ABSTRACT

Narron, Robert Hardwicke. Biorefinery Processing of Diverse Protolignins to Assist Biorefinery Lignin Valorization. (Under the direction of Dr. Sunkyu Park and Dr. Hasan Jameel).

Carbohydrate-derived products are currently the primary and sole output of biorefinery processes. Biorefinery lignin co-products, if properly developed, are expected to supplement biorefinery income and reduce reliance upon the market of a single output product. A broad survey of biomasses have been comprehensively analyzed, focusing upon differing lignin's responses to biorefinery processing. Characterization results are intended to elucidate biorefinery lignin properties and aid biomass selection for biorefineries looking to valorize lignin.

The pretreatment utilized in the investigated cellulase-based biorefinery process was autohydrolysis. Autohydrolysis is a simple pretreatment that utilizes only water, temperature, and pressure to facilitate biomass degradation and deconstruction. Chemical degradation is enabled by cleavage of hemicellulosic acetates to produce acetic acid in the water hydrolyzate ("autohydrolyzate").

Autohydrolysis was found to render minor quantities of hardwood and non-wood lignin soluble in autohydrolyzate. A method for soluble lignin segregation was developed and applied across all studied biomasses in order to elucidate this mostly unknown biorefinery lignin fraction. Autohydrolyzate-soluble lignin is of particular interest based upon its low molecular weight, low polydispersity, high phenolic functionality, and aqueous solubility. These properties are ripe for innovation of co-products based upon autohydrolyzate-soluble lignin. In addition, sequestration of autohydrolyzate-solute lignin

backhandedly revealed a means of separating lignin from hemicellulosic xylooligosaccharides. These oligosaccharides are also attractive biorefinery co-products based on the existing market for similar non-digestible oligosaccharide food additives.

Insoluble residues after pretreatment and cellulolytic hydrolysis are lignin enriched due to the carbohydrate-targeting unit operations preceding its generation. This material has been partially characterized to describe it relative to protolignin, with conclusions being drawn that identify key markers in protolignin that are most influential upon residual lignin's chemical and molecular properties. For both hardwood and non-wood, trends have been identified that relate lignin reactivity with monolignol constituency.

Biorefinery Processing of Diverse Protolignins to Assist Biorefinery Lignin Valorization

by
Robert Hardwicke Narron

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Forest Biomaterials

Raleigh, North Carolina

2017

APPROVED BY:

Dr. Sunkyu Park Committee Co-chair	Dr. Hasan Jameel Committee Co-chair
Dr. Hou-min Chang	Dr. Jonathan Lindsey

DEDICATION

To the love of my life and future bride, Natalie Griffie.

To my parents and family for supporting me during my work. Thank you Elizabeth Narron, John Narron, Suzanne “Nana” Macklen, Park Narron, and Welch Narron.

To my closest friend Baxter Wells III, who has been my best friend since growing up in Wendell.

To my colleagues Xiao Jiang, Seunghyun Yoo, Wenhui Geng, Kelsey Boes, Caoxing Huang, Saara Hanhikoski, Adriana Ribeiro, Valeria Gomes, and Runkun Sun.

To my advisers and committee members, Hasan Jameel, Sunkyu Park, Hou-min Chang, and Jonathan Lindsey.

To my mentors, Barbara White, Jie Liu, Steve Kelley, Hanna Gracz, Jim McMurray, Linda McMurray, and Mike Maltby.

To Dan, Stu (Juan), Mike, Guillermo, Roy, Fat Chris, Allyson, Greg Cote, Ron Magill, Meech, Tom Kurkjian, Mina, Amin, Domonique, Pablo, Bomani, and Stan van Gundy for providing me with free entertainment that I would have paid for.

BIOGRAPHY

Robert Hardwicke Narron was born April 7th, 1991. He grew up in Wendell, North Carolina until he moved to Raleigh, North Carolina in 2009 to attend North Carolina State University for undergraduate studies. After earning a B.A. in Chemistry (2013), he was serendipitously introduced to Drs. Sunkyu Park and Hasan Jameel. These two had recently secured funding to provide free doctoral studies to students seeking advanced degrees in the field of lignocellulosic biomass. Robert became quickly enamored by lignin, the complex binding material nature produces to provide strength to wood and non-woods alike. Robert will finally depart NCSU in fall 2017 to begin his career with Borregaard LignoTech, a company that monetizes lignin acquired from various pulping operations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. Biomass pretreatments capable of enabling lignin valorization in a biorefinery process

Robert H. Narron, Hoyong Kim, Hou-min Chang, Hasan Jameel, Sunkyu Park

Current Opinion in Biotechnology (2016), 38, 39-46

Abstract

Actualization of the lignocellulosic biorefinery, which converts renewable biomass into valuable fuels, chemicals, and materials, continues development towards industrial realization. Recent economic and technical feasibility studies have concluded that creating value from biomass' lignin will assist realization of such processes due to co-valorization and the new revenues beyond carbohydrates thus enabled. The pretreatment step within a biorefinery process is indispensable towards recovering carbohydrates, but different techniques and intensities have a variety of effects on lignin. Acidic and alkaline pretreatments have shown to produce diverse lignins based on delignification chemistry. The valorization potential of pretreated lignin is affected by its chemical structure, which is known to degrade with inter-lignin condensation under high-severity pretreatment. Co-valorization of lignin and carbohydrates will require dampening of pretreatment intensities to avoid such effects, in spite of tradeoffs in carbohydrate production.

1. Introduction

To cultivate economic value from renewable biomass, a pretreatment operation of incoming biomass is a critical step to overcome the recalcitrant properties of lignocellulosic biomass for downstream biological conversion. It is known that lignin can inhibit both enzymatic saccharification and fermentation of carbohydrates, leading to the necessity of biomass pretreatment for lignin removal prior to said processes.[1] The diverse array of pretreatments that can be optimized for carbohydrate conversion have been reviewed over recent years.[2-12] In a newly published techno-economic analysis of a hypothetical industrial biorefinery, return on investment was highly susceptible to the incoming carbohydrate fraction in biomass.[13] In the best case, marginal returns were predicted in addition to addressing the market risks. These findings support the actively occurring paradigm shift in lignocellulosic biorefinery research towards valorization of the lignin fraction from biomass, both soluble in the pretreatment hydrolyzate or insoluble as residue after enzymatic hydrolysis of carbohydrates.[14-26]

Despite recently successful investigation into novel methods for valorizing biorefinery lignin by conversion into precursors for different chemical and material applications, the published results remain at the laboratory scale.[27-31] In order to advance lignin-valorizing technologies, a comprehensive understanding of biorefinery lignin streams must be further established. Currently there exists a void of explicit knowledge regarding the modifications to lignin's properties after pretreatment across techniques. This dearth of knowledge lends itself to the practice of techno-economic models of industrial-scale biorefineries treating lignin solely as a burnable source of power for the biorefinery itself or for consumers (fuel pellets).

With a better understanding of how lignin's structure is affected by pretreatment conditions, potential applications (beyond combustion) for lignin streams would be more readily innovated and developed. Furthermore, if any stream of biorefinery lignin can be valorized, then profitability of the entire lignocellulosic biorefinery schema would rise due to reduced dependence upon market prices for carbohydrate-derived products.

The two potential streams from which lignin can be valorized within a biorefinery are pretreatment hydrolyzate (solubilized lignins) and post-saccharification residue (insoluble lignins). It is impossible to correlate the amount of lignin solubilized during pretreatment with valorization potential of soluble and insoluble lignins due to inadequate data regarding the characteristics of the pretreated lignin. When comparing diverse pretreatment processes and intensities across woody and non-woody biomasses, great variability in lignin solubilization is observed. Aside from percent delignification during pretreatment, there tends to be sparse discussion regarding the chemical properties of either lignin stream due to carbohydrate-centric research objectives. The absence of understanding regarding the whole lignin structure after pretreatment is highlighted by the dissimilarity of lignin structural characterizations reported across recent literature (Table 1). Therefore it can be stated that current knowledge of lignin's chemical structure after pretreatment is only in its fledgling state, marred by selective analysis which yields incomplete understanding of the entire lignin structure after pretreatment.

2. Lignin properties for high-value applications

Some of the potential applications for biorefinery lignin under investigation involve carbon fibers for automobiles[32], hydrocarbons for transportation fuels and biochemicals[33], and

petro-phenol substitution for resin adhesives.[34] In order to evaluate applicability of pretreated biorefinery lignins for valorization, it is first necessary to define the characteristics of pretreated lignin that are sought for applications. Pretreated lignin can be considered to have greater valorization potential if the following criteria are met: 1) low degree of condensation, 2) chemically labile moieties for functionalization, and 3) low impurity.

For lignin, the degree of condensation indicates the ratio of carbon-carbon bonds versus carbon-hydrogen bonds and carbon-oxygen bonds. If the degree of condensation of lignin is high, then its application can be significantly affected due to the strength of the carbon-carbon bonds, which reduce lignin's potential for further modification. Also, it is known that increasing degrees of condensation will decrease enzymatic saccharification yields[35], leaving the saccharification residue with high amounts of carbohydrate impurity. The degree of condensation for lignin varies with different pretreatments depending upon pretreatment chemistry and intensity. However, the relationship between different pretreatments and the genesis of inter-lignin condensation is not yet understood. Concerning lignin's functional groups and labile moieties, targeted applications could potentially be developed if biorefinery lignin featured an appreciable amount of functionalities or reactive moieties that can be functionalized for application. The lignin biopolymer is formed through radical coupling between precursor monomers, yielding different chemically-bonded lignin *in situ* depending on the distributions of lignin monomeric precursor (*para*-hydroxyphenyl, guaiacyl, syringyl) a given biomass biosynthesizes at a given point during lignification. To have the residual lignin from a biorefinery remain reactive and thus valorizable, it is ideal to maintain the native and labile alkyl-aryl ether bonds and not generate additional carbon-carbon linkages beyond

those formed naturally during plant cell lignification. Regarding impurity, any stream of lignin in a biorefinery process will require an isolation process. The extent of association between lignin and any non-lignin matter such as carbohydrates, silicates, and others, will drive the cost of producing adequately pure lignin. Therefore any lignin stream that has inherently low impurity bears greater potential for valorization.

