.

The number of aromatic hydroxyl groups can be seen to have a negative relationship with BLD (**Table 5.10**). Where gallic acid, catechol, and protocatechuic acid had more than one aromatic hydroxyl group and the lowest BLD values compared to all the other compounds. So, with an increase in the number of aromatic hydroxyl groups, there will be a decrease in BLD, correlating to a decrease in the compound's antibacterial

properties against *B. subtilis*. Bouareab-Chibane et al. [160] found a negative relationship between the number of hydrogen donors and the BLD of plant-based polyphenols screened against *B. subitilis*. Since the number of aromatic hydroxyl groups and the number of hydrogen donors have a direct positive relationship [292], we can see that in general phenolics with higher overall polarity will have a decrease in antimicrobial properties. This is supported by the experimental data seen here, where highly lipophilic compounds like 4-ethylphenol had high BLD values. However, this model does not provide an explanation for the 3-hydroxy-2-(2-methoxyphenoxy)-1-(4-methoxyphenyl)-1-propanone lignin dimers high BLD value compared to the monomers, highlighting the issue QSAR models can have with limited data sizes and breadth of compound variability. Additionally, when comparing the results from the QSAR models for the B-Meta and B-Experimental datasets, we can see that the presence of certain compounds and how antimicrobial activity was measured will influence which descriptors play the most important role in describing antimicrobial activity.

For example, even though both datasets used the same descriptors and organisms, the best OLS regressions for each dataset found that different descriptors played an important role in describing activity. We saw that the hydroxycinnamic derivatives in the B-Meta dataset drove the negative relationship between the number of aliphatic carboxylic acid groups and pMIC. At the same time, a higher number of aromatic hydroxyl groups were shown to decrease the BLD in the B-Experimental dataset. This emphasizes the fact that using different measures of antimicrobial properties and different lignin compounds to develop QSARs for predicting the antimicrobial properties of lignin will tell different stories. While this is intuitive, the data here provide more concrete support for the need to

169

develop a comprehensive and cohesive dataset with lignin derivatives and their antimicrobial properties. Without such a dataset, we can see that our ability to accurately predict the antimicrobial potential of lignin and its ever-growing variety of derivatives is extremely difficult.

Table 5.10: Statistical performance of the best OLS models obtained through backwards elimination of descriptors, for predicting BLD (%) values of lignin phenolics against *B. subtilis* in the B-Experimental dataset. The compounds used and their BLD values can be found in **Table 5.7** and the descriptor meaning can be found in **Supplementary Table 5.S1**. N: number of compounds; R^2: coefficient of determination.

Dataset	N	R^2	Descriptor	Coefficient	Standard Error	*p*-value
B-Experimental	28	0.831	MinABSEStateIndex	24.7939	7.229	0.002
			PEOE_VSA13	32.0858	11.202	0.009
			VSA_EState8	25.1929	7.769	0.004
			fr_Ar_OH	-43.8297	10.158	0.000
			Intercept	43.9883	4.545	0.000
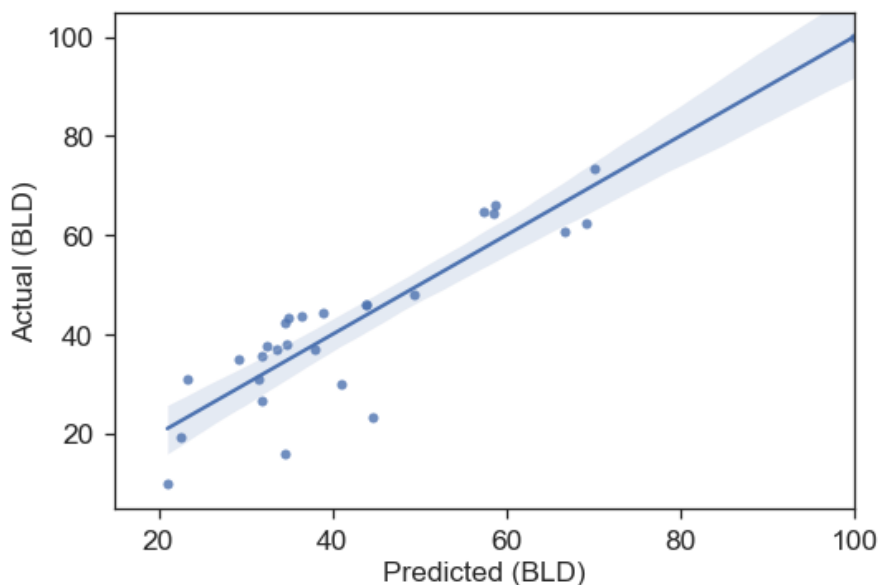
Figure 5.5: Predicted vs actual BLD (%) regression from the OLS QSAR model for the B-Experimental dataset, whose parameters can be found in **Table 5.8**. The shaded region represents the 95% confidence interval for the regression.

Conclusions

The open-access database ChEMBL was used to create three different datasets of compounds with MIC activity measurements against both *B. subtilis* and *E. coli*. Despite the fact these compounds are not lignin-based, two of these datasets had compounds with only C, H, and O or phenolic based structures to increase their chemical similarity to lignin. The QSAR models developed using machine learning algorithms for each of these datasets were found to underpredict the antimicrobial activity of actual lignin compounds compiled from a metanalysis of the literature. Therefore, this data creates a more concrete conclusion that a dataset of lignin derivatives with antimicrobial measurements must be used to develop accurate QSARs. Consequently, more traditional QSAR models using OLS regressions were created from datasets with actual lignin compounds.

The metanalysis data used to validate the ChEMBL dataset's QSAR models for *B. subtills* and *E. coli* were used to build these more traditional QSARs. An accurate QSAR model for *E. coli* was not found, but a satisfactory model was obtained for the *B. subtilis* metanalysis dataset. MOE-type descriptors and the number of aliphatic carboxylic acid groups were the descriptors that showed strong correlations to the pMIC values (R^2 of 0.759). Specifically, as the number of aliphatic carboxylic acid groups increased, the model predicted an increase in antimicrobial activity (i.e. lower MIC). Comparatively, an additional dataset was experimentally derived by screening 25 lignin monomers and three dimers against *B. subtilis* by measuring BLD. The experimentally based QSAR found that MOE-type descriptors and the number of aromatic hydroxyl groups were better predictors of BLD (R^2 of 0.831). Thus, we see that these smaller datasets and their QSARs show how the variability in antimicrobial measurements and the specific compounds used will impact the predictive nature of the resulting QSARs. In combination, the results from this study strongly support that future studies using QSAR to predict the antimicrobial properties of lignin-derived compounds must utilize a more comprehensive and cohesive dataset.

Supplemental Information

Figure 5.S1: Molecular descriptors calculated using RDKit and their descriptions.

Descriptor Name	Description	Dimension	Extended class
BalabanJ	Balaban's J value for a molecule,Chem. Phys. Lett. 89:399-404 (1982).	2	Topological descriptors
BertzCT	A topological index meant to quantify "complexity" of molecules.J. Am. Chem. Soc. 103:3599-601 (1981).	2	Topological descriptors
Chi0	From equations (1),(9) and (10) of Rev. Comp. Chem. vol 2, 367-422, (1991)	2	Connectivity descriptors
Chi1	From equations (1),(11) and (12) of Rev. Comp. Chem. vol 2, 367-422, (1991)	2	Connectivity descriptors
Chi0v	From equations (5),(9) and (10) of Rev. Comp. Chem. vol 2, 367-422, (1991)	2	Connectivity descriptors
Chi1v	From equations (5),(11) and (12) of Rev. Comp. Chem. vol 2, 367-422, (1991)	2	Connectivity descriptors
Chi2v	From equations (5),(15) and (16) of Rev. Comp. Chem. vol 2, 367-422, (1991)	2	Connectivity descriptors
Chi3v	From equations (5),(15) and (16) of Rev. Comp. Chem. vol 2, 367-422, (1991)	2	Connectivity descriptors
Chi4v	From equations (5),(15) and (16) of Rev. Comp. Chem. vol 2, 367-422, (1991)	2	Connectivity descriptors
Chi0n	Similar to Hall Kier Chi0v, but uses nVal instead of valence This makes a big difference after we get out of the first row.Rev. Comput. Chem. 2:367-422 (1991).	2	Connectivity descriptors
Chi1n	Similar to Hall Kier Chi1v, but uses nVal instead of valence.Rev. Comput. Chem. 2:367-422 (1991).	2	Connectivity descriptors
Chi2n	Similar to Hall Kier Chi2v, but uses nVal instead of valence This makes a big difference after we get out of the first row.Rev. Comput. Chem. 2:367-422 (1991).	2	Connectivity descriptors
Chi3n	Similar to Hall Kier Chi3v, but uses nVal instead of valence This makes a big difference after we get out of the first row.Rev. Comput. Chem. 2:367-422 (1991).	2	Connectivity descriptors
Chi4n	Similar to Hall Kier Chi4v, but uses nVal instead of valence.This makes a big difference after we get out of	2	Connectivity descriptors

	the first row.Rev. Comput. Chem. 2:367-422 (1991).		
EState_VSA1	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
EState_VSA2	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
EState_VSA3	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
EState_VSA4	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
EState_VSA5	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
EState_VSA6	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
EState_VSA7	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
EState_VSA8	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
EState_VSA9	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
EState_VSA10	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
EState_VSA11	MOE-type descriptors using EState indices and surface area contributions (developed at RD, not described in the CCG paper).	2	MOE-type descriptors
ExactMolWt	The molecule's exact molecular weight.	2	Molecular property descriptors
FractionCSP3	The fraction of C atoms that are SP3 hybridized.	1	Constitutional descriptors
HallKierAlpha	The Hall-Kier alpha value for a molecule.Rev. Comput. Chem. 2:367-422 (1991).	2	Topological descriptors
HeavyAtomCount	Number of heavy atoms of a molecule.	1	Constitutional descriptors

HeavyAtomMolWt	The average molecular weight of the molecule ignoring hydrogens	1	Constitutional descriptors
Ipc	the information content of the coefficients of the characteristic polynomial of the adjacency matrix of a hydrogen-suppressed graph of a molecule.	2	Topological descriptors
Kappa1	Hall-Kier Kappa1 value	2	Topological descriptors
Kappa2	Hall-Kier Kappa2 value	2	Topological descriptors
Kappa3	Hall-Kier Kappa2 value	2	Topological descriptors
LabuteASA	Labute's Approximate Surface Area (ASA from MOE)	2	MOE-type descriptors
MolLogP	Wildman-Crippen LogP value.Wildman and Crippen JCICS 39:868-73 (1999)	2	Molecular property descriptors
MolMR	Wildman-Crippen MR value.Wildman and Crippen JCICS 39:868-73 (1999)	2	Molecular property descriptors
MolWt	The average molecular weight of the molecule	2	Molecular property descriptors
NHOHCount	Number of NHs or OHs	1	Constitutional descriptors
NOCount	Number of Nitrogens and Oxygens	1	Constitutional descriptors
NumAliphaticCarbocycles	The number of aliphatic (containing at least one non-aromatic bond) carbocycles for a molecule	1	Constitutional descriptors
NumAliphaticHeterocycles	The number of aliphatic (containing at least one non-aromatic bond) heterocycles for a molecule	1	Constitutional descriptors
NumAliphaticRings	The number of aliphatic (containing at least one non-aromatic bond) rings for a molecule	1	Constitutional descriptors
NumAromaticCarbocycles	The number of aromatic carbocycles for a molecule	1	Constitutional descriptors
NumAromaticHeterocycles	The number of aromatic heterocycles for a molecule	1	Constitutional descriptors
NumAromaticRings	The number of aromatic rings for a molecule	1	Constitutional descriptors
NumHAcceptors	Number of Hydrogen Bond Acceptors	1	Constitutional descriptors
NumHDonors	Number of Hydrogen Bond Donors	1	Constitutional descriptors
NumHeteroatoms	Number of Heteroatoms	1	Constitutional descriptors
NumRadicalElectrons	The number of radical electrons the molecule has (says nothing about spin state)	1	Constitutional descriptors
NumRotatableBonds	Number of Rotatable Bonds	1	Constitutional descriptors
NumSaturatedCarbocycles	The number of saturated carbocycles for a molecule	1	Constitutional descriptors
NumSaturatedHeterocycles	The number of saturated heterocycles for a molecule	1	Constitutional descriptors

NumSaturatedRings	The number of saturated rings for a molecule	1	Constitutional descriptors
NumValenceElectrons	The number of valence electrons the molecule has	1	Constitutional descriptors
PEOE_VSA1	MOE Charge VSA Descriptor 1 ($-inf < x < -0.30$)	2	MOE-type descriptors
PEOE_VSA2	MOE Charge VSA Descriptor 2 ($-0.30 <= x < -0.25$)	2	MOE-type descriptors
PEOE_VSA3	MOE Charge VSA Descriptor 3 ($-0.25 <= x < -0.20$)	2	MOE-type descriptors
PEOE_VSA4	MOE Charge VSA Descriptor 4 ($-0.20 <= x < -0.15$)	2	MOE-type descriptors
PEOE_VSA5	MOE Charge VSA Descriptor 5 ($-0.15 <= x < -0.10$)	2	MOE-type descriptors
PEOE_VSA6	MOE Charge VSA Descriptor 6 ($-0.10 <= x < -0.05$)	2	MOE-type descriptors
PEOE_VSA7	MOE Charge VSA Descriptor 7 ($-0.05 <= x < 0.00$)	2	MOE-type descriptors
PEOE_VSA8	MOE Charge VSA Descriptor 8 ($0.00 <= x < 0.05$)	2	MOE-type descriptors
PEOE_VSA9	MOE Charge VSA Descriptor 9 ($0.05 <= x < 0.10$)	2	MOE-type descriptors
PEOE_VSA10	MOE Charge VSA Descriptor 10 ($0.10 <= x < 0.15$)	2	MOE-type descriptors
PEOE_VSA11	MOE Charge VSA Descriptor 11 ($0.15 <= x < 0.20$)	2	MOE-type descriptors
PEOE_VSA12	MOE Charge VSA Descriptor 12 ($0.20 <= x < 0.25$)	2	MOE-type descriptors
PEOE_VSA13	MOE Charge VSA Descriptor 13 ($0.25 <= x < 0.30$)	2	MOE-type descriptors
PEOE_VSA14	MOE Charge VSA Descriptor 14 ($0.30 <= x < inf$)	2	MOE-type descriptors
RingCount	The number of rings for a molecule	1	Constitutional descriptors
SMR_VSA1	MOE MR VSA Descriptor 1 ($-inf < x < 1.29$)	2	MOE-type descriptors
SMR_VSA2	MOE MR VSA Descriptor 2 ($1.29 <= x < 1.82$)	2	MOE-type descriptors
SMR_VSA3	MOE MR VSA Descriptor 3 ($1.82 <= x < 2.24$)	2	MOE-type descriptors
SMR_VSA4	MOE MR VSA Descriptor 4 ($2.24 <= x < 2.45$)	2	MOE-type descriptors
SMR_VSA5	MOE MR VSA Descriptor 5 ($2.45 <= x < 2.75$)	2	MOE-type descriptors
SMR_VSA6	MOE MR VSA Descriptor 6 ($2.75 <= x < 3.05$)	2	MOE-type descriptors
SMR_VSA7	MOE MR VSA Descriptor 7 ($3.05 <= x < 3.63$)	2	MOE-type descriptors
SMR_VSA8	MOE MR VSA Descriptor 8 ($3.63 <= x < 3.80$)	2	MOE-type descriptors
SMR_VSA9	MOE MR VSA Descriptor 9 ($3.80 <= x < 4.00$)	2	MOE-type descriptors
SMR_VSA10	MOE MR VSA Descriptor 10 ($4.00 <= x < inf$)	2	MOE-type descriptors