3. Lignin from acidic pretreatment

Due to the significant difference in behaviors of lignin spanning the aqueous pH range due to its ionizable phenol groups, pretreatments are classified by environmental pH in this review. Pretreatment using mineral acids (e.g. sulfuric acid[36] and nitric acid[37]) or organic acids (e.g. oxalic acid[38] and maleic acid[39]) continue to be investigated due to the ability to remove hemicellulose and lignin for positive effects on enzymatic hydrolysis. Concerning lignin in an acidic medium, the driving force for chemical reactivity begins with hydronium protonation of a hydroxyl functional group which exists in chemical equilibrium with a reactive carbocation. The acidolysis of lignin's alkyl-aryl ethers has been recently revisited using lignin dimer model compounds and was successful in elucidating the major and minor reaction routes for non-phenolic β-O-4' ethers in acidic media, Figure 1.[40] Acidolysis is considered to be one of the major reaction pathways for lignin during acidic pretreatment, cleaving β-O-4' ethers to generate either a β-ketone (Hibbert's ketone) or an aldehyde functionality on one lignin molecule and liberating a free phenol functionality on another.[41] However, in addition to acidolysis, there are potential reaction pathways towards undesired inter-lignin condensation reactions which decreases lignin's valorization potential and incur cellulolytic inhibition.[42]

Inter-lignin condensation occurs when the carbocation does not proceed with acidolysis but rather electrophilic aromatic substitution involving another aromatic proton. Autohydrolysis pretreatment, which uses hot water and pressure to generate a weak acid solution by cleavage of acetyl groups in hemicellulose to produce acetic acid, presents a unique chemical-free pretreatment option.[43] Sequential autohydrolysis (180 °C, 30 min) and organosolv (50% ethanol, 1% hydrochloric acid, 180 °C, 30 min) pretreatment was studied in attempt to generate two streams: one of carbohydrates and the other of high-purity lignin.[44] In this study, a remarkable 68.7% yield of high-purity lignin was achieved as well as the 89.3% carbohydrate yield from enzymatic hydrolysis based on the pretreated hardwood. The amount of condensed G and S moieties bearing hydroxyl functionalities in the dually-pretreated lignin did increase compared to lignin from the raw material (0.07 to 0.96 mmol / g lignin), however, a greater increase in the same moieties from non-condensed G and S was also observed (0.62 to 2.16 mmol / g lignin). Regarding labile chemical bonds, it was found that the entirety of the original β-O-4' and β-5' substructures had been disrupted, leaving mainly native β-β' substructures in the lignin. In this case the labile moieties are already functionalized to hydroxyl groups with more than double the amount existing on non-condensed structures. Finally, a significant discovery involving the chemical structure of pine lignin after wet explosion pretreatment was recently reported.[45] The pretreatment, consisting of cooking loblolly pine in pure water under a pressurized oxygenated atmosphere followed by rapid explosion, was found to increase the methoxy content in lignin by conversion of H moieties into S moieties by oxygen. This finding sheds light on a new chemistry of lignin transformation during pretreatment, which renders a less condensable and thus more-valorizable residual lignin stream.

3. Lignin from alkaline pretreatment

Pretreatment with alkaline chemicals (e.g. sodium hydroxide[36] and ammonia[46]) is also under investigation due to the established alkaline delignification chemistry from the pulping industry. Large quantities of raw lignin are dissolved into alkaline solution creating a "black liquor". Under the conditions of alkaline pretreatment, nucleophilic cleavage of lignin's phenolic alkyl-aryl ethers occurs, promoting lignin solvation into alkali. These reaction pathways are summarized in Figure 2. The initially-high pH of black liquor continuously decreases as neutralizing compounds from the biomass are solubilized during pretreatment; for example, acetic acid, sugar acids (iso-, meta-saccharinic acid), and new phenol groups on lignin. The pH of the black liquor during and post-pretreatment affects the degree of lignin solubility, consequently pretreatment vessels with metered alkali input are required for delignification optimization. In the case of Kraft pulping, the sulfur-corrupted and highly condensed Kraft lignin solvated in black liquor is combusted for the purpose of alkaline chemical recovery, an integral step that renders industrial Kraft pulping an economically-viable process. Kraft lignin recovery from black liquor has historically presented obstacles towards generating any additional value apart from heat, although Domtar (Plymouth, NC, USA) opened the first industrial-scale LignoBoost^{TM} operation for recovering Kraft lignin from black liquor in 2014. There also exists active research towards other means of utilization for Kraft lignin.[47-50] For a biorefinery, the pH an alkaline pretreatment will require careful maintenance. Excessively alkaline conditions with respect to the chosen nucleophile's pKa will promote nucleophilicity, fostering inter-lignin condensation. In addition, the means of recovery concerning solubilized lignin will need to be established for its valorization.

Concerning alkaline pretreatment chemistries, non-phenolic alkyl-aryl ether cleavage was studied recently using lignin dimer model compounds with and without γ carbon hydroxyl functionalities, concluding that the presence of said groups provide steric hindrance towards bond scission of S lignin as the second aromatic ring.[51] Another study compared alkaline pretreatment and enzymatic hydrolysis of switchgrass and corn stover feedstocks.[52] It was found that in order to achieve identical delignification from switchgrass with respect to corn stover, twice the soda charge would be required (154 mg NaOH / g dry switchgrass at 130 °C for 30 minutes to achieve 72% delignification). Pretreated switchgrass' recalcitrance to enzymatic digestion indicates that the lignin in the saccharification residue is susceptible to bearing carbohydrate impurity depending upon the cellulolytic digestibility of the pretreated solids. The water-soluble components from AFEX (ammonia fiber explosion) pretreated material was recently analyzed in terms of molecular weight and fractionated along the following ranges: < 3 kDa, 3-10 kDa, and >10 kDa.[53] The authors concluded that the < 3 kDa fraction had the greatest contribution to cellulase inhibition. Of this fraction, it was found that based on dry weight up to 25% consisted of inseparable oligomeric lignins and oligomeric hemicelluloses linked either covalently or non-covalently and the remainder lignin and ash. Elsewhere, the black liquor stream from alkaline pretreatment was studied for differences in solubilized lignin molecular weight for downstream valorization.[54] Dissolved pine lignin was observed to have higher molecular weight compared to birch (~5.5 kDa vs. ~4 kDa), yet at the best delignification only 15% of the softwood and 16% of the hardwood lignin solubilized under the conditions of 8% NaOH at 160 °C/60 min and 150 °C/90 min, respectively. Through combination of ionic liquid pre-pretreatment and alkali pretreatment, up to 36% delignification

of eucalyptus was reported as dissolved into black liquor, but more significantly, the lignin in the saccharification residue was found to be highly similar to lignin *in situ*, with 77% and 20% of all lignin side chains being native β-O-4' and β-β' substructures, respectively.[55] This indicates that the ionic liquid and alkaline pretreatment could render valorizable lignin in the saccharification residue due to its well-preserved structure, however, it was concluded to be a synergistic effect. This reduces industrial applicability of such a scheme due to the added pre-pretreatment step that utilizes expensive reagents. Using hot deionized water (140 °C, 20 min) with alkali (1% NaOH), 81% delignification of rice straw was reported[56] with the residual solid lignin capable of bearing its original β-O-4' substructures due to the retardation effect from hydroxide ions on alkyl-aryl ether acidolysis reactions.[57] In a study of pretreatment effects on the lignin structure of olive tree prunings, solvent-extracted solid lignin after alkaline pretreatment (7.5% NaOH at 90°C/90 min) was found to contain high amounts of impurity (23% sugars, 44% ash, 8% total lignin by weight).[58] In addition to the great impurity, the extracted lignin was mainly composed of G lignin (S/G ratio = 0.75) and had ~2-2.5x higher molecular weight compared to the lignin obtained from hydrothermal pretreatment. These findings show a low-yield and highly impure lignin with condensable G moieties is generated from alkaline pretreatment of this agricultural waste.

4. Closing remarks

Figure 3 presents a graphical representation of where established pretreatment technologies exist with respect to valorization of carbohydrates along the lines of pretreatment pH and intensity. There exists optimal pretreatment conditions for carbohydrate production,

but the optimal pretreatments for co-valorization involving carbohydrates and lignins remains unexposed. By establishing which pretreatments can be used to valorize lignin, it will then become possible to investigate the happy medium at which a biorefinery can pretreat incoming biomass to enable simultaneous valorization of carbohydrates and lignins. If pretreatments are optimized from the lens of co-valorization with decreased severities and milder pH's, then the economic prospects of the lignocellulosic biorefinery process will increase in concurrence with likelihood of industrial investment in such an operation.

REFERENCES

1.Zeng, Y.; Zhao, S.; Yang, S.; Ding, S.-Y. Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. *Current Opinion in Biotechnology* **2014,** *27*, 38-45.

2. Barakat, A.; de Vries, H.; Rouau, X. Dry fractionation process as an important step in current and future lignocellulose biorefineries: a review. *Bioresource Technology* **2013,** *134*, 362-373.

3. Bhagwat, S.; Ratnaparkhe, S.; Kumar, A. Biomass Pre-treatment Methods and Their Economic Viability for Efficient Production of Biofuel. *British Biotechnology Journal* **2015,** *8* (2), 1-17.

4. Brodeur, G.; Yau, E.; Badal, K.; Collier, J.; Ramachandran, K.; Ramakrishnan, S. Chemical and physicochemical pretreatment of lignocellulosic biomass: a review. *Enzyme Research* **2011,** *2011*, 1-17.

5.Davis, R.; Tao, L.; Scarlata, C.; Tan, E.; Ross, J.; Lukas, J.; Sexton, D. Process Design and Economics for the Conversion of Lignocellulosic Biomass to Hydrocarbons: Dilute-Acid and Enzymatic Deconstruction of Biomass to Sugars and Catalytic Conversion of Sugars to Hydrocarbons; NREL/TP-5100-62498; National Renewable Energy Laboratory, Golden, CO: 2015.

6.Humbird, D.; Davis, R.; Tao, L.; Kinchin, C.; Hsu, D.; Aden, A.; Schoen, P.; Lukas, J.; Olthof, B.; Worley, M. Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol: dilute-acid pretreatment and enzymatic hydrolysis of corn stover; NREL/TP-5100-47764; National Renewable Energy Laboratory, Golden, CO: 2011.

7. Khoo, H. H. Review of bio-conversion pathways of lignocellulose-to-ethanol: Sustainability assessment based on land footprint projections. *Renewable and Sustainable Energy Reviews* **2015,** *46*, 100-119.

8.Sathitsuksanoh, N.; George, A.; Zhang, Y.; Percival, H. New lignocellulose pretreatments using cellulose solvents: a review. *Journal of Chemical Technology and Biotechnology* **2013,** *88* (2), 169-180.

9.Shi, J.; George, K. W.; Sun, N.; He, W.; Li, C.; Stavila, V.; Keasling, J. D.; Simmons, B. A.; Lee, T. S.; Singh, S. Impact of Pretreatment Technologies on Saccharification and Isopentenol Fermentation of Mixed Lignocellulosic Feedstocks. *BioEnergy Research* **2015,** *8* (3), 1004-1013.

10. Singh, J.; Suhag, M.; Dhaka, A. Augmented digestion of lignocellulose by steam explosion, acid and alkaline pretreatment methods: a review. *Carbohydrate Polymers* **2015,** *117*, 624-631.

11. van der Pol, E. C.; Bakker, R. R.; Baets, P.; Eggink, G. By-products resulting from lignocellulose pretreatment and their inhibitory effect on fermentations for (bio) chemicals and fuels. *Applied Microbiology and Biotechnology* **2014,** *98* (23), 9579-9593.

12.Xu, Z.; Huang, F. Pretreatment methods for bioethanol production. *Applied Biochemistry and Biotechnology* **2014,** *174* (1), 43-62.

13.Treasure, T.; Gonzalez, R.; Jameel, H.; Phillips, R. B.; Park, S.; Kelley, S. Integrated conversion, financial, and risk modeling of cellulosic ethanol from woody and non-woody biomass via dilute acid pre-treatment. *Biofuels, Bioproducts, and Biorefining* **2014,** *8* (6), 755-769.

14.Zamudio, M. A.; Alfaro, A.; de Alva, H. E.; García, J. C.; García-Morales, M.; López, F. Biorefinery of paulownia by autohydrolysis and soda-anthraquinone delignification process. Characterization and application of lignin. *Journal of Chemical Technology and Biotechnology* **2015,** *90* (3), 534-542.

15.Li, H.; Sivasankarapillai, G.; McDonald, A. G. Lignin valorization by forming toughened thermally stimulated shape memory copolymeric elastomers: Evaluation of different fractionated industrial lignins. *Journal of Applied Polymer Science* **2015,** *132* (5), 1-12.

16.Albarelli, J. Q.; Mian, A.; Santos, D. T.; Ensinas, A. V.; Maréchal, F.; Meireles, M. A. A. Valorization of sugarcane biorefinery residues using supercritical water gasification: A case study and perspectives. *Journal of Supercritical Fluids* **2015,** *96*, 133-143.

17.Santos, J. I.; Martín-Sampedro, R.; Fillat, Ú.; Oliva, J. M.; Negro, M. J.; Ballesteros, M.; Eugenio, M. E.; Ibarra, D. Evaluating Lignin-Rich Residues from Biochemical Ethanol Production of Wheat Straw and Olive Tree Pruning by FTIR and 2D-NMR. *International Journal of Polymer Science* **2015,** *2015*, 1-11.

18.Espinoza Acosta, J. L.; Torres Chávez, P. I.; Ramírez-Wong, B.; Bello-Pérez, L. A.; Vega Ríos, A.; Carvajal Millán, E.; Plascencia Jatomea, M.; Ledesma Osuna, A. I. Mechanical, thermal, and antioxidant properties of composite films prepared from durum wheat starch and lignin. *Starch-Stärke* **2015,** *67* (5-6), 502-511.

19.Ten, E.; Vermerris, W. Recent developments in polymers derived from industrial lignin. *Journal of Appled Polymer Science* **2015,** *132* (24), 1-13.

20.Hu, L.; Stevanovic, T.; Rodrigue, D. Comparative Study of Polyethylene Composites Containing Industrial Lignins. *Polymer-Polymer Composites* **2015,** *23* (6), 369.

21. Xie, J.; Hse, C. Y.; Shupe, T. F.; Hu, T. Physicochemical characterization of lignin recovered from microwave-assisted delignified lignocellulosic biomass for use in biobased materials. *Journal of Appled Polymer Science* **2015,** *132* (40), 1-7.

22.Ma, R.; Xu, Y.; Zhang, X. Catalytic Oxidation of Biorefinery Lignin to Value-added Chemicals to Support Sustainable Biofuel Production. *ChemSusChem* **2015,** *8* (1), 24-51.

23.Leskinen, T.; Kelley, S. S.; Argyropoulos, D. S. Refining of ethanol biorefinery residues to isolate value added lignins. *ACS Sustainable Chemistry and Engineering* **2015,** *3* (7), 1632-1641.

24.Rahimi, A.; Ulbrich, A.; Coon, J. J.; Stahl, S. S. Formic-acid-induced depolymerization of oxidized lignin to aromatics. *Nature* **2014,** *515*, 249-252.

25.Öhrman, O. G.; Weiland, F.; Pettersson, E.; Johansson, A.-C.; Hedman, H.; Pedersen, M. Pressurized oxygen blown entrained flow gasification of a biorefinery lignin residue. *Fuel Processing Technology* **2013,** *115*, 130-138.

26.Zhang, W.; Ma, Y.; Wang, C.; Li, S.; Zhang, M.; Chu, F. Preparation and properties of lignin–phenol–formaldehyde resins based on different biorefinery residues of agricultural biomass. *Industrial Crops and Products* **2013,** *43*, 326-333.

27.Deepa, A. K.; Dhepe, P. L. Lignin Depolymerization into Aromatic Monomers over Solid Acid Catalysts. *ACS Catalysis* **2015,** *5* (1), 365-379.

28.Krutov, S. M.; Evtuguin, D. V.; Ipatova, E. V.; Santos, S. A. O.; Sazanov, Y. N. Modification of acid hydrolysis lignin for value-added applications by micronization followed by hydrothermal alkaline treatment. *Holzforschung* **2015,** *69* (6), 761-768.
29.Ma, R.; Hao, W.; Ma, X.; Tian, Y.; Li, Y. Catalytic Ethanolysis of Kraft Lignin into High-Value Small-Molecular Chemicals over a Nanostructured a-Molybdenum Carbide Catalyst. *Angewandte Chemie International Edition* **2014,** *126*, 7438-7443.

30.Singh, S. K.; Ekhe, J. D. Towards effective lignin conversion: HZSM-5 catalyzed one-pot solvolytic depolymerization/hydrodeoxygenation of lignin into value added compounds. *RSC Advances* **2014,** *4* (53), 27971-27978.

31.Zeng, J.; Yoo, C. G.; Wang, F.; Pan, X.; Vermerris, W.; Tong, Z. Biomimetic Fenton-Catalyzed Lignin Depolymerization to High-Value Aromatics and Dicarboxylic Acids. *ChemSusChem* **2015,** *8* (5), 861-871.

32. Mainka, H.; Hilfert, L.; Busse, S.; Edelmann, F.; Haak, E.; Herrmann, A. S. Characterization of the major reactions during conversion of lignin to carbon fiber. *Journal of Material Research and Technology* **2015**, In press.

33.Kim, J.-Y.; Lee, J. H.; Park, J.; Kim, J. K.; An, D.; Song, I. K.; Choi, J. W. Catalytic pyrolysis of lignin over HZSM-5 catalysts: Effect of various parameters on the production of aromatic hydrocarbon. *Journal of Analyticla and Applied Pyrolysis* **2015,** *114*, 273-280.

34. Yang, S.; Zhang, Y.; Yuan, T. Q.; Sun, R. C. Lignin–phenol–formaldehyde resin adhesives prepared with biorefinery technical lignins. *Journal of Applied Polymer Science* **2015,** *132* (36), 1-8.

35.Yu, Z.; Gwak, K. S.; Treasure, T.; Jameel, H.; Chang, H. m.; Park, S. Effect of lignin chemistry on the enzymatic hydrolysis of woody biomass. *ChemSusChem* **2014,** *7* (7), 1942-1950.

36.Cabrera, E.; Muñoz, M. J.; Martín, R.; Caro, I.; Curbelo, C.; Díaz, A. B. Comparison of industrially viable pretreatments to enhance soybean straw biodegradability. *Bioresource Technology* **2015,** *194*, 1-6.

37.Yang, F.; Afzal, W.; Cheng, K.; Liu, N.; Pauly, M.; Bell, A. T.; Liu, Z.; Prausnitz, J. M. Nitric-acid hydrolysis of Miscanthus giganteus to sugars fermented to bioethanol. *Biotechnology and Bioprocess Engineering* **2015,** *20* (2), 304-314.

38.Lacerda, T. M.; Zambon, M. D.; Frollini, E. Oxalic acid as a catalyst for the hydrolysis of sisal pulp. *Industrial Crops and Products* **2015,** *71*, 163-172.

39.Jung, Y. H.; Park, H. M.; Kim, K. H. Whole slurry saccharification and fermentation of maleic acid-pretreated rice straw for ethanol production. *Bioprocess Biosystems and Engineering* **2015,** *38* (9), 1639-1644.

40.Yokoyama, T. Revisiting the Mechanism of β-O-4 Bond Cleavage During Acidolysis of Lignin. Part 6: A Review. *Journal of Wood Chemistry and Technology* **2014,** *35* (1), 27-42.

41. Mitchell, V. D.; Taylor, C. M.; Bauer, S. Comprehensive analysis of monomeric phenolics in dilute acid plant hydrolysates. *BioEnergy Research* **2014,** *7* (2), 654-669.