SlogP_VSA1	MOE logP VSA Descriptor 1 (-inf < x < -0.40)	2	MOE-type descriptors
SlogP_VSA2	MOE logP VSA Descriptor 2 (-0.40 <= x < -0.20)	2	MOE-type descriptors
SlogP_VSA3	MOE logP VSA Descriptor 3 (-0.20 <= x < 0.00)	2	MOE-type descriptors
SlogP_VSA4	MOE logP VSA Descriptor 4 (0.00 <= x < 0.10)	2	MOE-type descriptors
SlogP_VSA5	MOE logP VSA Descriptor 5 (0.10 <= x < 0.15)	2	MOE-type descriptors
SlogP_VSA6	MOE logP VSA Descriptor 6 (0.15 <= x < 0.20)	2	MOE-type descriptors
SlogP_VSA7	MOE logP VSA Descriptor 7 (0.20 <= x < 0.25)	2	MOE-type descriptors
SlogP_VSA8	MOE logP VSA Descriptor 8 (0.25 <= x < 0.30)	2	MOE-type descriptors
SlogP_VSA9	MOE logP VSA Descriptor 9 (0.30 <= x < 0.40)	2	MOE-type descriptors
SlogP_VSA10	MOE logP VSA Descriptor 10 (0.40 <= x < 0.50)	2	MOE-type descriptors
SlogP_VSA11	MOE logP VSA Descriptor 11 (0.50 <= x < 0.60)	2	MOE-type descriptors
SlogP_VSA12	MOE logP VSA Descriptor 12 (0.60 <= x < inf)	2	MOE-type descriptors
TPSA	The polar surface area of a molecule based upon fragments	2	Molecular property descriptors
VSA_EState1	VSA EState Descriptor 1 (-inf < x < 4.78)	2	MOE-type descriptors
VSA_EState2	VSA EState Descriptor 2 (4.78 <= x < 5.00)	2	MOE-type descriptors
VSA_EState3	VSA EState Descriptor 3 (5.00 <= x < 5.41)	2	MOE-type descriptors
VSA_EState4	VSA EState Descriptor 4 (5.41 <= x < 5.74)	2	MOE-type descriptors
VSA_EState5	VSA EState Descriptor 5 (5.74 <= x < 6.00)	2	MOE-type descriptors
VSA_EState6	VSA EState Descriptor 6 (6.00 <= x < 6.07)	2	MOE-type descriptors
VSA_EState7	VSA EState Descriptor 7 (6.07 <= x < 6.45)	2	MOE-type descriptors
VSA_EState8	VSA EState Descriptor 8 (6.45 <= x < 7.00)	2	MOE-type descriptors
VSA_EState9	VSA EState Descriptor 9 (7.00 <= x < 11.00)	2	MOE-type descriptors
VSA_EState10	VSA EState Descriptor 10 (11.00 <= x < inf)	2	MOE-type descriptors
fr_Al_COO	Number of aliphatic carboxylic acids	1	Constitutional descriptors
fr_Al_OH	Number of aliphatic hydroxyl groups	1	Constitutional descriptors
fr_Al_OH_noTert	Number of aliphatic hydroxyl groups excluding tert-OH	1	Constitutional descriptors
fr_ArN	Number of N functional groups attached to aromatics	1	Constitutional descriptors

fr_Ar_COO	Number of Aromatic carboxylic acide	1	Constitutional descriptors
fr_Ar_N	Number of aromatic nitrogens	1	Constitutional descriptors
fr_Ar_NH	Number of aromatic amines	1	Constitutional descriptors
fr_Ar_OH	Number of aromatic hydroxyl groups	1	Constitutional descriptors
fr_COO	Number of carboxylic acids	1	Constitutional descriptors
fr_COO2	Number of carboxylic acids	1	Constitutional descriptors
fr_C_O	Number of carbonyl O	1	Constitutional descriptors
fr_C_O_noCOO	Number of carbonyl O, excluding COOH	1	Constitutional descriptors
fr_C_S	Number of thiocarbonyl	1	Constitutional descriptors
fr_HOCCN	Number of C(OH)CCN-Ctert-alkyl or C(OH)CCNcyclic	1	Constitutional descriptors
fr_Imine	Number of Imines	1	Constitutional descriptors
fr_NH0	Number of Tertiary amines	1	Constitutional descriptors
fr_NH1	Number of Secondary amines	1	Constitutional descriptors
fr_NH2	Number of Primary amines	1	Constitutional descriptors
fr_N_O	Number of hydroxylamine groups	1	Constitutional descriptors
fr_Ndealkylation1	Number of XCCNR groups	1	Constitutional descriptors
fr_Ndealkylation2	Number of tert-alicyclic amines (no heteroatoms, not quinine-like bridged N)	1	Constitutional descriptors
fr_Nhpyrrole	Number of H-pyrrole nitrogens	1	Constitutional descriptors
fr_SH	Number of thiol groups	1	Constitutional descriptors
fr_aldehyde	Number of aldehydes	1	Constitutional descriptors
fr_alkyl_carbamate	Number of alkyl carbamates (subject to hydrolysis)	1	Constitutional descriptors
fr_alkyl_halide	Number of alkyl halides	1	Constitutional descriptors
fr_allylic_oxid	Number of allylic oxidation sites excluding steroid dienone	1	Constitutional descriptors
fr_amide	Number of amides	1	Constitutional descriptors
fr_amidine	Number of amidine groups	1	Constitutional descriptors
fr_aniline	Number of anilines	1	Constitutional descriptors
fr_aryl_methyl	Number of aryl methyl sites for hydroxylation	1	Constitutional descriptors

fr_azide	Number of azide groups	1	Constitutional descriptors
fr_azo	Number of azo groups	1	Constitutional descriptors
fr_barbitur	Number of barbiturate groups	1	Constitutional descriptors
fr_benzene	Number of benzene rings	1	Constitutional descriptors
fr_benzodiazepine	Number of benzodiazepines with no additional fused rings	1	Constitutional descriptors
fr_bicyclic	Bicyclic	1	Constitutional descriptors
fr_diazo	Number of diazo groups	1	Constitutional descriptors
fr_dihydropyridine	Number of dihydropyridines	1	Constitutional descriptors
fr_epoxide	Number of epoxide rings	1	Constitutional descriptors
fr_ester	Number of esters	1	Constitutional descriptors
fr_ether	Number of ether oxygens (including phenoxy)	1	Constitutional descriptors
fr_furan	Number of furan rings	1	Constitutional descriptors
fr_guanido	Number of guanidine groups	1	Constitutional descriptors
fr_halogen	Number of halogens	1	Constitutional descriptors
fr_hdrzine	Number of hydrazine groups	1	Constitutional descriptors
fr_hdrzone	Number of hydrazone groups	1	Constitutional descriptors
fr_imidazole	Number of imidazole rings	1	Constitutional descriptors
fr_imide	Number of imide groups	1	Constitutional descriptors
fr_isocyan	Number of isocyanates	1	Constitutional descriptors
fr_isothiocyan	Number of isothiocyanates	1	Constitutional descriptors
fr_ketone	Number of ketones	1	Constitutional descriptors
fr_ketone_Topliss	Number of ketones excluding diaryl, a,b-unsat.	1	Constitutional descriptors
fr_lactam	Number of beta lactams	1	Constitutional descriptors
fr_lactone	Number of cyclic esters (lactones)	1	Constitutional descriptors
fr_methoxy	Number of methoxy groups -OCH3	1	Constitutional descriptors
fr_morpholine	Number of morpholine rings	1	Constitutional descriptors
fr_nitrile	Number of nitriles	1	Constitutional descriptors

fr_nitro	Number of nitro groups	1	Constitutional descriptors
fr_nitro_arom	Number of nitro benzene ring substituents	1	Constitutional descriptors
fr_nitro_arom_nonortho	Number of non-ortho nitro benzene ring substituents	1	Constitutional descriptors
fr_nitroso	Number of nitroso groups, excluding NO2	1	Constitutional descriptors
fr_oxazole	Number of oxazole rings	1	Constitutional descriptors
fr_oxime	Number of oxime groups	1	Constitutional descriptors
fr_para_hydroxylation	Number of para-hydroxylation sites	1	Constitutional descriptors
fr_phenol	Number of phenols	1	Constitutional descriptors
fr_phenol_noOrthoHbond	Number of phenolic OH excluding ortho intramolecular Hbond substituents	1	Constitutional descriptors
fr_phos_acid	Number of phosphoric acid groups	1	Constitutional descriptors
fr_phos_ester	Number of phosphoric ester groups	1	Constitutional descriptors
fr_piperdine	Number of piperdine rings	1	Constitutional descriptors
fr_piperzine	Number of piperzine rings	1	Constitutional descriptors
fr_priamide	Number of primary amides	1	Constitutional descriptors
fr_prisulfonamd	Number of primary sulfonamides	1	Constitutional descriptors
fr_pyridine	Number of pyridine rings	1	Constitutional descriptors
fr_quatN	Number of quarternary nitrogens	1	Constitutional descriptors
fr_sulfide	Number of thioether	1	Constitutional descriptors
fr_sulfonamd	Number of sulfonamides	1	Constitutional descriptors
fr_sulfone	Number of sulfone groups	1	Constitutional descriptors
fr_term_acetylene	Number of terminal acetylenes	1	Constitutional descriptors
fr_tetrazole	Number of tetrazole rings	1	Constitutional descriptors
fr_thiazole	Number of thiazole rings	1	Constitutional descriptors
fr_thiocyan	Number of thiocyanates	1	Constitutional descriptors
fr_thiophene	Number of thiophene rings	1	Constitutional descriptors
fr_unbrch_alkane	Number of unbranched alkanes of at least 4 members (excludes halogenated alkanes)	1	Constitutional descriptors

fr_urea	Number of urea groups	1	Constitutional descriptors
MaxAbsEStateIndex	Returns a tuple of EState indices for the molecule, Reference: Hall, Mohney and Kier. JCICS _31_ 76-81 (1991)	2	Topological descriptors
MaxAbsPartialCharge	Returns molecular charge descriptors	2	Topological descriptors
MaxEStateIndex	Returns a tuple of EState indices for the molecule, Reference: Hall, Mohney and Kier. JCICS _31_ 76-81 (1991)	2	Topological descriptors
MaxPartialCharge	Returns molecular charge descriptors	2	Topological descriptors
MinAbsEStateIndex	Returns a tuple of EState indices for the molecule, Reference: Hall, Mohney and Kier. JCICS _31_ 76-81 (1991)	2	Topological descriptors
MinAbsPartialCharge	Returns molecular charge descriptors	2	Topological descriptors
MinEStateIndex	Returns a tuple of EState indices for the molecule, Reference: Hall, Mohney and Kier. JCICS _31_ 76-81 (1991)	2	Topological descriptors
MinPartialCharge	Returns molecular charge descriptors	2	Topological descripto

Figure 5.S2: Selected descriptors used for the ChEMBL database QSAR model development. Descriptor meaning can be found in Supplementary Table S1.

Bacillus subtilis			*Escherichia coli*		
B-All	**B-Sort**	**B-Phenol**	**E-All**	**E-Sort**	**E-Phenol**
MaxEStateIndex	MaxEStateIndex	MaxEStateIndex	MaxEStateIndex	MaxEStateIndex	MaxEStateIndex
MinEStateIndex	MinEStateIndex	MinEStateIndex	MinEStateIndex	MinEStateIndex	MinEStateIndex
MinAbsEStateIndex	MinAbsEStateIndex	MinAbsEStateIndex	MinAbsEStateIndex	MinAbsEStateIndex	MinAbsEStateIndex
qed	qed	qed	qed	qed	qed
MolWt	MolWt	MolWt	MolWt	MolWt	MolWt
MaxPartialCharge	MaxPartialCharge	MaxPartialCharge	MaxPartialCharge	MaxPartialCharge	MaxPartialCharge
MinPartialCharge	MinPartialCharge	MinPartialCharge	MinPartialCharge	MinPartialCharge	MinPartialCharge
FpDensityMorgan1	FpDensityMorgan1	FpDensityMorgan1	FpDensityMorgan1	FpDensityMorgan1	FpDensityMorgan1
BalabanJ	BalabanJ	BalabanJ	BalabanJ	BalabanJ	BalabanJ
Ipc	Ipc	Ipc	Ipc	HallKierAlpha	HallKierAlpha
Kappa3	PEOE_VSA10	PEOE_VSA10	Kappa3	Ipc	Ipc