42.Ko, J. K.; Kim, Y.; Ximenes, E.; Ladisch, M. R. Effect of liquid hot water pretreatment severity on properties of hardwood lignin and enzymatic hydrolysis of cellulose. *Biotechnology and Bioengineering* **2015,** *112* (2), 252-262.

43. Han, Q.; Jin, Y.; Jameel, H.; Chang, H.-m.; Phillips, R.; Park, S. Autohydrolysis Pretreatment of Waste Wheat Straw for Cellulosic Ethanol Production in a Co-located Straw Pulp Mill. *Applied Biochemistry and Biotechnology* **2015,** *175* (2), 1193-1210.

44.Zhu, M.-Q.; Wen, J.-L.; Su, Y.-Q.; Wei, Q.; Sun, R.-C. Effect of structural changes of lignin during the autohydrolysis and organosolv pretreatment on Eucommia ulmoides Oliver for an effective enzymatic hydrolysis. *Bioresource Technology* **2015,** *185*, 378-385.

45.Rana, D.; Laskar, D.; Srinivas, K.; Ahring, B. Wet explosion pretreatment of loblolly pine leads to an increase in methoxylation of the lignin. *Bioresources and Bioprocesses* **2015,** *2* (1), 1-10.

46.Antonopoulou, G.; Gavala, H. N.; Skiadas, I. V.; Lyberatos, G. The Effect of Aqueous Ammonia Soaking Pretreatment on Methane Generation Using Different Lignocellulosic Biomasses. *Waste and Biomass Valorization* **2015,** *6* (3), 281-291.

47.Ma, X.; Cui, K.; Hao, W.; Ma, R.; Tian, Y.; Li, Y. Alumina supported molybdenum catalyst for lignin valorization: Effect of reduction temperature. *Bioresource Technology* **2015,** *192*, 17-22.

48.Awungacha Lekelefac, C.; Busse, N.; Herrenbauer, M.; Czermak, P. Photocatalytic based degradation processes of lignin derivatives. *International Journal of Photoenergy* **2014,** *2015*, 1-18.

49.Zhou, X. F. Catalytic oxidation and conversion of kraft lignin into phenolic products using zeolite-encapsulated Cu (II)[H4] salen and [H2] salen complexes. *Environmental Progress & Sustainable Energy* **2015,** *34* (4), 1120–1128.

50. Li, H.; McDonald, A. G. Fractionation and characterization of industrial lignins. *Industrial Crops and Products* **2014,** *62*, 67-76.

51.Shimizu, S.; Posoknistakul, P.; Yokoyama, T.; Matsumoto, Y. Quantitative difference in the rates of the β-O-4 bond cleavage between lignin model compounds with and without γ-hydroxymethyl groups during the alkaline pulping process. *BioResources* **2013,** *8* (3), 4312-4322.

52.Karp, e.; Resch, M. G.; Donohoe, B. S.; Ciesielski, P. N.; O'Brien, M. H.; Nill, J. E.; Mittal, A.; Biddy, M.; Beckham, G. T. Alkaline pretreatment of switchgrass. *ACS Sustainable Chemistry and Engineering* **2015,** *3* (7), 1479–1491.

53.Humpula, J. F.; Uppugundla, N.; Vismeh, R.; Sousa, L.; Chundawat, S. P. S.; Jones, A. D.; Balan, V.; Dale, B. E.; Cheh, A. M. Probing the nature of AFEX-pretreated corn stover

derived decomposition products that inhibit cellulase activity. *Bioresource Technology* **2014,** *152*, 38-45.

54.Lehto, J.; Pakkanen, H.; Alén, R. Characterization of Lignin Dissolved During Alkaline Pretreatment of Softwood and Hardwood. *Journal of Wood Chemistry and Technology* **2015,** *35* (5), 337-347.

55. Xu, J.-K.; Sun, Y.-C.; Sun, R.-C. Synergistic effects of ionic liquid plus alkaline pretreatments on eucalyptus: Lignin structure and cellulose hydrolysis. *Process Biochemistry* **2015,** *50* (6), 955-965.

56.Imman, S.; Arnthong, J.; Burapatana, V.; Champreda, V.; Laosiripojana, N. Influence of alkaline catalyst addition on compressed liquid hot water pretreatment of rice straw. *Chemical Engineering Journal* **2014,** *278* (15), 85-91.

57.Li, J.; Gellerstedt, G. Improved lignin properties and reactivity by modifications in the autohydrolysis process of aspen wood. *Ind Crop Prod* **2008,** *27* (2), 175-181.
58.Toledano, A.; Erdocia, X.; Serrano, L.; Labidi, J. Influence of extraction treatment on olive tree (Olea europaea) pruning lignin structure. *Environmental Progress & Sustainable Energy* **2013,** *32* (4), 1187-1194.

59.Romero, I.; López-Linares, J. C.; Delgado, Y.; Cara, C.; Castro, E. Ethanol production from rape straw by a two-stage pretreatment under mild conditions. *Bioprocess and Biosystem Engineering* **2015,** *38* (8), 1469-1478.

60.Hendrianie, N.; Juliastuti, S. R.; Iwani, M. I.; Eka, A. Lignocellulosic Processing with Acid Pretreatment and Enzymatic Hydrolysis for Improving the Acquisition of Sugar Fermentation. *Modern Applied Science* **2015,** *9* (7), 23-28.

61. Samuel, R.; Pu, Y.; Raman, B.; Ragauskas, A. J. Structural characterization and comparison of switchgrass ball-milled lignin before and after dilute acid pretreatment. *Applied Biochemistry & Biotechnology* **2010,** *162* (1), 62-74.

62. Wang, K.; Yang, H.; Yao, X.; Xu, F.; Sun, R.-c. Structural transformation of hemicelluloses and lignin from triploid poplar during acid-pretreatment based biorefinery process. *Bioresource Technology* **2012,** *116*, 99-106.

63.Kaparaju, P.; Felby, C. Characterization of lignin during oxidative and hydrothermal pre-treatment processes of wheat straw and corn stover. *Bioresource Technology* **2010,** *101* (9), 3175-3181.

64.Li, J.; Henriksson, G.; Gellerstedt, G. Lignin depolymerization/repolymerization and its critical role for delignification of aspen wood by steam explosion. *Bioresource Technology* **2007,** *98* (16), 3061-3068.

65.Zeng, J.; Tong, Z.; Wang, L.; Zhu, J.; Ingram, L. Isolation and structural characterization of sugarcane bagasse lignin after dilute phosphoric acid plus steam explosion pretreatment and its effect on cellulose hydrolysis. *Bioresource Technology* **2014,** *154*, 274-281.

66.Sun, S.-N.; Li, H.-Y.; Cao, X.-F.; Xu, F.; Sun, R.-C. Structural variation of eucalyptus lignin in a combination of hydrothermal and alkali treatments. *Bioresource Technology* **2015,** *176*, 296-299.

67.Bozell, J. J.; Astner, A.; Baker, D.; Biannic, B.; Cedeno, D.; Elder, T.; Hosseinaei, O.; Delbeck, L.; Kim, J.-W.; O'Lenick, C. Integrating Separation and Conversion—Conversion of Biorefinery Process Streams to Biobased Chemicals and Fuels. *BioEnergy Research* **2014,** *7* (3), 856-866.

68. Bozell, J. J.; O'Lenick, C.; Warwick, S. Biomass fractionation for the biorefinery: Heteronuclear multiple quantum coherence–Nuclear magnetic resonance investigation of lignin isolated from solvent fractionation of switchgrass. *Journal of Agricultural and Food Chemistry* **2011,** *59* (17), 9232-9242.

69.Hu, G.; Cateto, C.; Pu, Y.; Samuel, R.; Ragauskas, A. J. Structural characterization of switchgrass lignin after ethanol organosolv pretreatment. *Energy & Fuels* **2011,** *26* (1), 740-745.

70.Gu, F.; Wang, W.; Jing, L.; Jin, Y. Effects of green liquor pretreatment on the chemical composition and enzymatic digestibility of rice straw. *Bioresource Technology* **2013,** *149*, 375-382.

71. Hussin, M. H.; Rahim, A. A.; Ibrahim, M. N. M.; Brosse, N. Physicochemical characterization of alkaline and ethanol organosolv lignins from oil palm (Elaeis guineensis) fronds as phenol substitutes for green material applications. *Industrial Crops and Products* **2013,** *49*, 23-32.

72.de Carvalho, D. M.; Sevastyanova, O.; Penna, L. S.; da Silva, B. P.; Lindström, M. E.; Colodette, J. L. Assessment of chemical transformations in eucalyptus, sugarcane bagasse and straw during hydrothermal, dilute acid, and alkaline pretreatments. *Industrial Crops and Products* **2015,** *73*, 118-126.

73.Yuan, T.-Q.; You, T.-T.; Wang, W.; Xu, F.; Sun, R.-C. Synergistic benefits of ionic liquid and alkaline pretreatments of poplar wood. Part 2: Characterization of lignin and hemicelluloses. *Bioresource Technology* **2013,** *136*, 345-350.

74. Wen, J.-L.; Sun, S.-L.; Xue, B.-L.; Sun, R.-C. Quantitative structures and thermal properties of birch lignins after ionic liquid pretreatment. *Journal of Agricultrual Food and Chemistry* **2013,** *61* (3), 635-645.