PEOE_VSA11	PEOE_VSA11	PEOE_VSA12	PEOE_VSA11	PEOE_VSA1	PEOE_VSA10
PEOE_VSA13	PEOE_VSA12	PEOE_VSA13	PEOE_VSA13	PEOE_VSA10	PEOE_VSA11
PEOE_VSA14	PEOE_VSA13	PEOE_VSA14	PEOE_VSA3	PEOE_VSA11	PEOE_VSA12
PEOE_VSA3	PEOE_VSA14	PEOE_VSA2	PEOE_VSA4	PEOE_VSA12	PEOE_VSA13
PEOE_VSA4	PEOE_VSA2	PEOE_VSA5	PEOE_VSA5	PEOE_VSA13	PEOE_VSA14
PEOE_VSA5	PEOE_VSA3	PEOE_VSA6	PEOE_VSA6	PEOE_VSA14	PEOE_VSA2
PEOE_VSA6	PEOE_VSA5	PEOE_VSA7	PEOE_VSA8	PEOE_VSA2	PEOE_VSA3
PEOE_VSA8	PEOE_VSA6	PEOE_VSA8	PEOE_VSA9	PEOE_VSA3	PEOE_VSA5
PEOE_VSA9	PEOE_VSA7	PEOE_VSA9	SMR_VSA10	PEOE_VSA5	PEOE_VSA6
SMR_VSA10	PEOE_VSA8	SMR_VSA10	SMR_VSA2	PEOE_VSA6	PEOE_VSA7
SMR_VSA2	PEOE_VSA9	SMR_VSA4	SMR_VSA4	PEOE_VSA7	PEOE_VSA8
SMR_VSA4	SMR_VSA10	SMR_VSA5	SMR_VSA6	PEOE_VSA8	PEOE_VSA9
SMR_VSA6	SMR_VSA4	SMR_VSA6	SMR_VSA7	PEOE_VSA9	SMR_VSA10
SMR_VSA7	SMR_VSA6	SMR_VSA7	SMR_VSA9	SMR_VSA10	SMR_VSA6
SMR_VSA9	SMR_VSA7	SlogP_VSA1	SlogP_VSA10	SMR_VSA6	SMR_VSA7
SlogP_VSA10	SlogP_VSA1	SlogP_VSA3	SlogP_VSA11	SMR_VSA7	SMR_VSA9
SlogP_VSA11	SlogP_VSA3	SlogP_VSA4	SlogP_VSA12	SMR_VSA9	SlogP_VSA1
SlogP_VSA12	SlogP_VSA8	SlogP_VSA8	SlogP_VSA4	SlogP_VSA1	SlogP_VSA3
SlogP_VSA4	EState_VSA2	EState_VSA2	SlogP_VSA7	SlogP_VSA3	SlogP_VSA8
SlogP_VSA7	EState_VSA3	EState_VSA3	SlogP_VSA8	SlogP_VSA8	EState_VSA11
SlogP_VSA8	EState_VSA4	EState_VSA4	EState_VSA11	EState_VSA11	EState_VSA2
EState_VSA11	EState_VSA5	EState_VSA5	EState_VSA3	EState_VSA2	EState_VSA3
EState_VSA3	EState_VSA6	EState_VSA6	EState_VSA4	EState_VSA3	EState_VSA4
EState_VSA4	EState_VSA7	EState_VSA7	EState_VSA5	EState_VSA4	EState_VSA5
EState_VSA5	EState_VSA8	EState_VSA8	EState_VSA6	EState_VSA5	EState_VSA6
EState_VSA6	EState_VSA9	EState_VSA9	EState_VSA7	EState_VSA6	EState_VSA7
EState_VSA7	VSA_EState2	VSA_EState6	EState_VSA8	EState_VSA7	EState_VSA8
EState_VSA8	VSA_EState4	VSA_EState7	EState_VSA9	EState_VSA8	EState_VSA9
EState_VSA9	VSA_EState5	VSA_EState8	VSA_EState1	EState_VSA9	VSA_EState2
VSA_EState1	VSA_EState6	NumAliphatic Carbocycles	VSA_EState4	VSA_EState2	VSA_EState4
VSA_EState3	VSA_EState7	NumAliphatic Heterocycles	VSA_EState7	VSA_EState4	VSA_EState5
VSA_EState4	VSA_EState8	NumAliphatic Rings	VSA_EState8	VSA_EState5	VSA_EState6
VSA_EState5	VSA_EState9	NumAromatic Carbocycles	VSA_EState9	VSA_EState6	VSA_EState7
VSA_EState7	FractionCSP3	NumAromatic Heterocycles	FractionCSP3	VSA_EState7	VSA_EState8
VSA_EState8	NumAliphatic Heterocycles	NumSaturated Heterocycles	NumAliphatic Carbocycles	VSA_EState8	NumAliphatic Heterocycles
VSA_EState9	NumAromatic Heterocycles	NumSaturated Rings	NumAliphatic Heterocycles	VSA_EState9	NumAromatic Heterocycles
FractionCSP3	RingCount	MolLogP	NumAromatic Heterocycles	NumAliphatic Heterocycles	NumRotatable Bonds
NumAliphatic Carbocycles	MolLogP	fr_Al_COO	RingCount	NumAromatic Heterocycles	RingCount

NumAliphatic Heterocycles	fr_Al_COO	fr_Al_OH	MolLogP	RingCount	MolLogP
NumAromatic Heterocycles	fr_Ar_COO	fr_Ar_COO	fr_Al_COO	MolLogP	fr_Al_COO
RingCount	fr_aldehyde	fr_aldehyde	fr_Al_OH	fr_Al_COO	fr_Ar_COO
MolLogP	fr_allylic_oxid	fr_allylic_oxid	fr_ArN	fr_Ar_COO	fr_Ar_OH
fr_Al_COO	fr_aryl_methyl	fr_aryl_methyl	fr_Ar_COO	fr_Ar_OH	fr_aldehyde
fr_Al_OH	fr_bicyclic	fr_bicyclic	fr_Ar_NH	fr_aldehyde	fr_allylic_oxid
fr_ArN	fr_epoxide	fr_epoxide	fr_Ar_OH	fr_allylic_oxid	fr_aryl_methyl
fr_Ar_COO	fr_furan	fr_furan	fr_COO	fr_aryl_methyl	fr_bicyclic
fr_Ar_NH	fr_ketone	fr_ketone	fr_C_S	fr_bicyclic	fr_epoxide
fr_Ar_OH	fr_ketone_Topliss	fr_lactone	fr_HOCCN	fr_epoxide	fr_ester
fr_COO	fr_lactone	fr_para_hydroxylation	fr_Imine	fr_ester	fr_furan
fr_C_S	fr_para_hydroxylation	fr_unbrch_alkane	fr_NH0	fr_furan	fr_ketone
fr_HOCCN	fr_term_acetylene		fr_N_O	fr_ketone	fr_ketone_Topliss
fr_Imine			fr_Ndealkylation1	fr_ketone_Topliss	fr_lactone
fr_NH0			fr_Ndealkylation2	fr_lactone	fr_para_hydroxylation
fr_N_O			fr_SH	fr_para_hydroxylation	fr_term_acetylene
fr_Ndealkylation1			fr_aldehyde	fr_term_acetylene	fr_unbrch_alkane
fr_Ndealkylation2			fr_alkyl_carbamate	fr_unbrch_alkane	
fr_SH			fr_alkyl_halide		
fr_aldehyde			fr_allylic_oxid		
fr_alkyl_carbamate			fr_amidine		
fr_alkyl_halide			fr_aniline		
fr_allylic_oxid			fr_aryl_methyl		
fr_amidine			fr_azide		
fr_aniline			fr_azo		
fr_aryl_methyl			fr_barbitur		
fr_azide			fr_bicyclic		
fr_azo			fr_dihydropyridine		
fr_barbitur			fr_epoxide		
fr_bicyclic			fr_ester		
fr_dihydropyridine			fr_ether		
fr_epoxide			fr_furan		
fr_ester			fr_halogen		
fr_ether			fr_hdrzine		
fr_furan			fr_hdrzone		

fr_guanido			fr_imidazole		
fr_halogen			fr_imide		
fr_hdrzine			fr_isothiocyan		
fr_hdrzone			fr_ketone		
fr_imidazole			fr_lactam		
fr_imide			fr_lactone		
fr_isothiocyan			fr_methoxy		
fr_ketone			fr_morpholine		
fr_lactam			fr_nitrile		
fr_lactone			fr_nitro		
fr_methoxy			fr_nitroso		
fr_morpholine			fr_oxazole		
fr_nitrile			fr_oxime		
fr_nitro			fr_para_hydro xylation		
fr_nitroso			fr_phos_acid		
fr_oxazole			fr_piperdine		
fr_oxime			fr_piperzine		
fr_para_hydro xylation			fr_priamide		
fr_phos_acid			fr_pyridine		
fr_piperdine			fr_quatN		
fr_piperzine			fr_sulfide		
fr_priamide			fr_sulfonamd		
fr_pyridine			fr_sulfone		
fr_quatN			fr_term_acetyl ene		
fr_sulfide			fr_tetrazole		
fr_sulfonamd			fr_thiazole		
fr_sulfone			fr_thiocyan		
fr_term_acetyl ene			fr_thiophene		
fr_tetrazole			fr_unbrch_alka ne		
fr_thiazole			fr_urea		
fr_thiocyan					
fr_thiophene					
fr_unbrch_alka ne					
fr_urea					

CHAPTER 6. CONCLUSIONS AND FUTURE WORK

Conclusions

This study aimed to elucidate the antimicrobial potential of lignin derivates as a practical strategy for valorizing lignin waste and to improve the economic viability of lignocellulosic biorefineries. Knowledge gaps into how lignin can be effectively incorporated into different materials, what specific lignin derivatives retain antimicrobial properties in materials and as depolymerization products, and which have increased activity were addressed in this study. Firstly, how different lignocellulosic components (lignin and hemicellulose) affect the formation and properties of physically cross-linked cellulose-based hydrogels was examined. This evaluation sought to understand if whole biomass-based hydrogels can be formed using ionic liquid dissolution and if lignin will retain its antimicrobial properties when incorporated into cellulose-based hydrogels. Due to the lack of research investigating the antimicrobial properties of lignin depolymerization products, we also explored the use of reductive and oxidative depolymerization methods to produce lignin-based bio-oils and tested their antimicrobial effects. The reduction process of catalytic transfer hydrogenolysis (CTH) was used to depolymerize lignin biorefinery waste into a phenolic rich bio-oil. The antimicrobial properties of this bio-oil and liquid-liquid extracted fractions were examined to better understand the antimicrobial potential of different lignin derivatives. Furthermore, an oxidative depolymerization strategy using peracetic acid was used on the same biorefinery lignin to create a bio-oil with antimicrobial applications in the fuel ethanol industry. Finally, quantitative structure−activity relationship (QSAR) models were developed to predict the antimicrobial

properties of lignin derivatives. This work provides critical knowledge and guidance on using lignin as an antimicrobial source in different industrial processes/products and to identify lignin derivatives with enhanced activity.

The potential of using ILs in the facile preparation of physically crosslinked lignocellulose-based hydrogels was evaluated and compared to a traditional chemical crosslinking method. Isolated kraft lignin and xylan were added to pure cellulose-based hydrogels and were found to improve the mechanical strength and stiffness compared to using just cellulose. However, the physically crosslinked hydrogels had less elastic strength than the chemical crosslinker method. Conversely, utilizing raw biomass for hydrogel formation provided increased mechanical strength (poplar) and similar water retention abilities (poplar and sorghum) when compared to the cellulose chemical crosslinker method. The kraft lignin-containing and poplar-based hydrogels provided significant antimicrobial properties against *E. coli*, illustrating the retention of lignin's antimicrobial properties when incorporated into hydrogels. Collectively, results from this study demonstrated the potential of using ILs to make physically crosslinked hydrogels directly from lignocellulosic biomass with increased mechanical and antimicrobial properties.

Purified alkali-enzymatic corn stover lignin (AEL) was depolymerized by catalytic transfer hydrogenolysis using supercritical ethanol and a Ru/C catalyst. The resulting bio-oil was produced at high conversion yields (~50 wt%) with a large number of monomers present (>30 wt%) in the form of alkylated phenols, hydrogenated hydroxycinnamic acid derivatives, syringol, and guaiacol-type lignins. Sequential extraction using hexane, petroleum ether, chloroform, and ethyl acetate extracted the raw bio-oil into five different fractions at 50.7-5.8 wt% yields of total bio-oil in the order of chloroform > hexane >

petroleum ether ≈ water > ethyl acetate. The hexane and petroleum ether SEF's were primarily composed of alkylated phenolics (i.e., syringyl propane, 4-ethylphenol, and 4-propylguiacol), while the more polar solvent (chloroform, ethyl acetate, and water) SEF's contained large amounts of phenolics with increased oxygenated functional groups (i.e., ethyl homovanillate and homovanillyl alcohol). The Molecular weights of the raw bio-oil and sequential extraction fractions (SEF) were much lower than the purified AEL, highlighting the depolymerization that occurred after CTH. The antimicrobial results suggested that the total monomer concentration and the presence of specific monomers (i.e., syringyl propane) may correlate to the antimicrobial activity and that cell death occurred due to membrane damage. However, the exact mode of action or antimicrobial activity caused by the synergism of the complex mixtures of monomers and unidentified oligomers remains unclear. This study provided insights into the types of lignin-derived compounds that confer antimicrobial activity and that compounds can be preferentially extracted from lignin bio-oils using simple LLE methods.

The same AEL lignin was depolymerized through oxidative procedures using peracetic acid. The resulting bio-oil showed a low degree of depolymerization into identifiable monomeric (<2%) with mostly unidentifiable lignin oligomers being produced. Nonetheless, this bio-oil displayed highly selective antimicrobial properties with up to 90% inhibition of commercially sampled LAB at 4 mg/ml and no inhibition of yeast. Thus, the larger oligomers produced after oxidative depolymerization are responsible for this selective activity. Based on fluorescent staining, the bio-oil's mode of action may be attributed to an ionophoric or membrane damaging activity that results in cell death. Furthermore, the bio-oil also showed no signs of inhibiting the hydrolytic activity of the

saccharification enzymes α-amylase or glucoamylase. Using the bio-oil (4 mg/ml) as an alternative antibiotic treatment during SSF of raw corn starch showed an increase (up to 8%) in ethanol production as bacterial contamination increased, compared to untreated contaminated controls. While the bio-oil did not completely inhibit lactobacillus growth, which still resulted in net losses of ethanol production (9%) compared to the uncontaminated control, the ability of the bio-oil to improve ethanol yields clearly show its efficacy as an alternative antibiotic. Taken together with the CTH study, these works show that different depolymerization methods of the same lignin will produce drastically different lignin monomers and oligomers with varying antimicrobial properties and applications.

The final study was the first attempt in the literature to predict the antimicrobial properties of lignin derivatives based on their phenolic structure. The open-access database ChEMBL was used to create three different datasets of compounds with MIC activity measurements against both *B. subtilis* and *E. coli*. The first datasets used all compounds, and then to increase their chemical similarity to lignin, compounds with only C, H, and O or phenolic based structures were utilized. These larger ChEMBL datasets employed machine learning algorithms to develop QSAR models that predict pMIC (-LogMIC) using selected molecular descriptors. These models were ultimately found to underpredict the antimicrobial activity (pMIC values) of actual lignin compounds found in a metanalysis of the literature. Next, more traditional QSAR models using ordinary least square regressions were created using the compounds in the metanalysis of the literature, which contained MIC data of lignin monomers against both *B. subtilis* and *E. coli*. An accurate QSAR model for *E. coli* was not found, but a satisfactory model was obtained for the *B. subtilis*

metanalysis dataset. The best performing model found MOE-type descriptors and the number of aliphatic carboxylic acid groups to be the best predictors of the lignin monomer's pMIC values (R^2 of 0.759). Finally, an additional dataset was experimentally derived by screening 25 lignin monomers and three dimers against *B. subtilis* by measuring the bacterial load difference (BLD). OLS regressions were also employed to develop the experimental datasets' QSAR models. The best OLS model for the experimental dataset found that MOE-type descriptors and the number of aromatic hydroxyl groups were better predictors of BLD (R^2 of 0.831). Thus, the smaller datasets highlighted how the variability in antimicrobial measurements and the specific compounds used will impact the predictive nature of the resulting QSARs.

Future Work

Although the lignocellulosic based hydrogels were found to have tunable properties, the use of ILs as a solvent adds extremely high costs to their production. Thus, s further work on utilizing lower-cost ILs or cheaper solvents such as deep eutectic solvents (DES) with the ability to dissolve all fractions of lignocellulose and whole biomass should be further evaluated. Additionally, only two biomass feedstocks were examined here, and based on their good performance/properties, future research should include a wider variety of feedstocks to create whole biomass-based hydrogels. Furthermore, analyses such as NMR should be utilized to examine the exact structures of lignin present in the hydrogels to correlate lignin structures to the observed antimicrobial properties.

The CTH derived bio-oils were shown to be effectively separated into different fractions with different lignin monomer composition using LLE. However, there were

significant losses in specific monomers due to the removal of the initial solvents after the drying stage, so different drying methods at lower temperature and under atmospheric pressure should be evaluated. While some correlations were made in relating total monomer content and even specific monomers with the antimicrobial properties of the bio-oil fractions, more work needs to be done. Specifically, future work should examine various mixtures of the identifiable monomers in the bio-oils, at low concentrations, to elucidate their antimicrobial interactions. With the absence of larger oligomers, these types of experiments could reveal the true importance of the unidentified oligomers and interactions of model monomers in determining their antimicrobial properties.

The peracetic acid oxidized lignin bio-oils showed very promising selective antimicrobial properties for reducing LAB contamination in fuel ethanol fermentation systems. Further research must improve the bio-oils solubility during fermentation as the bio-oil was seen to be in a solid form throughout the fermentation process. Thus, by improving its solubility using different solvent or surfactants, there could be an increase its antimicrobial action and resulting beneficial effects on ethanol production. Additionally, this work utilized pure corn starch instead of a traditional corn mash for the SSF experiments, and they were performed at extremely small scales (i.e. 30 ml). To determine the true efficacy of this bio-oil at industrial scales, a traditional corn mash and larger fermentation volumes should be evaluated.

The results from the modeling study strongly support that using QSARs to predict the antimicrobial properties of lignin-derived compounds must utilize a large and more comprehensive/cohesive dataset. For example, the results show that utilizing larger open-access databases (>10,000 non-lignin compounds) and machine learning algorithms could

not accurately predict lignin specific compounds. While smaller datasets using direct lignin derivates were compiled, with satisfactory/predictive QSARs developed, the low degree of variability amongst the compounds and use of different bioactivity measurements (i.e. BLD vs. MIC) in the datasets resulted in different descriptors being identified as main contributors to their activity. Therefore, future work should create a larger dataset (>100 compounds) with lignin only derivatives and a consistent bioactivity measurement to accurately predict the antimicrobial potential of future lignin depolymerization products. Additionally, model lignin dimers and even larger oligomers should be used in these models/datasets to help elucidate their antimicrobial potential, not just monomers.

REFERENCES

1.	Hadar, Y., *Sources for Lignocellulosic Raw Materials for the Production of Ethanol*. 2013. p. 21-38.
2.	Bajpai, P., *Structure of Lignocellulosic Biomass*. 2016.
3.	Canam, T., et al., *Pretreatment of Lignocellulosic Biomass Using Microorganisms: Approaches, Advantages, and Limitations*, in *Sustainable Degradation of Lignocellulosic Biomass - Techniques, Applications and Commercialization*. 2013.
4.	Chen, H., *Chemical Composition and Structure of Natural Lignocellulose*, in *Biotechnology of Lignocellulose: Theory and Practice*, H. Chen, Editor. 2014, Springer Netherlands: Dordrecht. p. 25-71.
5.	Brandt, A., et al., *Deconstruction of lignocellulosic biomass with ionic liquids.* Green Chemistry, 2013. **15**(3).
6.	Menon, V. and M. Rao, *Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery concept.* Progress in Energy and Combustion Science, 2012. **38**(4): p. 522-550.
7.	De Bhowmick, G., A.K. Sarmah, and R. Sen, *Lignocellulosic biorefinery as a model for sustainable development of biofuels and value added products.* Bioresource Technology, 2018. **247**: p. 1144-1154.
8.	Boerjan, W., J. Ralph, and M. Baucher, *Lignin biosynthesis.* Annual review of plant biology, 2003. **54**(1): p. 519-546.
9.	Pandey, M.P. and C.S. Kim, *Lignin Depolymerization and Conversion: A Review of Thermochemical Methods.* Chemical Engineering & Technology, 2011. **34**(1): p. 29-41.
10.	Li, C., et al., *Catalytic Transformation of Lignin for the Production of Chemicals and Fuels.* Chem Rev, 2015. **115**(21): p. 11559-624.
11.	Chen, F. and R.A. Dixon, *Lignin modification improves fermentable sugar yields for biofuel production.* Nature biotechnology, 2007. **25**(7): p. 759-761.
12.	Calvo-Flores, F.G., et al., *Applications of Modified and Unmodified Lignins*, in *Lignin and Lignans as Renewable Raw Materials*. 2015, John Wiley & Sons, Ltd. p. 247-288.
13.	Rauber, D., et al., *Electrochemical Lignin Degradation in Ionic Liquids on Ternary Mixed Metal Electrodes*, in *Zeitschrift für Physikalische Chemie*. 2018. p. 189.
14.	Pinto, P.C., D.V. Evtuguin, and C.P. Neto, *Structure of hardwood glucuronoxylans: modifications and impact on pulp retention during wood kraft pulping.* Carbohydrate Polymers, 2005. **60**(4): p. 489-497.
15.	Azadi, P., et al., *Liquid fuels, hydrogen and chemicals from lignin: A critical review.* Renewable and Sustainable Energy Reviews, 2013. **21**: p. 506-523.
16.	Hatakeyama, H. and T. Hatakeyama, *Lignin structure, properties, and applications*, in *Biopolymers*. 2009, Springer. p. 1-63.
17.	Schutyser, W., et al., *Catalysis in Lignocellulosic Biorefineries: The Case of Lignin Conversion: Applications in the Chemical Industry, Energy Development, and Environment Protection*. 2017. p. 537-584.
18.	Wang, H., et al., *From lignin to valuable products–strategies, challenges, and prospects.* Bioresource Technology, 2019. **271**: p. 449-461.

19.   Zakzeski, J., et al., *The Catalytic Valorization of Lignin for the Production of Renewable Chemicals.* Chemical Reviews, 2010. **110**(6): p. 3552-3599.

20.   Schutyser, W., et al., *Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading.* Chemical Society Reviews, 2018. **47**(3): p. 852-908.

21.   Lancefield, C.S., et al., *Pre-treatment of lignocellulosic feedstocks using biorenewable alcohols: towards complete biomass valorisation.* Green Chemistry, 2017. **19**(1): p. 202-214.

22.   Bhagia, S., et al., *Flowthrough pretreatment with very dilute acid provides insights into high lignin contribution to biomass recalcitrance.* Biotechnology for Biofuels, 2016. **9**(1): p. 245.

23.   Ibáñez, A.B. and S. Bauer, *Downscaled method using glass microfiber filters for the determination of Klason lignin and structural carbohydrates.* Biomass and Bioenergy, 2014. **68**: p. 75-81.

24.   Kim, J.S., Y.Y. Lee, and T.H. Kim, *A review on alkaline pretreatment technology for bioconversion of lignocellulosic biomass.* Bioresource Technology, 2016. **199**: p. 42-48.

25.   Jiang, B., et al., *Comparison of the Structural Characteristics of Cellulolytic Enzyme Lignin Preparations Isolated from Wheat Straw Stem and Leaf.* ACS Sustainable Chemistry & Engineering, 2017. **5**(1): p. 342-349.

26.   Hassan, S.S., G.A. Williams, and A.K. Jaiswal, *Emerging technologies for the pretreatment of lignocellulosic biomass.* Bioresource Technology, 2018. **262**: p. 310-318.

27.   Shrotri, A., H. Kobayashi, and A. Fukuoka, *Chapter Two - Catalytic Conversion of Structural Carbohydrates and Lignin to Chemicals*, in *Advances in Catalysis*, C. Song, Editor. 2017, Academic Press. p. 59-123.

28.   Gutiérrez-Macías, P., B. MontañezBarragán, and B. Barragan, *A review of agro-food waste transformation into feedstock for use in fermentation.* Fresenius Environmental Bulletin, 2015. **24**.

29.   *Pulp*, in *Ullmann's Encyclopedia of Industrial Chemistry*. p. 1-92.

30.   Collins, M.N., et al., *Valorization of lignin in polymer and composite systems for advanced engineering applications – A review.* International Journal of Biological Macromolecules, 2019. **131**: p. 828-849.

31.   Jung, Y.H. and K.H. Kim, *Chapter 3 - Acidic Pretreatment*, in *Pretreatment of Biomass*, A. Pandey, et al., Editors. 2015, Elsevier: Amsterdam. p. 27-50.

32.   He, Y., et al., *Lipid Production from Dilute Alkali Corn Stover Lignin by Rhodococcus Strains.* ACS Sustainable Chemistry & Engineering, 2017. **5**(3): p. 2302-2311.

33.   Zevallos Torres, L.A., et al., *Lignin as a potential source of high-added value compounds: A review.* Journal of Cleaner Production, 2020. **263**: p. 121499.

34.   Farrán, A., et al., *Green Solvents in Carbohydrate Chemistry: From Raw Materials to Fine Chemicals.* Chemical Reviews, 2015. **115**(14): p. 6811-6853.

35.   Plechkova, N.V. and K.R. Seddon, *Ionic liquids:"designer" solvents for green chemistry.* Methods and Reagents for Green Chemistry, 2007: p. 105-130.

36. Elgharbawy, A.A., et al., *Ionic liquid pretreatment as emerging approaches for enhanced enzymatic hydrolysis of lignocellulosic biomass.* Biochemical Engineering Journal, 2016. **109**: p. 252-267.

37. Tao, L., et al., *Process and technoeconomic analysis of leading pretreatment technologies for lignocellulosic ethanol production using switchgrass.* Bioresour Technol, 2011. **102**(24): p. 11105-14.

38. Satlewal, A., et al., *Natural deep eutectic solvents for lignocellulosic biomass pretreatment: Recent developments, challenges and novel opportunities.* Biotechnology Advances, 2018. **36**(8): p. 2032-2050.

39. Agbor, V.B., et al., *Biomass pretreatment: fundamentals toward application.* Biotechnol Adv, 2011. **29**(6): p. 675-85.

40. Alvira, P., et al., *Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review.* Bioresource technology, 2010. **101**(13): p. 4851-4861.

41. Conde-Mejía, C., A. Jiménez-Gutiérrez, and M. El-Halwagi, *A comparison of pretreatment methods for bioethanol production from lignocellulosic materials.* Process Safety and Environmental Protection, 2012. **90**(3): p. 189-202.

42. Sarkar, N., et al., *Bioethanol production from agricultural wastes: An overview.* Renewable Energy, 2012. **37**(1): p. 19-27.

43. Kumari, D. and R. Singh, *Pretreatment of lignocellulosic wastes for biofuel production: A critical review.* Renewable and Sustainable Energy Reviews, 2018. **90**: p. 877-891.

44. Mathew, A.K., et al., *Chapter 9 - Lignocellulosic Biorefinery Wastes, or Resources?*, in *Waste Biorefinery*, T. Bhaskar, et al., Editors. 2018, Elsevier. p. 267-297.

45. Sun, Z., et al., *Bright Side of Lignin Depolymerization: Toward New Platform Chemicals.* Chemical Reviews, 2018. **118**(2): p. 614-678.

46. Wang, H., M. Tucker, and Y. Ji, *Recent Development in Chemical Depolymerization of Lignin: A Review.* Journal of Applied Chemistry, 2013. **2013**: p. 9.

47. Huber, G.W., S. Iborra, and A. Corma, *Synthesis of Transportation Fuels from Biomass: Chemistry, Catalysts, and Engineering.* Chemical Reviews, 2006. **106**(9): p. 4044-4098.

48. Yang, H., et al., *Characteristics of hemicellulose, cellulose and lignin pyrolysis.* Fuel, 2007. **86**(12-13): p. 1781-1788.

49. Kawamoto, H., *Lignin pyrolysis reactions.* Journal of Wood Science, 2017. **63**(2): p. 117-132.

50. Lavoie, J.-M., W. Baré, and M. Bilodeau, *Depolymerization of steam-treated lignin for the production of green chemicals.* Bioresource Technology, 2011. **102**(7): p. 4917-4920.

51. Roberts, V.M., et al., *Towards Quantitative Catalytic Lignin Depolymerization.* Chemistry – A European Journal, 2011. **17**(21): p. 5939-5948.

52. Beauchet, R., F. Monteil-Rivera, and J.M. Lavoie, *Conversion of lignin to aromatic-based chemicals (L-chems) and biofuels (L-fuels).* Bioresource Technology, 2012. **121**: p. 328-334.

53.   Gasson, J.R., et al., *Modeling the Lignin Degradation Kinetics in an Ethanol/Formic Acid Solvolysis Approach. Part 1. Kinetic Model Development.* Industrial & Engineering Chemistry Research, 2012. **51**(32): p. 10595-10606.

54.   Forchheim, D., et al., *Modeling the Lignin Degradation Kinetics in a Ethanol/Formic Acid Solvolysis Approach. Part 2. Validation and Transfer to Variable Conditions.* Industrial & Engineering Chemistry Research, 2012. **51**(46): p. 15053-15063.

55.   Xu, W., et al., *Depolymerization and Hydrodeoxygenation of Switchgrass Lignin with Formic Acid.* ChemSusChem, 2012. **5**(4): p. 667-675.