TABLES

Table 1. Summary of pretreatments, severities, delignification, and impacts on lignin structure

Pretreatment	Conditions	pH	Biomass	Wt% Delignification	Impact on lignin structure	Ref.
Dilute acid (Sulfuric acid)	130 °C, 60 min	Acidic	Rape straw	42%	-	[59]
	30 °C, 7 h	Acidic	Banana peel	38%	-	[60]
	190 °C, 1 min	Acidic	Switchgrass	0%	Cβ in β-O-4, Cα in β-5 and β–β 0.74 (untreated); 0.54 (pretreated) Cγ in β-5 and β-O-4 with Cα═O in G and S units 0.35 (untreated); 0.20 (pretreated) Cγ in β-O-4 without Cα═O 0.39 (untreated); 0.25 (pretreated)	[61]
	100~190 °C, 2 h	Acidic	Poplar	-	Nitrobenzene oxidation products Vanillin 53.3%, Syringaldehyde 22.0% (100°C) Vanillin 41.9%, Syringaldehyde 16.8% (170°C) Vanillin 17.8%, Syringaldehyde 12.6% (190°C)	[62]
Steam explosion	170 °C, 3 min	Acidic	Wheat straw	64%	-	[63]
	185 °C, 5 min	Acidic	Aspen	10%	Total β-O-4 54/100 Ar (untreated) 42/100 Ar (pretreated)	[64]
	200 °C, 15 min	Acidic	Aspen	20%	18/100 Ar (pretreated)	
	180 °C, 10 min	Acidic	Sugarcane bagasse	9%	Total β-O-4 25.3–31.7/100 Ar (untreated) 16.7–20.7/100 Ar (pretreated)	[65]
Autohydrolysis	177 °C, 3.5 h	Acidic	Aspen	62%	Total β-O-4 54/100 Ar (untreated) 34/100 Ar (pretreated)	[57]
	195 °C, 6 min	Acidic	Wheat straw	42%	-	[63]
	180 °C, 15 min	Acidic	Eucalyptus alkali lignin	26%	Total β-O-4 61.2/100 Ar (untreated) 30.8/100 Ar (pretreated)	[66]
Organosolv (Methyl isobutyl ketone, Ethanol, Water, Sulfuric acid)	140 °C, 56 min	Acidic	Mixed hardwoods	97%	-	[67]
	160 °C, 56 min	Acidic	Mixed hardwoods	98%	-	[67]
	140 °C, 56 min	Acidic	Switchgrass	64%	Total β-O-4 37/100 Ar (untreated) 17/100 Ar (organosolv lignin)	[68]
	160 °C, 56 min	Acidic	Switchgrass	73%	4/100 Ar (organosolv lignin)	[68]

Table 1 (continued)

	190 °C, 1 h	Acidic	Switchgrass	61%	Total phenolic OH 0.86 mmol/g of lignin (untreated) 1.90 mmol/g of lignin (organosolv treated) 1.65 mmol/g of lignin (organosolv lignin)	[69]
Alkaline green liquor (Sodium sulfide & Sodium carbonate)	160 °C, 60 min	Alkaline	Poplar	18%	-	[70]
	140 °C, 60 min	Alkaline	Miscanthus	54%	-	[70]
	170°C, 3 h	Alkaline	Oil palm frond	-	Nitrobenzene oxidation products Vanillin 0.5%, Syringaldehyde 0.7% (Kraft) Vanillin 0.7%, Syringaldehyde 1.5% (Soda)	[71]
	190°C, 1 h	Alkaline		-	Vanillin 1.6%, Syringaldehyde 4.6% (Organosolv)	
Mild alkaline (Sodium hydroxide; 15% soda charge)	175 °C, 15 min	Alkaline	Eucalyptus	36%	-	[72]
	175 °C, 15 min	Alkaline	Sugarcane bagasse	80%	-	[72]
Ionic liquid and alkaline	110 °C, 12 h 75 °C, 3 h		Poplar	-	Relative abundance of β-O-4 79.0% (untreated); 80.2% (pretreated)	[73]
Ionic liquid	110 °C, 12 h		Birch	26%	Total β-O-4 65/100 Ar (cellulolytic lignin) 56/100 Ar (ionic liquid lignin) 59/100 Ar (residual lignin)	[74]

FIGURES

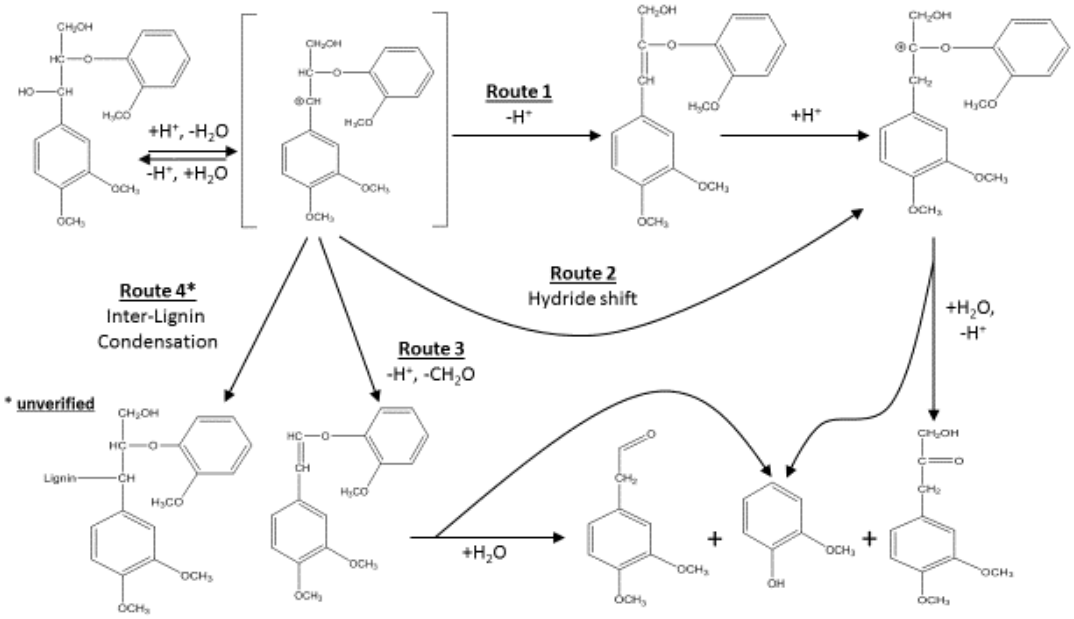

Figure 1. Acidolysis reaction pathways of model non-phenolic β-O-4' lignin dimers

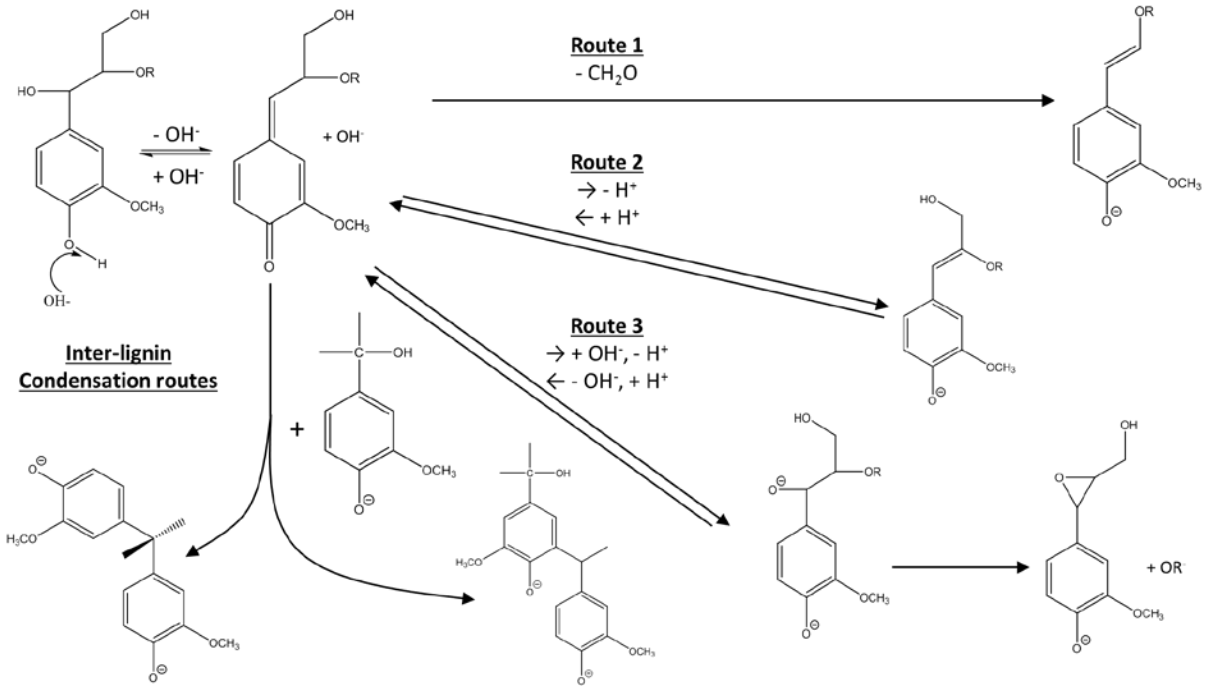

Figure 2. Hydroxide ion cleavage reaction pathways of phenolic β-O-4' lignin structures

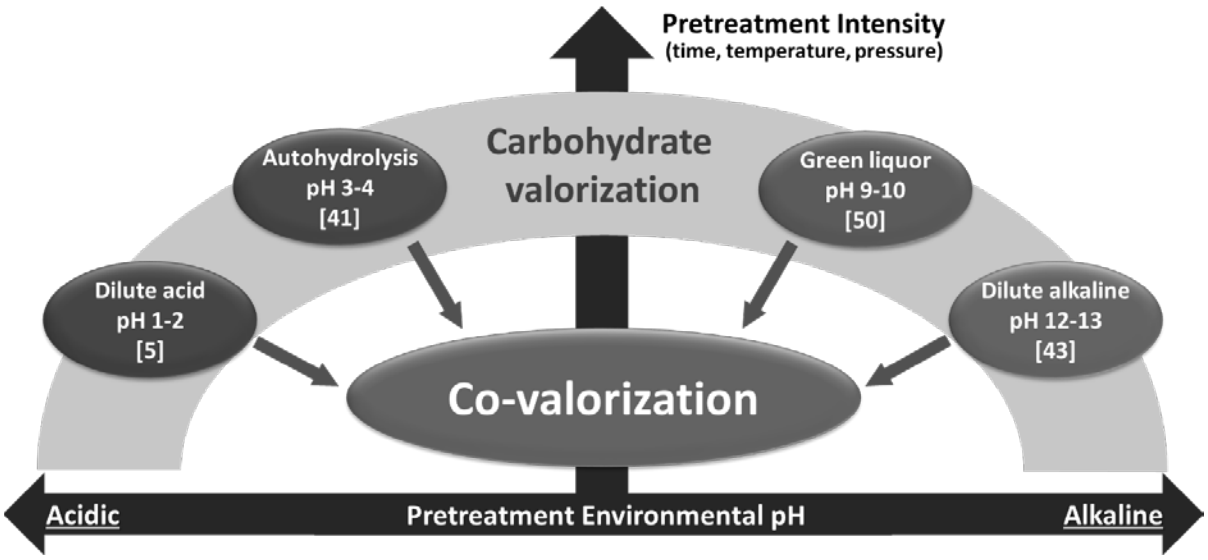

Figure 3. Pretreatment pH and intensities versus valorized fractions of lignocellulosic biomass

CHAPTER 2. Lignocentric analysis of a carbohydrate-producing lignocellulosic biorefinery process

Robert H. Narron, Qiang Han, Sunkyu Park, Hou-min Chang, Hasan Jameel

Bioresource Technology (2017), 241, 857-867

Abstract

A biologically-based lignocellulosic biorefinery process for obtaining carbohydrates from raw biomass was investigated across six diverse biomasses (three hardwoods & three nonwoods) for the purpose of decoding lignin's influence on sugar production. Acknowledging that lignin could positively alter the economics of an entire process if valorized appropriately, we sought to correlate the chemical properties of lignin within the process to the traditional metrics associated with carbohydrate production- cellulolytic digestibility and total sugar recovery. Based on raw carbohydrate, enzymatic recovery ranged from 40-64% w/w and total recovery ranged from 70-87% w/w. Using nitrobenzene oxidation to quantify non-condensed lignin structures, it was found that raw hardwoods bearing increasing non-condensed S/V ratios (2.5-5.1) render increasing total carbohydrate recovery from hardwood biomasses. This finding indicates that the chemical structure of hardwood lignin influences the investigated biorefinery process' ability to generate carbohydrates from a given raw hardwood feedstock.

1. Introduction

In the name of sustainability, researchers and innovators across the international community have been investigating and developing processes which generate valuable chemicals, fuels, and materials from renewable lignocellulosic biomass feedstocks in place of presently consumed fossil resources.[1-7] Biologically-based biorefinery processes employing cellulolytic enzyme operations for depolymerization of biomass' structural carbohydrates have been proven capable of producing monomeric carbohydrate products that can be sold as is or fermented to generate higher-value chemicals such as ethanol, lactic acid, succinic acid, among several others.[8-10] Intrinsic to lignocellulosic biomass are varying degrees of recalcitrance to said enzyme systems, which must be conquered through a series of unit operations referred to as pretreatment. Enzymatic digestibility of lignocellulosic biomasses can be significantly improved by pretreatment-derived removal and/or structural disruption of hemicellulose and lignin. We have thoroughly investigated a series of biorefinery operations which yield high carbohydrate recovery from raw lignocellulosic biomass, with the added benefit of said carbohydrate recovery being robust to multiple lignocellulosic feedstocks. The process investigated by this work is sequential autohydrolysis and mechanical refining prior to enzymatic hydrolysis by cellulolytic enzyme cocktails, selected for study due to the favorable financial prospects generated from modeling and technoeconomic analysis of similarly proposed industrial-scale biorefinery processes[11-12]. Unique to our effort, however, is due diligence towards lignin in proper recognition of lignin's importance to process economics. Pursuit of lignin-derived profit streams has recently been agreed upon as a viable salve for overall biorefinery economics for three

reasons; 1) the profitability from carbohydrates alone does not sufficiently justify investment in an industrial-scale process, 2) the existence of lignin-enriched process streams present after carbohydrate isolation beckon for utilization, and finally, 3) the myriad of potential applications where lignin's chemical properties are well suited for incorporation.[13-16]

Autohydrolysis is a capital-effective method for reducing biomass recalcitrance because of its simplicity and effectiveness when compared to more complex and capital-intensive pretreatment technologies.[17-19] Employing the triumvirate of water, temperature, and pressure, autohydrolysis catalyzes cleavage of hemicellulosic acetyl esters to produce a mildly acidic aqueous solution for facilitating the deconstruction of biomass. Elevated temperature and pressure additionally generates hydronium ions which act in unison with acetic acid to facilitate hydrolysis of acid-labile bonds within the hemicellulose and lignin fractions. In contrast to the glut of literature concerning solubilization and characterization of hemicellulose-derived compounds during autohydrolysis, limited data is published regarding how lignin is modified across feedstock varieties.[20-23]

When appropriately situated within a pretreatment sequence, mechanical refining of pretreated lignocellulosic biomass has been shown to enhance the enzymatic conversion of structural polysaccharides at mild enzyme dosages.[24-25] Mechanical refining of lignocellulosic biomass provides morphological changes to biomass fibers in the form of external fibrillation and internal delamination- all of which favorably enhance cellulase's productive interactions with biomass.[26-27] By coupling mechanical refining with autohydrolysis, dampening of autohydrolysis intensity is enabled, leading to a lesser extent of pentose and hexose degradation. Besides decreased amounts of carbohydrate degradation,

reducing pretreatment intensity can aid in-house lignin valorization efforts by reducing the extent to which devaluing inter-lignin condensation reactions will occur.[28]

Three hardwoods (maple, sweetgum, nitens) and three nonwoods (sugarcane bagasse, wheat straw, and switchgrass) were selected for testing to represent a wide array of candidate feedstocks for industrial biorefinery utilization. These raw biomasses were individually subjected to sequential autohydrolysis, mechanical refining, and cellulolytic enzyme hydrolysis. As previously discussed, the extent of lignin's influence on carbohydrate production within such a process remains unclear. Therefore, the goal of this work was not only to complete a full mass balance of carbohydrates and their derivatives within the process, but to also carefully note lignin's relocations and abundances. We seek to establish a better understanding of how botanically-diverse lignins respond to the process under investigation. Specifically, we aim to uncover further discoveries concerning lignin's influences upon enzymatic digestibility of autohydrolysis-pretreated solids and total carbohydrate recovery from raw biomass.

2. Materials and methods

2.1. Raw materials

The moisture content of each sample was measured by the weight difference before and after drying in a convective oven at 105 °C until constant mass was achieved. All raw materials were air-dried at room temperature for two weeks to establish constant moisture content. Hardwood chips from maple (*Acer rubrum*) and sweetgum (*Liquidambar styraciflua*) were provided by the Tree Improvement Program at North Carolina State University (Raleigh, NC,

USA). Nitens (*Eucalyptus nitens*) woodchips were graciously supplied by Celulosa Arauco y Constitución (Arauco, Chile). All wood chips were screened according to Scandinavian Pulp, Paper, and Board Testing Committee SCAN-CM 40:01, collecting all chips below (and including) the large accept chip classification for experimentation. The non-woody biomasses: sugarcane bagasse, wheat straw, and switchgrass (*Panicum virgatum*), were each obtained from unique sources. Our supply of sugarcane bagasse was catered by a sugarcane manufacturer located in southeastern Brazil. Wheat straw was purchased at home-improvement retailer Lowe's (Raleigh, NC, USA). Finally, switchgrass was generously gifted for experimentation from a nearby farm in southern Wake County, North Carolina. All nonwoods were hand-cut into 2-3 cm size while retaining all small particles. After cutting, all feedstocks were stored in plastic bags at room temperature prior to pretreatment. In addition, part of the feedstocks were ground by a Wiley Mill (Model No. 4, Thomas Scientific, USA), and the sawdust was screened to particle size between 20-40 mesh. The screened raw sawdust was used for compositional analysis and alkaline nitrobenzene oxidation.

2.2. Compositional analysis

The extractives content of the raw biomasses was removed using a soxhlet extraction apparatus with benzene and ethanol mixture at a ratio of 2:1 v/v for 24 hours. All samples were extracted by this method to remove any possible interference with experimental results by obstructive non-structural extractives. The chemical compositions of raw and autohydrolyzed biomasses were determined in accordance with the National Renewable Energy Laboratory (NREL) laboratory analytical procedure with modification.[29] All analyses

were performed in duplicate. Prior to experimentation, samples were dried to zero moisture content in a vacuum oven at 35 °C over fresh P_2O_5 (s) and analyzed upon reaching constant weight. ~300 mg of dry biomass was carefully weighed for analysis. The acid-soluble lignin and acid-insoluble (Klason) lignin content was determined by a two-stage acid hydrolysis (72% w/w sulfuric acid at room temperature for 2 hours, followed by dilution to 4% w/w sulfuric acid and then cooked at 121 °C for 1.5 hours). Each material's ash content was quantified by weight difference after cauterization in a muffle furnace at 575 °C for 4 hours.

To quantify carbohydrate contents, the filtrates obtained after autoclaving were neutralized to pH ~5 using $CaCO_3$ (s) and passed through 0.2 µm nylon filters. Filtrate carbohydrate concentrations were analyzed by Agilent HPLC (Agilent HPLC 1200 series, Agilent, USA), equipped with a Shodex SP081 column (8 x 300 mm, Showa Denko, Japan), and a refractive index detector operating at 50 °C. The column was operated at 80 °C using Milli-Q water as the mobile phase, which was flowing at the rate of 0.5 mL/min. Each sample's run time was 60 minutes. A five point calibration curve using commercial standard carbohydrates (glucose, xylose, galactose, mannose, and arabinose) was used to quantify carbohydrate concentrations in the filtrates. Mannose and arabinose are observed to co-elute using this HPLC system, which is accounted for. When converting between arabinose + mannose to arabinan + mannan, a molar dehydration factor of 0.89 was applied due to mannose being a hexose and arabinose being a pentose. This is an assumption that contributes error to quantification of arabinan and mannan, especially in the case when the concentrations of these co-eluting carbohydrates are skewed towards either arabinose or mannose alone.