56.   Liguori, L. and T. Barth, *Palladium-Nafion SAC-13 catalysed depolymerisation of lignin to phenols in formic acid and water.* Journal of Analytical and Applied Pyrolysis, 2011. **92**(2): p. 477-484.

57.   Song, Q., et al., *Lignin depolymerization (LDP) in alcohol over nickel-based catalysts via a fragmentation–hydrogenolysis process.* Energy & Environmental Science, 2013. **6**(3): p. 994-1007.

58.   Das, L., et al., *Characterization and Catalytic Transfer Hydrogenolysis of Deep Eutectic Solvent Extracted Sorghum Lignin to Phenolic Compounds.* ACS Sustainable Chemistry & Engineering, 2018. **6**(8): p. 10408-10420.

59.   Zhou, M., et al., *Catalytic in Situ Hydrogenolysis of Lignin in Supercritical Ethanol: Effect of Phenol, Catalysts, and Reaction Temperature.* ACS Sustainable Chemistry & Engineering, 2018. **6**(5): p. 6867-6875.

60.   Xu, C., et al., *Lignin depolymerisation strategies: towards valuable chemicals and fuels.* Chemical Society Reviews, 2014. **43**(22): p. 7485-7500.

61.   Barta, K., et al., *Catalytic disassembly of an organosolv lignin via hydrogen transfer from supercritical methanol.* Green Chemistry, 2010. **12**(9): p. 1640-1647.

62.   Zhang, J., *Catalytic transfer hydrogenolysis as an efficient route in cleavage of lignin and model compounds.* Green Energy & Environment, 2018. **3**(4): p. 328-334.

63.   Ma, R., et al., *Peracetic Acid Depolymerization of Biorefinery Lignin for Production of Selective Monomeric Phenolic Compounds.* Chemistry – A European Journal, 2016. **22**(31): p. 10884-10891.

64.   Ma, R., M. Guo, and X. Zhang, *Recent advances in oxidative valorization of lignin.* Catalysis Today, 2018. **302**: p. 50-60.

65.   Swern, D., *Organic Peracids.* Chemical Reviews, 1949. **45**(1): p. 1-68.

66.   Ren, S., X.P. Ye, and A.P. Borole, *Separation of chemical groups from bio-oil water-extract via sequential organic solvent extraction.* Journal of Analytical and Applied Pyrolysis, 2017. **123**: p. 30-39.

67.   Wei, Y., et al., *Liquid–Liquid Extraction of Biomass Pyrolysis Bio-oil.* Energy & Fuels, 2014. **28**(2): p. 1207-1212.

68.   Matsumoto, M., Y. Inomoto, and K. Kondo, *Selective separation of aromatic hydrocarbons through supported liquid membranes based on ionic liquids.* Journal of Membrane Science, 2005. **246**(1): p. 77-81.

69.   Aghazadeh, M. and A.S. Engelberth, *Techno-economic analysis for incorporating a liquid-liquid extraction system to remove acetic acid into a proposed commercial scale biorefinery.* Biotechnol Prog, 2016. **32**(4): p. 971-7.

70.   Miyamoto, K., et al., *High-Efficiency Liquid Chromatographic Separation Utilizing Long Monolithic Silica Capillary Columns.* Analytical Chemistry, 2008. **80**(22): p. 8741-8750.

71.   Fele Žilnik, L. and A. Jazbinšek, *Recovery of renewable phenolic fraction from pyrolysis oil.* Separation and Purification Technology, 2012. **86**: p. 157-170.

72.   Mahfud, F.H., et al., *Acetic Acid Recovery from Fast Pyrolysis Oil. An Exploratory Study on Liquid-Liquid Reactive Extraction using Aliphatic Tertiary Amines.* Separation Science and Technology, 2008. **43**(11-12): p. 3056-3074.

73.   Zeng, F., et al., *Separation of phthalate esters from bio-oil derived from rice husk by a basification–acidification process and column chromatography.* Bioresource Technology, 2011. **102**(2): p. 1982-1987.

74.   Mantilla, S.V., A.M. Manrique, and P. Gauthier-Maradei, *Methodology for Extraction of Phenolic Compounds of Bio-oil from Agricultural Biomass Wastes.* Waste and Biomass Valorization, 2015. **6**(3): p. 371-383.

75.   Li, J., C. Wang, and Z. Yang, *Production and separation of phenols from biomass-derived bio-petroleum.* Journal of Analytical and Applied Pyrolysis, 2010. **89**(2): p. 218-224.

76.   Wang, S., et al., *Multi-step separation of monophenols and pyrolytic lignins from the water-insoluble phase of bio-oil.* Separation and Purification Technology, 2014. **122**: p. 248–255.

77.   Meier, D., R. Ante, and O. Faix, *Catalytic hydropyrolysis of lignin: Influence of reaction conditions on the formation and composition of liquid products.* Bioresource Technology, 1992. **40**(2): p. 171-177.

78.   Maggi, R. and B. Delmon, *Comparison between 'slow' and 'flash' pyrolysis oils from biomass.* Fuel, 1994. **73**(5): p. 671-677.

79.   Mante, O.D., et al., *Isolation and Purification of Monofunctional Methoxyphenols from Loblolly Pine Biocrude.* ACS Sustainable Chemistry & Engineering, 2019. **7**(2): p. 2262-2269.

80.   Mante, O.D., et al., *A selective extraction method for recovery of monofunctional methoxyphenols from biomass pyrolysis liquids.* Green Chemistry, 2019. **21**(9): p. 2257-2265.

81.   Organization, W.H., *Antimicrobial resistance: global report on surveillance.* 2014: World Health Organization.

82.   Harvey, A.L., R. Edrada-Ebel, and R.J. Quinn, *The re-emergence of natural products for drug discovery in the genomics era.* Nature reviews drug discovery, 2015. **14**(2): p. 111-129.

83.   Upadhyay, A., et al., *Combating Pathogenic Microorganisms Using Plant-Derived Antimicrobials: A Minireview of the Mechanistic Basis.* BioMed Research International, 2014. **2014**.

84.   Terzioğlu, P., F.N. Parın, and Y. Sıcak, *Lignin Composites for Biomedical Applications: Status, Challenges and Perspectives*, in *Lignin: Biosynthesis and Transformation for Industrial Applications*, S. Sharma and A. Kumar, Editors. 2020, Springer International Publishing: Cham. p. 253-273.

85.   Chandna, S., et al., *Synthesis and Applications of Lignin-Derived Hydrogels*, in *Lignin: Biosynthesis and Transformation for Industrial Applications*, S. Sharma and A. Kumar, Editors. 2020, Springer International Publishing: Cham. p. 231-252.

86.   Thakur, V.K. and M.K. Thakur, *Recent advances in green hydrogels from lignin: a review.* International Journal of Biological Macromolecules, 2015. **72**: p. 834-847.

87.   Spasojević, D., et al., *Lignin model compound in alginate hydrogel: a strong antimicrobial agent with high potential in wound treatment.* International Journal of Antimicrobial Agents, 2016. **48**(6): p. 732-735.

88.   Zmejkoski, D., et al., *Bacterial cellulose-lignin composite hydrogel as a promising agent in chronic wound healing.* International journal of biological macromolecules, 2018. **118 Pt A**: p. 494-503.

89.   Aadil, K.R., S.I. Mussatto, and H. Jha, *Synthesis and characterization of silver nanoparticles loaded poly(vinyl alcohol)-lignin electrospun nanofibers and their antimicrobial activity.* International Journal of Biological Macromolecules, 2018. **120**: p. 763-767.

90.   Kim, S., et al., *Chitosan–lignosulfonates sono-chemically prepared nanoparticles: Characterisation and potential applications.* Colloids and Surfaces B: Biointerfaces, 2013. **103**: p. 1-8.

91.   Lee, E., Y. Song, and S. Lee, *Crosslinking of lignin/poly(vinyl alcohol) nanocomposite fiber webs and their antimicrobial and ultraviolet-protective properties.* Textile Research Journal, 2017. **89**(1): p. 3-12.

92.   Lee, E.-S., et al., *Antimicrobial properties of lignin-decorated thin multi-walled carbon nanotubes in poly(vinyl alcohol) nanocomposites.* European Polymer Journal, 2018. **105**: p. 79-84.

93.   Mehta, M.J. and A. Kumar, *Ionic Liquid Stabilized Gelatin–Lignin Films: A Potential UV-Shielding Material with Excellent Mechanical and Antimicrobial Properties.* Chemistry – A European Journal, 2019. **25**(5): p. 1269-1274.

94.   Yang, W., et al., *Polyvinyl alcohol/chitosan hydrogels with enhanced antioxidant and antibacterial properties induced by lignin nanoparticles.* Carbohydrate Polymers, 2018. **181**: p. 275-284.

95.   Jaganathan, G., et al., *Fabrication and characterization of Artocarpus heterophyllus waste derived lignin added chitosan biocomposites for wound dressing application.* Sustainable Chemistry and Pharmacy, 2018. **10**: p. 27-32.

96.   Gabov, K., et al., *Preparation, characterization and antimicrobial application of hybrid cellulose-lignin beads.* Cellulose, 2017. **24**(2): p. 641-658.

97.   Freitas, F.M.C., et al., *Green synthesis of lignin nano- and micro-particles: Physicochemical characterization, bioactive properties and cytotoxicity assessment.* International Journal of Biological Macromolecules, 2020. **163**: p. 1798-1809.

98.   Yang, W., et al., *Effect of cellulose and lignin on disintegration, antimicrobial and antioxidant properties of PLA active films.* International Journal of Biological Macromolecules, 2016. **89**: p. 360-368.

99.   Yang, W., et al., *Valorization of Acid Isolated High Yield Lignin Nanoparticles as Innovative Antioxidant/Antimicrobial Organic Materials.* ACS Sustainable Chemistry & Engineering, 2018. **6**(3): p. 3502-3514.

100.   Pal, K., A.K. Banthia, and D.K. Majumdar, *Polymeric Hydrogels: Characterization and Biomedical Applications.* Designed Monomers and Polymers, 2009. **12**(3): p. 197-220.

101. Kayra, N. and A.Ö. Aytekin, *Synthesis of Cellulose-Based Hydrogels: Preparation, Formation, Mixture, and Modification*, in *Cellulose-Based Superabsorbent Hydrogels*, M.I.H. Mondal, Editor. 2018, Springer International Publishing: Cham. p. 1-28.

102. Ahmed, E.M., *Hydrogel: Preparation, characterization, and applications: A review.* Journal of Advanced Research, 2015. **6**: p. 105-121.

103. Zohuriaan-Mehr, M.J., et al., *Protein- and homo poly(amino acid)-based hydrogels with super-swelling properties.* Polymers for Advanced Technologies, 2009. **20**(8): p. 655-671.

104. Bhatnagar, A., et al., *Hydrogels:A Boon for Increasing Agricultural Productivity in Water-Stressed Environment.* Current Science, 2016. **111**(11).

105. Yue, M., et al., *Temperature-Responsive Microgel Films as Reversible Carbon Dioxide Absorbents in Wet Environment.* Angewandte Chemie, 2014. **126**(10): p. 2692-2695.

106. Caló, E. and V.V. Khutoryanskiy, *Biomedical applications of hydrogels: A review of patents and commercial products.* European Polymer Journal, 2015. **65**: p. 252-267.

107. Kamoun, E.A., E.S. Kenawy, and X. Chen, *A review on polymeric hydrogel membranes for wound dressing applications: PVA-based hydrogel dressings.* J Adv Res, 2017. **8**(3): p. 217-233.

108. Utech, S., *A review of hydrogel-based composites for biomedical applications: enhancement of hydrogel properties by addition of rigid inorganic fillers.* Journal of materials science, 2016. **v. 51**(no. 1): p. pp. 271-310-2016 v.51 no.1.

109. Ullah, F., et al., *Classification, processing and application of hydrogels: A review.* Materials Science and Engineering: C, 2015. **57**: p. 414-433.

110. Larrañeta, E., et al., *Synthesis and Characterization of Lignin Hydrogels for Potential Applications as Drug Eluting Antimicrobial Coatings for Medical Materials.* ACS sustainable chemistry & engineering, 2018. **6**(7): p. 9037-9046.

111. Yang, W., et al., *Antioxidant and antibacterial lignin nanoparticles in polyvinyl alcohol/chitosan films for active packaging.* Industrial Crops and Products, 2016. **94**: p. 800-811.

112. Sláviková, E. and B. Košíková, *Inhibitory effect of lignin by-products of pulping on yeast growth.* Folia Microbiologica, 1994. **39**(3): p. 241-243.

113. Dizhbite, T., et al., *Characterization of the radical scavenging activity of lignins—natural antioxidants.* Bioresource Technology, 2004. **95**(3): p. 309-317.

114. Baurhoo, B., C.A. Ruiz-Feria, and X. Zhao, *Purified lignin: Nutritional and health impacts on farm animals—A review.* Animal Feed Science and Technology, 2008. **144**(3): p. 175-184.

115. Nada, A.M.A., A.I. El-Diwany, and A.M. Elshafei, *Infrared and antimicrobial studies on different lignins.* Acta Biotechnologica, 1989. **9**(3): p. 295-298.

116. Özgen, C., *Isolation of antimicrobial molecules from agricultural biomass and utilization in xylan-based biodegradable films.* Middle East Technical University, 2010.

117. Kaur, R., S.K. Uppal, and P. Sharma, *Antioxidant and Antibacterial Activities of Sugarcane Bagasse Lignin and Chemically Modified Lignins.* Sugar Tech, 2017. **19**(6): p. 675-680.

118. Coral Medina, J.D., et al., *Biological activities and thermal behavior of lignin from oil palm empty fruit bunches as potential source of chemicals of added value.* Industrial Crops and Products, 2016. **94**: p. 630-637.

119. Dong, X., et al., *Antimicrobial and antioxidant activities of lignin from residue of corn stover to ethanol production.* Industrial Crops and Products, 2011. **34**(3): p. 1629-1634.

120. García, A., G. Spigno, and J. Labidi, *Antioxidant and biocide behaviour of lignin fractions from apple tree pruning residues.* Industrial Crops and Products, 2017. **104**: p. 242-252.

121. Wang, G., et al., *Successive ethanol–water fractionation of enzymatic hydrolysis lignin to concentrate its antimicrobial activity.* Journal of Chemical Technology & Biotechnology, 2018. **93**(10): p. 2977-2987.

122. Cazacu, G., M. Capraru, and V.I. Popa, *Advances Concerning Lignin Utilization in New Materials*, in *Advances in Natural Polymers*. 2013, Springer, Berlin, Heidelberg. p. 255-312.