2.3. Autohydrolysis pretreatment and post-treatment mechanical refining

Autohydrolysis pretreatment was carried out in a 1.0 L alloy reactor with an agitator (Parr Model C-276, Parr Instrument Company, USA). For each run, ~50 air-dry g of biomass was loaded into the reactor and supplemented with the appropriate amount of deionized water (accounting for moisture content) to set the starting liquid to solids ratio to 10:1. This ratio was selected to ensure effective mixing within the lab-scale reactor. In effort to enhance mass and heat transfer, air in the reactor was removed via vacuum prior to heating. The target temperature was set at 180 °C and the average temperature ramp-up time was 30 minutes. Target temperature was maintained for 40 minutes, followed by rapid temperature quenching in an ice water bath. After sufficient cooling, solid and liquid were separated by vacuum filtration using Whatman 40 grade filter paper. Residual autohydrolyzate liquid on the biomass was squeezed out using cheese cloth and removed using vacuum filtration. Squeezed solids were then extensively washed using fresh deionized water, centrifuged, and finally refrigerated at 4 °C.

Solid recovery yield was determined by measuring the total wet weight and the moisture content of autohydrolysis-pretreated samples. Extensively-washed autohydrolyzed solids were next disintegrated by a disc refiner (Bauer 148-2, Bauer Brothers Co., USA) at 0.005 inch opening, followed by fibrillation in a PFI refiner (Hamjern Maskin A/S, Norway) at 10% w/v consistency for 6000 revolutions. The refined solids were centrifuged, fluffed, and refrigerated at 4 °C.

Total carbohydrate quantities in each autohydrolyzate was determined according to the NREL laboratory analytical procedure.[30] First, an aliquot of autohydrolyzate is hydrolyzed

with 4% w/w sulfuric acid at 121 °C for one hour. After cooling, the acidic solution was neutralized to pH ~5 using $CaCO_3$ (s), followed by filtration through 0.2 μm nylon filters. Carbohydrate concentrations were then measured using the same HPLC system as described in Section 2.2.

The organic byproducts in autohydrolyzate (acetic acid, formic acid, furfural, and 5-hydroxymethylfurfural) had their concentrations determined using a different HPLC system (Dionex UltiMate 3000, Sunnyvale, CA). The system is equipped with a Bio-Rad Aminex HPX-87H column (300 mm x 7.8 mm) and operates at 65 °C using 0.05 M sulfuric acid as a mobile phase, flowing at a rate of 0.6 mL / min. Commercial standards were utilized to produce five-point calibration curves for each quantified byproduct.

2.4. Enzymatic hydrolysis

Enzymatic hydrolysis was carried out in an environmental incubator shaker (New Brunswick Scientific, USA) operating at 50 °C and 180 rpm. The total incubation time for enzymatic hydrolysis was 96 hours. An enzyme cocktail of Cellic CTec 2 and Cellic HTec 2 (Novozymes, USA) was used. The activity of CTec 2 was determined to be 139 FPU/g, determined using a published method.[31] Hydrolysis was performed at 5% w/v consistency in a 50 mM acetate buffer (pH = 4.8) with an enzyme dosing of 5 FPU per dry g substrate. CTec 2 was supplemented with HTec 2 at the amount of 1/9 v/v the dosage of CTec 2 as prescribed by the enzyme manufacturer. A miniscule peppering of sodium azide (0.1% w/w) was also added to function as antibiotic. After 96 hours, hydrolysis was ceased by immersion of the reaction tubes in boiling water for 10 mins. The reaction tubes were then vortexed and

centrifuged at 20,000 rpm (Avanti J-E High-Speed Centrifuge, Beckman Coulter, USA). To ensure enzyme removal from the substrate, the centrifuged solids were twice re-suspended in fresh acetate buffer and vortexed. After centrifugation and separation of the fresh buffer supernatant, the buffer-washed solids were then twice re-suspended in fresh deionized water, vortexed, and separated by centrifugation. These extensively-washed solids were then refrigerated at 4 °C. Part of solids were air-dried and ground by Wiley Mill and screened to a particle size of 20-40 mesh. The screened autohydrolyzed sawdust was used for compositional analysis and alkaline nitrobenzene oxidation.

2.5. Quantification of lignin's *para*-coumarate and ferulate esters

Quantification of *para*-coumarate and ferulate ester contents of lignin was performed as described in publication.[32] Approximately one gram of extractive-free sawdust was carefully weighed before being immersed in an environmental incubator containing 50 mL of 1N NaOH. Air in the incubator was purged out with nitrogen gas after solid loading. The incubator was shaken at a rate of ~200 rpm for 48 hours at room temperature. After completion, a known quantity of cinnamic acid was added to the mixture as an internal standard followed by vigorous mixing. The samples were then centrifuged to separate solid from liquid, and the liquid was acidified to pH 2.5 using 2N HCl. The acidified liquid was next extracted three consecutive times with 10 mL of ethyl acetate to isolate *para*-coumaric acid, ferulic acid, and the internal standard. Ethyl acetate extracts were then combined and filtered through 0.2 μm nylon filters prior to injection into HPLC. The HPLC system used to determine the product concentration (Dionex UltiMate 3000, Sunnyvale, CA) was equipped with a Zorbax SB-C3

column, a gradient pump, and a UV detector. A 30 min gradient was applied, utilizing two mobile phases of 10 mM formic acid in acetonitrile and 10 mM formic acid in water.

2.6. Alkaline nitrobenzene oxidation

An established protocol was applied for alkaline nitrobenzene oxidation of the examined biomass substrates.[33] Alkaline nitrobenzene oxidation was performed in a stainless steel bomb reactor and electronically heated in an aluminum block. Extractive-free raw and autohydrolyzed biomass was milled and screened as described in section 2.1. ~200 mg of screened sawdust was carefully weighed and mixed with 7 mL of 2 N NaOH solution and 0.4 mL of nitrobenzene in the bomb reactor and heated to 170 °C and held at temperature for 2.5 hours. An additional cooking of the non-woody biomasses was performed at 190 °C for 4 hours to more-accurately quantify *para*-hydroxybenzldehyde (H) yields from the nonwoods.[34] During cooking, the bomb reactors were carefully shaken every 30 min. Once the appropriate reaction time elapsed, the bomb reactor's temperature was rapidly quenched in an ice water bath, followed by manual addition of 1 mL of internal standard solution (5-iodovanillin in 1,4-dioxane; 16 mg/mL). The mixture was first extracted with 30 mL of CH_2Cl_2 three times to remove residual nitrobenzene. The aqueous phase was then acidified with 2 N HCl to pH ~2.5. The resulting acidic solution was next extracted with 30 mL of CH_2Cl_2 three additional times. Organic extracts from acidic solution were combined and dried over anhydrous Na_2SO_4. Two milliliters of the organic solution was dried by rotary evaporator at room temperature under reduced pressure. The obtained dry solids were next dissolved in 50 µL of pyridine and 50 µL of N,O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) and mildly agitated for 15 mins at 50

°C to ensure derivatization. Finally, the solution of trimethylsilylated products was injected into the GC. A five point calibration curve using commercial standard nitrobenzene oxidation products (syringaldehyde, syringic acid, vanillin, vanillic acid, *para*-hydroxybenzaldehyde) were similarly derivatized and used to quantify product yields. Results are expressed in units of mmol/g lignin. S/V was determined by the molar ratio of the summed yields of syringaldehyde and syringic acid divided by the summed yields of vanillin and vanillic acid. For non-woods, H/V was also determined from the molar ratio of the yield of *para*-hydroxybenzoic acid divided by the summed yields of vanillin and vanillic acid. All analyses were performed in duplicate.

3. Results and discussion

3.1. Chemical composition of raw materials

The chemical composition of the investigated raw materials is listed in Table 1 (units: g/100g oven-dry (OD) raw (biomass)). The total lignin content in the nonwood biomasses was found to be lesser (~20-22 g/100g OD raw) when compared to the hardwood species (~24-25 g/100g OD raw). The total carbohydrate content of all feedstocks ranged from ~59-67 g/100g OD raw. An unknown portion of the quantified glucan is known to be derived from starch, especially in nonwood biomasses.[35] Inorganic ash content in the nonwoods was significantly more pronounced (~3-6 g/100g OD raw) versus the hardwood species (~0.5 g/100g OD raw). These results showcase the increased biological inclusion of inorganic components into nonwoody biomasses compared to the tested woody biomasses. An

exhaustive review of biomass-derived ashes from wood and nonwood highlighted raw biomass ash content values that are in agreement with the raw ash quantities recorded.[36]

Sorely, repeated composition analyses continued to leave notable quantities of unidentified components. For hardwoods, the quantity of unidentified material ranged from ~12-14 g/100g OD raw. For the nonwoods, the quantity of unidentified components value was lower than hardwoods but equally pestilent due to its magnitude (~6-11 g/100g OD raw). These results absolutely indicate impartial raw biomass characterization. However, the quantity of unidentified components within the tested biomasses can be lowered through specific quantitation of biomass components which are not accounted for during the employed compositional analysis protocol. The following are known components of lignocellulosic biomasses which were not quantified in this work: acetyl, glucuronic acid, galactouronic acid, rhamnose, and the pentoses and hexoses which dehydrated to furfural, 5-hydroxymethylfurfural, formic acid, and levulinic acid during compositional analysis.

Unique to nonwood biomasses is the existence of *para*-coumarate and ferulate moieties linking various points between lignin as well as between lignin and carbohydrates. These linkages consist of ester and benzyl ether bonds that exhibit susceptibility to cleavage under acidic conditions. Table 3.1 also contains information regarding quantification of *para*-coumarate and ferulate moieties in the raw nonwood biomasses (units: % w/w of total lignin). Primarily, it can be seen that these components exist in varying amounts within the tested nonwood biomasses and are not detectable in the hardwood control. Within the nonwoods, a greater amount of *para*-coumaric acid was detected over ferulic acid in sugarcane bagasse and switchgrass, whereas the opposite trend was observed in wheat straw.