123. Zemek, J., et al., *Antibiotic properties of lignin components.* Folia Microbiologica, 1979. **24**(6): p. 483-486.

124. Barber, M.S., V.S. McConnell, and B.S. DeCaux, *Antimicrobial intermediates of the general phenylpropanoid and lignin specific pathways.* Phytochemistry, 2000. **54**(1): p. 53-6.

125. Greenberg, M., M. Dodds, and M. Tian, *Naturally Occurring Phenolic Antibacterial Compounds Show Effectiveness against Oral Bacteria by a Quantitative Structure−Activity Relationship Study.* Journal of Agricultural and Food Chemistry, 2008. **56**(23): p. 11151-11156.

126. Patra, J.K., et al., *Antibacterial Effects of Pyrolysis Oil Against Salmonella Typhimurium and Escherichia coli.* Foodborne Pathogens and Disease, 2015. **13**(1): p. 13-20.

127. Hwang, Y.H., et al., *Antimicrobial effect of the wood vinegar from Cryptomeria japonica sapwood on plant pathogenic microorganisms*. Vol. 15. 2005. 1106-1109.

128. de Souza Araujo, E., et al., *Antibacterial and antifungal activities of pyroligneous acid from wood of Eucalyptus urograndis and Mimosa tenuiflora.* J Appl Microbiol, 2018. **124**(1): p. 85-96.

129. Hou, X., et al., *Chemical constituents and antimicrobial activity of wood vinegars at different pyrolysis temperature ranges obtained from Eucommia ulmoides Olivers branches.* RSC Advances, 2018. **8**(71): p. 40941-40949.

130. Milly, P.J., *Antimicrobial properties of liquid smoke fractions*. 2003, University of Georgia.

131. Bedmutha, R., et al., *Insecticidal and bactericidal characteristics of the bio-oil from the fast pyrolysis of coffee grounds.* Journal of Analytical and Applied Pyrolysis, 2011. **90**(2): p. 224-231.

132. Mohan, D., et al., *Fungicidal values of bio-oils and their lignin-rich fractions obtained from wood/bark fast pyrolysis.* Chemosphere, 2008. **71**(3): p. 456-465.

133. Sari, E., et al., *Production of Liquid Smoke From the Process of Carbonization of Durian Skin Biomass, Coconut Shell and Palm Shell for Preservation of Tilapia Fish.* IOP Conference Series: Materials Science and Engineering, 2019. **543**: p. 012075.

134. Lourençon, T.V., et al., *Bio-oil from a fast pyrolysis pilot plant as antifungal and hydrophobic agent for wood preservation.* Journal of Analytical and Applied Pyrolysis, 2016. **122**: p. 1-6.

135. Patra, J.K., et al., *Bactericidal Mechanism of Bio-oil Obtained from Fast Pyrolysis of Pinus densiflora Against Two Foodborne Pathogens, Bacillus cereus and Listeria monocytogenes.* Foodborne Pathog Dis, 2015. **12**(6): p. 529-35.

136. Hossain, M.M., et al., *Insecticidal and anti-microbial activity of bio-oil derived from fast pyrolysis of lignin, cellulose, and hemicellulose.* Journal of Pest Science, 2015. **88**(1): p. 171-179.

137. Ngarmsak, M., et al., *Antimicrobial activity of vanillin against spoilage microorganisms in stored fresh-cut mangoes.* J Food Prot, 2006. **69**(7): p. 1724-7.

138. Xue, J., P.M. Davidson, and Q. Zhong, *Thymol nanoemulsified by whey protein-maltodextrin conjugates: the enhanced emulsifying capacity and antilisterial properties in milk by propylene glycol.* Journal of agricultural and food chemistry, 2013. **61**(51): p. 12720-12726.

139. Bittner, S., *When quinones meet amino acids: chemical, physical and biological consequences.* Amino Acids, 2006. **30**(3): p. 205-24.

140. Borges, A., et al., *Antibacterial activity and mode of action of ferulic and gallic acids against pathogenic bacteria.* Microb Drug Resist, 2013. **19**(4): p. 256-65.

141. Maddox, C.E., L.M. Laur, and L. Tian, *Antibacterial activity of phenolic compounds against the phytopathogen Xylella fastidiosa.* Curr Microbiol, 2010. **60**(1): p. 53-8.

142. Muthuswamy, S. and H.V. Rupasinghe, *Fruit phenolics as natural antimicrobial agents: Selective antimicrobial activity of catechin, chlorogenic acid and phloridzin.* 2007.

143. Nohynek, L.J., et al., *Berry phenolics: antimicrobial properties and mechanisms of action against severe human pathogens.* Nutr Cancer, 2006. **54**(1): p. 18-32.

144. Cushnie, T.T. and A.J. Lamb, *Antimicrobial activity of flavonoids.* International journal of antimicrobial agents, 2005. **26**(5): p. 343-356.

145. Helander, I.M., et al., *Characterization of the Action of Selected Essential Oil Components on Gram-Negative Bacteria.* Journal of Agricultural and Food Chemistry, 1998. **46**(9): p. 3590-3595.

146. Puupponen-Pimiä, R., et al., *Bioactive berry compounds-novel tools against human pathogens.* Appl Microbiol Biotechnol, 2005. **67**(1): p. 8-18.

147. Kubo, I., et al., *Anti-Salmonella activity of (2E)-alkenals.* J Appl Microbiol, 2004. **96**(4): p. 693-9.

148. Kubo, I., H. Muroi, and A. Kubo, *Structural functions of antimicrobial long-chain alcohols and phenols.* Bioorganic & Medicinal Chemistry, 1995. **3**(7): p. 873-880.

149. Devi, K.P., et al., *Eugenol (an essential oil of clove) acts as an antibacterial agent against Salmonella typhi by disrupting the cellular membrane.* J Ethnopharmacol, 2010. **130**(1): p. 107-15.

150. Salton, M.R., *Studies of the bacterial cell wall. IV. The composition of the cell walls of some Gram-positive and Gram-negative bacteria.* Biochim Biophys Acta, 1953. **10**(4): p. 512-23.

151. Campos, F.M., J.A. Couto, and T.A. Hogg, *Influence of phenolic acids on growth and inactivation of Oenococcus oeni and Lactobacillus hilgardii.* J Appl Microbiol, 2003. **94**(2): p. 167-74.

152. Rohn, S., H.M. Rawel, and J. Kroll, *Inhibitory Effects of Plant Phenols on the Activity of Selected Enzymes.* Journal of Agricultural and Food Chemistry, 2002. **50**(12): p. 3566-3571.

153. Ultee, A., M.H.J. Bennik, and R. Moezelaar, *The phenolic hydroxyl group of carvacrol is essential for action against the food-borne pathogen Bacillus cereus.* Applied and environmental microbiology, 2002. **68**(4): p. 1561-1568.

154. Shahlaei, M., *Descriptor Selection Methods in Quantitative Structure–Activity Relationship Studies: A Review Study.* Chemical Reviews, 2013. **113**(10): p. 8093-8103.

155. Griffin, S.G., S.G. Wyllie, and J.L. Markham, *Antimicrobially Active Terpenes Cause K+ Leakage in E. coli Cells.* Journal of Essential Oil Research, 2005. **17**(6): p. 686-690.

156. Dorman, H.J. and S.G. Deans, *Antimicrobial agents from plants: antibacterial activity of plant volatile oils.* J Appl Microbiol, 2000. **88**(2): p. 308-16.

157. Moleyar, V. and P. Narasimham, *Antifungal activity of some essential oil components.* Food Microbiology, 1986. **3**(4): p. 331-336.

158. Oyedemi, S., et al., *The proposed mechanism of bactericidal action of eugenol, A-terpineol and Y-terpinene against Listeria monocytogenes, Streptococcus pyogenes, Proteus vulgaris and Escherichia coli.* African Journal of Biotechnology, 2009. **8**: p. 1280-1286.

159. Larif, M., et al., *Potential effluent oil mills and antibacterial activity polyphenols against some pathogenic strains.* Research on Chemical Intermediates, 2015. **41**(2): p. 1213-1225.

160. Bouarab-Chibane, L., et al., *Antibacterial Properties of Polyphenols: Characterization and QSAR (Quantitative Structure–Activity Relationship) Models.* Frontiers in Microbiology, 2019. **10**(829).

161. Akhtar, M.F., M. Hanif, and N.M. Ranjha, *Methods of synthesis of hydrogels … A review.* Saudi Pharmaceutical Journal.

162. Isikgor, F.H. and C.R. Becer, *Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers.* Polymer Chemistry, 2015. **6**(25): p. 4497-4559.

163. Li, X. and X. Pan, *Hydrogels Based on Hemicellulose and Lignin from Lignocellulose Biorefinery: a Mini-review.* Journal of Biobased Materials and Bioenergy, 2010. **4**: p. 289-297.

164. Liang, X., et al., *Preparation of cellulose-based conductive hydrogels with ionic liquid.* Reactive and Functional Polymers, 2015. **86**: p. 1-6.

165. Lopez-Sanchez, P., et al., *Poroelastic mechanical effects of hemicelluloses on cellulosic hydrogels under compression.* PLoS One, 2015. **10**(3): p. e0122132.

166. *<Ciolacu et al. - 2013 - Morphological and surface aspects of cellulose-lig.pdf>*.

167. Ciolacu, D., et al., *Morphological and surface aspects of cellulose-lignin hydrogels.* Cellulose Chemistry and Technology, 2013. **47**: p. 377-386.

168. Ciolacu, D., et al., *New cellulose–lignin hydrogels and their application in controlled release of polyphenols.* Materials Science and Engineering: C, 2012. **32**(3): p. 452-463.

169. Park, S., et al., *Application of cellulose/lignin hydrogel beads as novel supports for immobilizing lipase.* Journal of Molecular Catalysis B: Enzymatic, 2015. **119**(Supplement C): p. 33-39.

170. Telysheva, G., et al., *Lignin products for decontamination of environment objects from pathogenic microorganisms and pollutants.* Proceedings of the 7th ILI Forum, Barcelona, Spain, 2005: p. 71-74.

171. Bakker, E.P., *Ionophore Antibiotics*, in *Mechanism of Action of Antibacterial Agents*. 1979, Springer, Berlin, Heidelberg. p. 67-97.

172. Russell, J.B. and A.J. Houlihan, *Ionophore resistance of ruminal bacteria and its potential impact on human health.* FEMS microbiology reviews, 2003. **27**(1): p. 65-74.

173. Ibrahim, M.M., M. Abd-Eladl, and N.H. Abou-Baker, *Lignocellulosic biomass for the preparation of cellulose-based hydrogel and its use for optimizing water resources in agriculture.* Journal of Applied Polymer Science, 2015. **132**: p. n/a-n/a.

174. Wu, X.-L., et al., *Biomass-Derived Sponge-like Carbonaceous Hydrogels and Aerogels for Supercapacitors.* ACS Nano, 2013. **7**(4): p. 3589-3597.

175. Zhang, M., et al., *Biomass based hydrogel as an adsorbent for the fast removal of heavy metal ions from aqueous solutions.* Journal of Materials Chemistry A, 2017. **5**(7): p. 3434-3446.

176. Puoci, F., et al., *Polymer in Agriculture: a Review.* Vol. 3. 2008.

177. Yang, K., et al., *Antimicrobial hydrogels: promising materials for medical application.* Int J Nanomedicine, 2018. **13**: p. 2217-2263.

178. Alvarez-Vasco, C., et al., *Unique low-molecular-weight lignin with high purity extracted from wood by deep eutectic solvents (DES): a source of lignin for valorization.* Green Chemistry, 2016. **18**(19): p. 5133-5141.

179. Chang, C. and L. Zhang, *Cellulose-based hydrogels: Present status and application prospects.* Carbohydrate Polymers, 2011. **84**: p. 40-53.

180. Selig, M., N. Weiss, and Y. Ji, *Enzymatic saccharification of lignocellulosic biomass. NREL Laboratory Analytical Procedure.* 2008: Technical Report NREL/TP-510-42629, NREL, Colorado, USA. http://www. nrel. gov/docs/gen/fy08/42629. pdf.

181. C. Segal, L., et al., *An Empirical Method for Estimating the Degree of Crystallinity of Native Cellulose Using the X-Ray Diffractometer.* Textile Research Journal - TEXT RES J, 1959. **29**: p. 786-794.

182. Nakayama, A., et al., *High Mechanical Strength Double-Network Hydrogel with Bacterial Cellulose.* Advanced Functional Materials, 2004. **14**(11): p. 1124-1128.

183. Liu, E., et al., *Impact of Dilute Sulfuric Acid, Ammonium Hydroxide, and Ionic Liquid Pretreatments on the Fractionation and Characterization of Engineered Switchgrass.* BioEnergy Research, 2017. **10**(4): p. 1079-1093.

184. Sun, N., et al., *Understanding pretreatment efficacy of four cholinium and imidazolium ionic liquids by chemistry and computation.* Green Chemistry, 2014. **16**(5): p. 2546-2557.

185. Abbott, A.P., et al., *Deep Eutectic Solvents Formed between Choline Chloride and Carboxylic Acids: Versatile Alternatives to Ionic Liquids.* Journal of the American Chemical Society, 2004. **126**(29): p. 9142-9147.

186. Francisco, M., A.v.d. Bruinhorst, and M.C. Kroon, *New natural and renewable low transition temperature mixtures (LTTMs): screening as solvents for lignocellulosic biomass processing.* Green Chemistry, 2012. **14**(8): p. 2153-2157.

187. Xu, G.C., et al., *Enhancing cellulose accessibility of corn stover by deep eutectic solvent pretreatment for butanol fermentation.* Bioresour Technol, 2016. **203**: p. 364-9.

188. Zhang, Q., et al., *Deep eutectic solvents: syntheses, properties and applications.* Chemical Society Reviews, 2012. **41**(21): p. 7108-7146.

189. Chen, Z. and C. Wan, *Ultrafast fractionation of lignocellulosic biomass by microwave-assisted deep eutectic solvent pretreatment.* Bioresour Technol, 2018. **250**: p. 532-537.

190. Osch, D.J.G.P.v., et al., *Ionic liquids and deep eutectic solvents for lignocellulosic biomass fractionation.* Physical Chemistry Chemical Physics, 2017. **19**(4): p. 2636-2665.

191. Das, L., et al., *Industrial hemp as a potential bioenergy crop in comparison with kenaf, switchgrass and biomass sorghum.* Bioresource Technology, 2017. **244**: p. 641-649.

192. Tan, X., et al., *Solubility of starch and microcrystalline cellulose in 1-ethyl-3-methylimidazolium acetate ionic liquid and solution rheological properties.* Physical Chemistry Chemical Physics, 2016. **18**(39): p. 27584-27593.

193. Chang, C., A. Lue, and L. Zhang, *Effects of Crosslinking Methods on Structure and Properties of Cellulose/PVA Hydrogels.* Macromolecular Chemistry and Physics, 2008. **209**(12): p. 1266-1273.

194. Park, S., et al., *Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance.* Biotechnology for Biofuels, 2010. **3**: p. 10.

195. Petti, C., et al., *Sorghum mutant RGdisplays antithetic leaf shoot lignin accumulation resulting in improved stem saccharification properties.* Biotechnology for Biofuels, 2013. **6**: p. 146.

196. Qadri, S.B., et al., *Silicon Carbide Synthesis.* 2014, The Government Of The United States Of America, As Represented By The Secretary Of The Navy.

197. Bodenberger, N., et al., *Evaluation of methods for pore generation and their influence on physio-chemical properties of a protein based hydrogel.* Biotechnol Rep (Amst), 2016. **12**: p. 6-12.

198. Ivanov, C., et al., *Synthesis of poly(vinyl alcohol) - Methyl cellulose hydrogel as possible scaffolds in tissue engineering.* Journal of Optoelectronics and Advanced Materials, 2007. **9**: p. 3440-3444.

199. Jiang, G., *Tensile Mechanical Properties of Hydrophobic Association Hydrogels: Effect of Crosslinking Method, Synthesis Temperature, Mineral Salt and Swelling.* Journal of Macromolecular Science, Part A, 2014. **51**(2): p. 165-172.

200. Nakasone, K. and T. Kobayashi, *Cytocompatible cellulose hydrogels containing trace lignin.* Materials Science and Engineering: C, 2016. **64**(Supplement C): p. 269-277.

201. Muniz, E.C. and G. Geuskens, *Compressive Elastic Modulus of Polyacrylamide Hydrogels and Semi-IPNs with Poly(N-isopropylacrylamide).* Macromolecules, 2001. **34**(13): p. 4480-4484.

202. Rudzinski, W.E., et al., *Hydrogels as controlled release devices in agriculture.* Designed Monomers and Polymers, 2002. **5**(1): p. 39-65.

203. Chang, C., et al., *Structure and properties of hydrogels prepared from cellulose in NaOH/urea aqueous solutions.* Carbohydrate Polymers, 2010. **82**(1): p. 122-127.

204. Varanasi, P., et al., *Understanding changes in lignin of Panicum virgatum and Eucalyptus globulus as a function of ionic liquid pretreatment.* Bioresource Technology, 2012. **126**: p. 156-161.

205. Shen, R., L. Tao, and B. Yang, *Techno-economic analysis of jet-fuel production from biorefinery waste lignin.* Biofuels, Bioproducts and Biorefining, 2018. **0**(0).

206. Mandal, S.M., R.O. Dias, and O.L. Franco, *Phenolic Compounds in Antimicrobial Therapy.* Journal of Medicinal Food, 2017. **20**(10): p. 1031-1038.

207. Espinoza-Acosta, J.L., et al., *Antioxidant, antimicrobial, and antimutagenic properties of technical lignins and their applications.* BioResources, 2016. **11**(2): p. 5452-5481.

208. Li, W., et al., *Understanding Low-Pressure Hydropyrolysis of Lignin Using Deuterated Sodium Formate.* ACS Sustainable Chemistry & Engineering, 2017. **5**(10): p. 8939-8950.

209. Riaz, A., et al., *High-yield and high-calorific bio-oil production from concentrated sulfuric acid hydrolysis lignin in supercritical ethanol.* Fuel, 2016. **172**: p. 238-247.

210. Lee, H.S., et al., *Hydro- and solvothermolysis of kraft lignin for maximizing production of monomeric aromatic chemicals.* Bioresour Technol, 2016. **203**: p. 142-9.

211. Zaheer, M. and R. Kempe, *Catalytic Hydrogenolysis of Aryl Ethers: A Key Step in Lignin Valorization to Valuable Chemicals.* ACS Catalysis, 2015. **5**(3): p. 1675-1684.

212. Cheng, S., et al., *Hydrothermal degradation of alkali lignin to bio-phenolic compounds in sub/supercritical ethanol and water–ethanol co-solvent.* Polymer Degradation and Stability, 2012. **97**(6): p. 839-848.

213. Huang, X., et al., *Ethanol as capping agent and formaldehyde scavenger for efficient depolymerization of lignin to aromatics.* Green Chemistry, 2015. **17**(11): p. 4941-4950.

214. Katahira, R., et al., *Base-Catalyzed Depolymerization of Biorefinery Lignins.* ACS Sustainable Chemistry & Engineering, 2016. **4**(3): p. 1474-1486.

215. Dodge, L.A., et al., *Sequential Extraction and Characterization of Lignin-Derived Compounds from Thermochemically Processed Biorefinery Lignins.* Energy & Fuels, 2019. **33**(5): p. 4322-4330.

216. McClelland, D.J., et al., *Functionality and molecular weight distribution of red oak lignin before and after pyrolysis and hydrogenation.* Green Chemistry, 2017. **19**(5): p. 1378-1389.

217. Bowman, A.S., S.O. Asare, and B.C. Lynn, *Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis for characterization of lignin oligomers using cationization techniques and 2,5-*

dihydroxyacetophenone (DHAP) matrix. Rapid Communications in Mass Spectrometry, 2019. **33**(8): p. 811-819.

218. Kim, K.H., B.A. Simmons, and S. Singh, *Catalytic transfer hydrogenolysis of ionic liquid processed biorefinery lignin to phenolic compounds.* Green Chemistry, 2017. **19**(1): p. 215-224.

219. Wen, J.L., et al., *Recent Advances in Characterization of Lignin Polymer by Solution-State Nuclear Magnetic Resonance (NMR) Methodology.* Materials (Basel), 2013. **6**(1): p. 359-391.

220. Ralph, S.A., L.L. Landucci, and J. Ralph. *NMR database of lignin and cell wall model compounds*. Available over Internet at http://ars.usda.gov/Services/docs.htm?docid=10491 (previously http://www.dfrc.ars.usda.gov/software.html), updated at least annually since 1993 2009 2009 [cited Available over Internet at https://www.glbrc.org/databases_and_software/nmrdatabase/ [previously at http://ars.usda.gov/Services/docs.htm?docid=10491 (and before that at http://www.dfrc.ars.usda.gov/software.html)];updated sporadically; Available from: https://www.glbrc.org/databases_and_software/nmrdatabase/.

221. Stan-Lotter, H., et al., *24 The Assessment of the Viability of Halophilic Microorganisms in Natural Communities*, in *Methods in Microbiology*. 2006, Academic Press. p. 569-584.

222. Wang, H., et al., *Effects of Sugars, Furans, and their Derivatives on Hydrodeoxygenation of Biorefinery Lignin-Rich Wastes to Hydrocarbons.* ChemSusChem, 2018. **11**(15): p. 2562-2568.

223. Miller. *Solvent Polarity Table*. 1998; Available from: https://sites.google.com/site/miller00828/in/solvent-polarity-table.

224. Liu, J.-Y., S. Wu, and R. Lou, *Chemical structure and pyrolysis response of β-O-4 lignin model polymer.* BioResources, 2011. **6**.

225. Tayier, M., et al., *Bamboo biochar-catalytic degradation of lignin under microwave heating.* Journal of Wood Chemistry and Technology, 2020. **40**(3): p. 190-199.

226. Anderson, E.M., et al., *Reductive Catalytic Fractionation of Corn Stover Lignin.* ACS Sustainable Chemistry & Engineering, 2016. **4**(12): p. 6940-6950.

227. Ralph, J., et al., *Peroxidase-dependent cross-linking reactions of p-hydroxycinnamates in plant cell walls.* Phytochemistry Reviews, 2004. **3**(1): p. 79-96.

228. Klein, I., B. Saha, and M.M. Abu-Omar, *Lignin depolymerization over Ni/C catalyst in methanol, a continuation: effect of substrate and catalyst loading.* Catalysis Science & Technology, 2015. **5**(6): p. 3242-3245.

229. Huang, X., et al., *Catalytic Depolymerization of Lignin in Supercritical Ethanol.* ChemSusChem, 2014. **7**(8): p. 2276-2288.

230. Tian, Q.K., et al., *Study on Solvent Extraction of Fast Pyrolysis Bio-Oil.* Advanced Materials Research, 2011. **306-307**: p. 1532-1536.

231. Bennour, N., et al., *Effect of solvent evaporation method on phenolic compounds and the antioxidant activity of Moringa oleifera cultivated in Southern Tunisia.* South African Journal of Botany, 2020. **129**: p. 181-190.

232. Del-Toro-Sánchez, C.L., et al., *Storage Effect on Phenols and on the Antioxidant Activity of Extracts from <i>Anemopsis californica</i> and Inhibition of Elastase Enzyme.* Journal of Chemistry, 2015. **2015**: p. 602136.

233. Shuai, L., et al., *Formaldehyde stabilization facilitates lignin monomer production during biomass depolymerization.* Science, 2016. **354**(6310): p. 329-333.

234. Kim, K.H. and C.S. Kim, *Recent Efforts to Prevent Undesirable Reactions From Fractionation to Depolymerization of Lignin: Toward Maximizing the Value From Lignin.* Frontiers in Energy Research, 2018. **6**(92).

235. Shuai, L. and B. Saha, *Towards high-yield lignin monomer production.* Green Chemistry, 2017. **19**(16): p. 3752-3758.

236. Lan, W. and J.S. Luterbacher, *Preventing lignin condensation to facilitate aromatic monomer production.* CHIMIA International Journal for Chemistry, 2019. **73**(7-8): p. 591-598.

237. Setlow, P., *Spores of Bacillus subtilis: their resistance to and killing by radiation, heat and chemicals.* Journal of Applied Microbiology, 2006. **101**(3): p. 514-525.

238. O'gara, J.P. and H. Humphreys, *Staphylococcus epidermidis biofilms: importance and implications.* Journal of medical microbiology, 2001. **50**(7): p. 582-587.

239. M Bischoff, K., et al., *Modeling Bacterial Contamination of Fuel Ethanol Fermentation.* Vol. 103. 2009. 117-22.

240. Yang, J.F., et al., *Chemical Composition, Antioxidant, and Antibacterial Activity of Wood Vinegar from Litchi chinensis.* Molecules, 2016. **21**(9).

241. Fitzgerald, D.J., et al., *Mode of antimicrobial action of vanillin against Escherichia coli, Lactobacillus plantarum and Listeria innocua.* Journal of Applied Microbiology, 2004. **97**(1): p. 104-113.

242. Zbikowska, B., et al., *Antimicrobial and Antiradical Activity of Extracts Obtained from Leaves of Five Species of the Genus Bergenia: Identification of Antimicrobial Compounds.* Microb Drug Resist, 2017. **23**(6): p. 771-780.

243. Faria, F.A., et al., *Lignin-based polyurethane doped with carbon nanotubes for sensor applications.* Polymer International, 2012. **61**(5): p. 788-794.

244. Kevin Ii, D.A., D.A. Meujo, and M.T. Hamann, *Polyether ionophores: broad-spectrum and promising biologically active molecules for the control of drug-resistant bacteria and parasites.* Expert opinion on drug discovery, 2009. **4**(2): p. 109-146.

245. Campos, F.M., et al., *Cell membrane damage induced by phenolic acids on wine lactic acid bacteria.* Int J Food Microbiol, 2009. **135**(2): p. 144-51.

246. Wu, Y., et al., *Antibacterial Activity and Membrane-Disruptive Mechanism of 3-p-trans-Coumaroyl-2-hydroxyquinic Acid, a Novel Phenolic Compound from Pine Needles of Cedrus deodara, against Staphylococcus aureus.* Molecules, 2016. **21**(8).

247. Fitzgerald, D.J., et al., *Mode of antimicrobial action of vanillin against Escherichia coli, Lactobacillus plantarum and Listeria innocua.* J Appl Microbiol, 2004. **97**(1): p. 104-13.

248. Lundevall, F.J., et al., *A Co2B Mediated NaBH4 Reduction Protocol Applicable to a Selection of Functional Groups in Organic Synthesis.* European Journal of Organic Chemistry, 2018. **2018**(26): p. 3416-3425.

249. Beckner, M., M.L. Ivey, and T.G. Phister, *Microbial contamination of fuel ethanol fermentations.* Letters in Applied Microbiology, 2011. **53**(4): p. 387-394.

250. Brexó, R.P. and A.S. Sant'Ana, *Impact and significance of microbial contamination during fermentation for bioethanol production.* Renewable and Sustainable Energy Reviews, 2017. **73**: p. 423-434.

251. Rich, J.O., et al., *Biofilm formation and ethanol inhibition by bacterial contaminants of biofuel fermentation.* Bioresource Technology, 2015. **196**: p. 347-354.

252. Skinner, K.A. and T.D. Leathers, *Bacterial contaminants of fuel ethanol production.* Journal of Industrial Microbiology and Biotechnology, 2004. **31**(9): p. 401-408.

253. Rich, J.O., et al., *Resolving bacterial contamination of fuel ethanol fermentations with beneficial bacteria – An alternative to antibiotic treatment.* Bioresource Technology, 2018. **247**: p. 357-362.

254. Lushia, W. and P. Heist, *Antibiotic resistant bacteria in fuel ethanol fermentations.* Ethanol Producer Magazine, 2005: p. 80-82.

255. Bischoff, K.M., Y. Zhang, and J.O. Rich, *Fate of virginiamycin through the fuel ethanol production process.* World J Microbiol Biotechnol, 2016. **32**(5): p. 76.

256. Lutgring, J.D., C.A.D. Granados, and J.E. McGowan, *Antimicrobial Resistance: An International Public Health Problem*, in *Antimicrobial Drug Resistance: Clinical and Epidemiological Aspects, Volume 2*, D.L. Mayers, et al., Editors. 2017, Springer International Publishing: Cham. p. 1519-1528.

257. Magalhães, L.M., et al., *Rapid microplate high-throughput methodology for assessment of Folin-Ciocalteu reducing capacity.* Talanta, 2010. **83**(2): p. 441-447.

258. Kazeem, M., J. Adamson, and I. Ogunwande, *Modes of inhibition of α-amylase and α-glucosidase by aqueous extract of Morinda lucida Benth leaf.* BioMed research international, 2013. **2013**.

259. Park, S.-Y., et al., *Effects of peracetic acid and hydrogen peroxide concentration on kraft lignin degradation at room temperature.* Bioresources, 2019. **14**: p. 4413-4420.

260. Park, S.-Y., et al., *Peracetic acid-induced kraft lignin solubilization and its characterization for selective production of macromolecular biopolymers.* International Journal of Biological Macromolecules, 2020. **161**: p. 1240-1246.

261. Chauret, C.P., *Sanitization*, in *Encyclopedia of Food Microbiology (Second Edition)*, C.A. Batt and M.L. Tortorello, Editors. 2014, Academic Press: Oxford. p. 360-364.

262. Rutala, W.A. and D.J. Weber, *301 - Disinfection, Sterilization, and Control of Hospital Waste*, in *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases (Eighth Edition)*, J.E. Bennett, R. Dolin, and M.J. Blaser, Editors. 2015, Content Repository Only!: Philadelphia. p. 3294-3309.e4.

263. Blainski, A., G.C. Lopes, and J.C.P. de Mello, *Application and analysis of the folin ciocalteu method for the determination of the total phenolic content from Limonium brasiliense L.* Molecules (Basel, Switzerland), 2013. **18**(6): p. 6852-6865.

264. Bjarnsholt, T., et al., *Antibiofilm Properties of Acetic Acid.* Advances in wound care, 2015. **4**(7): p. 363-372.

265. Sousa, M., et al., *Stress and cell death in yeast induced by acetic acid.* Cell Metabolism-Cell Homeostasis and Stress Response, 2012.

266. Khameneh, B., et al., *Review on plant antimicrobials: a mechanistic viewpoint.* Antimicrobial resistance and infection control, 2019. **8**: p. 118-118.

267. Bischoff, K., et al., *Modeling Bacterial Contamination of Fuel Ethanol Fermentation.* Biotechnology and bioengineering, 2009. **103**: p. 117-22.

268. Araya-Cloutier, C., et al., *QSAR-based molecular signatures of prenylated (iso)flavonoids underlying antimicrobial potency against and membrane-disruption in Gram positive and Gram negative bacteria.* Scientific Reports, 2018. **8**(1): p. 9267.

269. Camacho, D.M., et al., *Next-Generation Machine Learning for Biological Networks.* Cell, 2018. **173**(7): p. 1581-1592.

270. Lenselink, E.B., et al., *Beyond the hype: deep neural networks outperform established methods using a ChEMBL bioactivity benchmark set.* Journal of Cheminformatics, 2017. **9**(1): p. 45.

271. Gaulton, A., et al., *ChEMBL: a large-scale bioactivity database for drug discovery.* Nucleic Acids Res, 2012. **40**(Database issue): p. D1100-7.

272. Andrews, J.M., *Determination of minimum inhibitory concentrations.* J Antimicrob Chemother, 2001. **48 Suppl 1**: p. 5-16.

273. Weininger, D., A. Weininger, and J.L. Weininger, *SMILES. 2. Algorithm for generation of unique SMILES notation.* Journal of Chemical Information and Computer Sciences, 1989. **29**(2): p. 97-101.

274. Song, Y., et al., *Gold-catalyzed conversion of lignin to low molecular weight aromatics.* Chemical Science, 2018. **9**(42): p. 8127-8133.

275. Pedregosa, F., et al., *Scikit-learn: Machine learning in Python.* the Journal of machine Learning research, 2011. **12**: p. 2825-2830.

276. Golbraikh, A., et al., *Rational selection of training and test sets for the development of validated QSAR models.* Journal of Computer-Aided Molecular Design, 2003. **17**(2): p. 241-253.

277. Hira, Z.M. and D.F. Gillies, *A review of feature selection and feature extraction methods applied on microarray data.* Advances in bioinformatics, 2015. **2015**.

278. Seabold, S. and J. Perktold. *Statsmodels: Econometric and statistical modeling with python.* in *Proceedings of the 9th Python in Science Conference.* 2010. Austin, TX.

279. Mei, H., et al., *Support vector machine applied in QSAR modelling.* Chinese Science Bulletin, 2005. **50**(20): p. 2291-2296.

280. Basant, N., S. Gupta, and K.P. Singh, *QSAR modeling for predicting reproductive toxicity of chemicals in rats for regulatory purposes.* Toxicology research, 2016. **5**(4): p. 1029-1038.

281. Svetnik, V., et al., *Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling.* Journal of Chemical Information and Computer Sciences, 2003. **43**(6): p. 1947-1958.

282. Zheng, W. and A. Tropsha, *Novel Variable Selection Quantitative Structure−Property Relationship Approach Based on the k-Nearest-Neighbor Principle.* Journal of Chemical Information and Computer Sciences, 2000. **40**(1): p. 185-194.

283. Vamathevan, J., et al., *Applications of machine learning in drug discovery and development.* Nature reviews. Drug discovery, 2019. **18**(6): p. 463-477.

284. Svensson, F., U. Norinder, and A. Bender, *Modelling compound cytotoxicity using conformal prediction and PubChem HTS data.* Toxicology research, 2017. **6 1**: p. 73-80.

285. Speck-Planche, A., et al., *Predicting multiple ecotoxicological profiles in agrochemical fungicides: a multi-species chemoinformatic approach.* Ecotoxicol Environ Saf, 2012. **80**: p. 308-13.

286. Bengio, Y., I. Goodfellow, and A. Courville, *Deep learning.* Vol. 1. 2017: MIT press Massachusetts, USA:.

287. Taguri, T., T. Tanaka, and I. Kouno, *Antibacterial Spectrum of Plant Polyphenols and Extracts Depending upon Hydroxyphenyl Structure.* Biological & pharmaceutical bulletin, 2006. **29**: p. 2226-35.

288. Cynthia, et al., *Antibacterial and antioxidant activities of pyrogallol and synthetic pyrogallol dimer.* Research Journal of Chemistry and Environment, 2018. **22**: p. 39-47.

289. Aziz, N., et al., *Comparative antibacterial and antifungal effects of some phenolic compounds.* Microbios, 1998. **93**(374): p. 43-54.

290. Labute, P., *A widely applicable set of descriptors.* Journal of Molecular Graphics and Modelling, 2000. **18**(4): p. 464-477.

291. Hall, L.H., B. Mohney, and L.B. Kier, *The electrotopological state: structure information at the atomic level for molecular graphs.* Journal of Chemical Information and Computer Sciences, 1991. **31**(1): p. 76-82.

292. Bouarab-Chibane, L., et al., *Antibacterial Properties of Polyphenols: Characterization and QSAR (Quantitative Structure-Activity Relationship) Models.* Frontiers in microbiology, 2019. **10**: p. 829-829.

EDUCATION:

M.S., Biological Sciences, Eastern Illinois University, 2016

M.S., Sustainable Energy, Eastern Illinois University, 2016

B.S., Environmental Science, University of St. Francis, 2014

PROFESSIONAL EXPERIENCE:

- Graduate Research Assistant, University of Kentucky. Biosystems and Agricultural Engineering Department, Lexington, KY: August 2016 to present
- Part-Time Instructor, University of Kentucky. Biosystems and Agricultural Engineering Department, Lexington, KY: Spring 2019
- Graduate Research/Teaching Assistant, Eastern Illinois University. Department of Biology, Charleston, IL: August 2014 – May 2016
- Intern Operator, Charleston Illinois Wastewater Treatment Plant, Charleston, IL: Summer 2014
- Undergraduate Researcher, University of Nebraska -Lincoln Summer Research Program (NSF REU), Lincoln, NE: Summer 2013
- Undergraduate Lab/Research Assistant, University of St. Francis, Joliet, IL: Summer 2010-2014

PUBLICATIONS:

1. Kalinoski R. M., Wenqi Li, Justin K. Mobley, Shardrack O. Asare, Bert C. Lynn, Xiaowen Chen, Jian Shi. 2019. Antimicrobial properties of corn stover lignin fractions derived from catalytic transfer hydrogenolysis in supercritical ethanol with a Ru/C catalyst. *ChemSusChem*. In progress (Manuscript Submitted).

2. Mohamad Barekati-Goudarzi, Dorin Boldor, Lavrent Khachatryan, Lynn Bert, Ryan Kalinoski, Jian Shi. Heterogeneous and Homogenous Components in Gas Phase Pyrolysis of Hydrolytic Lignin. 2020. ACS Sustainable Chemistry and Engineering. 8, 34, 12891-12901. https://doi.org/10.1021/acssuschemeng.0c03366

3. Jorge A. Belgodere, Syed A. Zamin, Ryan M. Kalinoski, Carlos E. Astete, Joseph C. Penrod, Katie M. Hamel, Bert C. Lynn, Jai S. Rudra, Jian Shi, and Jangwook P. Jung. 2019. Modulating Mechanical Properties of Collagen–Lignin Composites. *ACS Applied Bio Materials*. DOI: 10.1021/acsabm.9b00444

4. Luke A. Dodge, Ryan M. Kalinoski, Lalitendu Das, Jacob Bursavich, Pranjali Muley, Dorin Boldor, Jian Shi. 2019. Sequential Extraction and Characterization of Lignin-Derived Compounds from Thermochemically Processed Biorefinery Lignins. *Energy Fuels*. DOI: https://doi.org/10.1021/acs.energyfuels.9b00376

5. Kalinoski R. M, Jian Shi. 2018. Hydrogels derived from lignocellulosic compounds: Evaluation of the compositional, structural, mechanical and antimicrobial properties. *Industrial Crops and Products*. DOI: https://doi.org/10.1016/j.indcrop.2018.11.002

6. Kalinoski R. M, Hector D. Flores, Sunil Thapa, Erin R. Tuegel, Michael A. Bilek, Evelin Y. Reyes-Mendez, Michael J. West, Tim J. Dumonceaux, Thomas Canam. 2017. Pretreatment of Hardwood and Miscanthus with Trametes versicolor for Bioenergy Conversion and Densification Strategies. *Appl Biochem Biotechnol.* doi:10.1007/s12010-017-2507-3

7. Kalinoski R. M., and DeLong J. P. Beyond body mass: how prey traits improve predictions of functional response parameters. *Oecologia*. 2015. DOI: 10.1007/s00442-015-3487-z.

8. Thapa, S., D. B. Johnson, E. R. Tuegel, R. M. Kalinoski, R. M. Ellingston, M. J. West, P. P. Liu, T. J. Dumonceaux, and T. Canam. 2015. Assessing the thermomechanical characteristics of lignocellulosic biomass. *Journal of Visualized Experiments* (accepted).

9. Novich, R. A., E. K. Erickson, R. M. Kalinoski, and J. P. DeLong. 2014. The temperature independence of interaction strength in a sit-and-wait predator. *Ecosphere* 5(10):137. http://dx.doi.org/10.1890/ES14-00216.1

PRESENTATIONS:

1. Kalinoski R.M., Shi J. 2019. Controlling Bacterial Contamination During Fuel Ethanol Fermentation by Utilizing Thermochemically Depolymerized Lignin Bio-Oils. *ASABE National Meeting.* Virtual. July 2020.

2. Kalinoski R.M., Shi J. 2019. Antimicrobial properties of lignin bio-oil derived from catalytic transfer hydrogenolysis of alkali treated corn stover lignin in supercritical ethanol with a Ru/C catalyst. *S-1075 Multistate Committee Annual Meeting and the Symposium on Science and Technology Driving the Bioeconomy.* Golden, CO. July.

3. Kalinoski R.M., Shi J. 2019. Antimicrobial Properties of Lignin Derivatives from Thermochemical Depolymerization. *ASABE National Meeting.* Boston, MA. July.

4. Kalinoski R.M., Shi J. 2019. Antimicrobial Properties of depolymerized lignin compounds. *ACS National Meeting & Expo*. Orlando, FL. March.

5. Kalinoski R.M., Olaleye A., Shi J. 2018. Optimization of Lipo-chitooligosaccharides Production by Solid State Cultivation of Bradyrhizobium japonicum on Sweet Sorghum Bagasse. *ASABE Annual Conference.* Detroit, MI. July.

6. Kalinoski R.M., Chirwa U., Shi J. 2018. Antimicrobial Properties of Lignin in Lignocellulosic Hydrogels and Lignin Derived Compounds. *ASABE Annual Conference*. Detroit, MI. July.

7. Kalinoski R.M., Shi J. 2018. Hydrogels Derived from Lignocellulosic Compounds: Evaluation of the Compositional, Structural, Mechanical and Antimicrobial Properties. A*SABE Annual Conference: Boyd Scott Award*. Detroit, MI. July.

8. Kalinoski R.M., Shi J. 2018. Antimicrobial Properties of Lignin in Lignocellulosic Hydrogels and Lignin Derived Compounds. *GRS: Nano-Enabled Technologies to Improve Efficiency, Quality, and Health in Food and Agriculture.* South Hadley, MA. June

9. Kalinoski R.M., Chirwa U., Shi J. 2017. Antimicrobial Hydrogels Derived from Lignocellulosic Compounds. *ASABE Annual Conference*. Spokane, WA. July.

10. Kalinoski R.M., Flores H.D., Priya N., Liu P.P., Canam T. 2015. The white-rot fungus, *Trametes versicolor*, as an environmentally benign pretreatment agent for Miscanthus straw. *European Biomass Conference and Exhibition*. Vienna, Austria. June.

11. Kalinoski R.M., Flores H.D., Priya N., Ellsington R., West M., Canam T. 2015. Fungi for Bioenergy: Development of a Sustainable Biomass Pretreatment System. *EPA P3 Expo*. Washington, D.C. April.

AWARDS AND HONORS:

1. Outstanding PhD Graduate Student Award, BAE, 2020

2. ASABE Oral/Poster Competition Award, 2019

3. Boyd-Scott Graduate Research Award, ASABE, 2018

4. Alpha Epsilon, scholastic honor society member, 2018 to present

5. Hamand Society Scholar, for outstanding achievement in scholarship and service, Eastern Illinois University, Spring 2016

6. Distinguished Graduate Student Award for Biological Sciences, Eastern Illinois University, Spring 2016

7. Lewis Hanford Tiffany Botany Graduate Research Fund Award, $300, Eastern Illinois University, Spring 2015

8. College of Science Graduate Travel Award, $650, Eastern Illinois University, Spring 2015

9. Graduate School Research/Creative Activity Grant, $750, Eastern Illinois University, Fall 2014

10. Delta Epsilon Sigma, national scholastic honor society member, University of St. Francis, 2012-present

11. Duns Scotus Honor's Program, University of St. Francis, 2010-2014

12. Science Fellows Scholarship, yearly tuition stipend $3000, University of St. Francis, 2010-2014