Utilizing the same means of quantification, it was previously reported that genetically unmodified switchgrass samples contain ~25% w/w *para*-coumarate and ferulate moieties per gram of lignin at a coumarate to ferulate ratio of 4:1.[32] A publication involving wheat straw found that solubilized whole cell-walls from wheat straw bear 4 *para*-coumarate and 11 ferulate structures per 100 aromatic rings, determined through integration of appropriate signals within 2D-HSQC nuclear magnetic resonance spectra.[37] For sugarcane bagasse, a publication studying milled wood lignin from raw sugarcane bagasse reported to 48 *para*-coumarates and 4 ferulates per 100 aromatic rings, determined using a similar 2D-HSQC nuclear magnetic resonance method.[38] Despite differences in methods of characterization and reported units between the cited publications and the present work, our quantification of *para*-coumarate and ferulate moieties in the selected raw nonwood biomasses are of similar magnitudes as well as of similar coumarate-to-ferulate ratios when compared with the aforementioned publications.

3.2. Autohydrolysis of botanically-diverse biomasses

3.2.1 Composition of autohydrolyzed solids and recovery of components

Table 3.2 provides the compositional analysis of the solid fractions treated by autohydrolysis. Decomposition of nonwood biomass was observed to be stronger than that of hardwood biomass, interpreted from nonwood's lower solid recovery after autohydrolysis. Xylan, typically assigned to broadly represent hardwood and nonwood hemicellulose, dominated the loss of solids across all feedstocks. Residual solid xylan quantities ranged from ~4-5 g/100g OD raw in the nonwoods and ~3-4 g/100g OD raw in hardwoods.

Recovery of xylan (based on raw xylan content) ranged from 17-23% w/w in nonwoods, and in the hardwoods, 18-24% w/w. The similar ranges of residual xylan and xylan recovery indicates that the extent of xylan degradation during autohydrolysis is more related to pretreatment severity instead of raw biomass. It must be emphasized that xylan, albeit the majority, is only one component of hemicellulose. Arabinan and mannan, additional carbohydrates comprising hemicellulose, were also present in differing amounts following autohydrolysis, depending upon raw biomass. For nonwood biomasses, total residual arabinan and mannan quantities were ~0.3-0.4 g/100g OD raw. In the hardwoods, this quantity was found to be higher: 0.6-1.6 g/100g OD raw. Percent recovery of arabinan and mannan (based on raw) of nonwoods ranged a mere ~10-15% w/w, providing additional proof of significant nonwood hemicellulose degradation during autohydrolysis. For the hardwoods, arabinan and mannan recoveries (based on raw) ranged from 43-44% w/w, with 73% w/w recovery from sweetgum serving as an outlier. Interestingly, the quantity of arabinan and mannan in each raw biomass was found to be somewhat similar, with nonwoods bearing ~3 g/100g OD raw and hardwoods ~1-3 g/100g OD raw. Despite the similar quantities of these hemicellulosic sugars in raw hardwood and nonwood, a notable difference can be seen in the recoveries of these carbohydrates in the solid fractions after autohydrolysis. This implies that there exists an arabinan-rich and/or mannan-rich portion of hardwood hemicellulose that resists degradation during autohydrolysis. This observation can be interpreted as a minor display of durability from hardwood hemicelluloses at the tested autohydrolysis conditions. Alternatively, it is possible that the increased recovery of arabinan and/or mannan in hardwoods indicates preferential degradation of lignin, casting a

preservation effect upon these portions of the hemicellulose structure. Galactan, ranging from 0.7-1.1 g/100g OD raw across all tested raw biomasses, was not detectable in any of the autohydrolyzed solids. Therefore, surviving galacto-saccharides in autohydrolyzate are the only descendants of raw galactan that can be recovered from the tested raw lignocellulosic biomasses after autohydrolysis.

Residual glucan and associated raw glucan recovery was found to be dissimilar when comparing the hardwoods and nonwoods. For the hardwoods, residual glucan ranged from ~38-43 g/100g OD raw with associated percent recoveries (based on raw) ranging from 93-98% w/w. These quantities indicate minimal damage to hardwood cellulose during autohydrolysis. In fact, when considering the previously established degradation effect that autohydrolysis has on hemicellulose, it is more likely that the un-recovered glucan in hardwood biomasses originated from glucose residues present within glucomannan as opposed to glucose residues from cellulose. For nonwoods, the quantity of residual glucan and their associated recoveries ranged from ~30-37 g/100g OD raw, corresponding to ~80-87% w/w raw glucan recovery. The additional ~10-20% w/w loss of glucan in nonwoody biomasses (compared to hardwoods) can be similarly attributed to the presence of glucomannan hemicelluloses. However, there exists additional conflict when considering the effects of starch being present in the raw material. Assuming solid starch is fully hydrolyzed at the tested autohydrolysis conditions, solid glucan recovery for autohydrolyzed nonwood solids could be shunted where the quantified raw glucan has been infiltrated by the presence of starch.

Concerning lignin, it is of significant interest that despite the lower lignin content in the raw nonwoods compared to the raw hardwoods, ~30-36% w/w delignification was observed in the nonwoods versus ~15-19% w/w of the hardwood lignin. This observation indicates that the differences in raw lignin (protolignin) structures in hardwood and nonwood biomasses are significant contributors to lignin's reactive behavior during autohydrolysis. The observed delignification percentages are similar in magnitude to those reported for autohydrolysis of comparable biomasses at conditions close to those utilized in this work.[39-42] Substantially more involved discussion of lignin's behavior during autohydrolysis is conducted in the latter half of this manuscript.

The amount of unidentified components in autohydrolyzed solids is also generally lower when compared to their respective raw biomasses. This effect is more pronounced in hardwoods, with unidentified quantities decreasing from ~12-14 g/100g OD raw in the raw hardwood to ~3-5 g/100g OD raw in autohydrolyzed hardwood solids. For nonwoods, unidentified components decreased from ~6-11 g/100g OD raw in raw biomass to ~4-7 g/100g OD raw in autohydrolyzed solids. From these observations, it is expected that the elusive unidentified components in the tested raw biomasses can be quantified within the autohydrolyzate produced by pretreatment. As previously mentioned, the amount of unidentified components within any lignocellulosic biomass can be decreased with extra characterization of the components that were not counted during the applied compositional analysis protocol.

After contemplating the differences in all solid component recoveries from hardwoods and nonwoods after autohydrolysis, it can be stated that the tested nonwood biomasses are

more susceptible to degradation by autohydrolysis when compared to the tested hardwoods. This claim is supported by nonwood's lower solid yield, lower glucan, arabinan, and mannan recoveries, and finally, enhanced delignification. When considering possible explanations for these observations, the most notable difference between all tested raw biomasses is the pronounced presence of ash in raw nonwoods compared to raw hardwoods. For the tested nonwoods, ash contents ranged from ~3-6 g/100g OD raw. These inorganics could be providing an unknown effect which enhances nonwood degradation during autohydrolysis. Complete understanding of the effects that these nonwood-derived ashes have on autohydrolysis will require specific characterization of these inorganic components. As a reminder: ash content is simply quantified in this work as the post-cauterized mass relative to the sample's starting mass.

3.2.2. Carbohydrates released into autohydrolyzate

The total quantity of carbohydrates produced in each autohydrolyzate is also included in Table 3.2 (subheader "liquid"). Total glucan (measured as glucose) concentrations in autohydrolyzate ranged from ~2-3 g/100g OD raw for nonwoods and ~1-2 g/100g OD raw for the hardwoods. As previously discussed, the origin of quantified glucan can be confused between minor hemicelluloses, starch, and cellulose. The greater range of glucan in nonwood autohydrolyzates is in support of a greater presence of starch in raw nonwoods and the permeating effect it has upon glucan quantification. The concentrations of glucan determined in hardwood autohydrolyzates represented ~1-2% w/w of raw glucan, with sweetgum again exhibiting outlier behavior with ~6% w/w glucan recovery in

autohydrolyzate. In line with previous discussion of hardwood hemicelluloses, it can be speculated that the hemicellulose structure of sweetgum responds to autohydrolysis differently when compared to the behaviors of maple and nitens hemicelluloses. This claim is derived from the combined observation that sweetgum displays higher arabinan and mannan recovery in autohydrolyzed solids and greater glucan recovery in autohydrolyzate compared to the other tested hardwoods.

Total xylan (measured as xylose) in autohydrolyzate from all biomasses ranged from ~9-11 g/100g OD raw, with switchgrass serving as an outlier at ~14 g/100 g OD raw. Despite similarities in concentrations, it should be emphasized that these quantities embody different quantities of the original xylan in raw biomass. Xylan recovery in autohydrolyzate for nonwoods ranged from only 50-58% w/w compared to hardwoods 66-75% w/w. The decreased xylan recovery from nonwoods reveals a greater extent of xylose degradation during autohydrolysis to produce dehydrated pentoses. Quantification of dehydrated carbohydrate molecules and ensuing discussion can be found in the next section, 3.2.3. The analyzed minor carbohydrates: galactan (measured as galactose), arabinan (measured as arabinose + mannose), and mannan (measured as arabinose + mannose) were each quantifiable in all autohydrolyzates. When considering recoveries of minor carbohydrates, sweetgum again exhibited outlier behavior when compared to the other hardwoods. Excluding sweetgum, galactan recovery ranged from 80-98% w/w across all biomasses, and arabinan and mannan recovery ranged from 45-63% w/w. Uniquely, sweetgum autohydrolyzate bears only 68% w/w of raw galactan and 24% w/w of raw arabinan and mannan. These observations are all in support of deeming sweetgum's raw hemicellulose

structure as a contributing factor to its unique residual hemicelluloses following autohydrolysis.

Total carbohydrate recovery in autohydrolyzate ranged from 21-30% w/w in nonwoods and 20-24% w/w in the hardwoods. Autohydrolysis is only intended as a pretreatment for increasing carbohydrate recovery in downstream unit operations. However, it can be seen that approximately a fourth of all raw carbohydrate is recoverable in autohydrolyzate. This significant quantity calls for treating autohydrolyzate as an additional xylan-rich stream for a biorefinery that must be accounted for when determining total carbohydrate recovery from a lignocellulosic biorefinery process. This sugar stream has already come under investigation as a valorization target for a novel biorefinery process focusing upon establishing multiple revenue streams.[43]

3.2.3. Organic byproducts present in autohydrolyzate

In response to clear experimental evidence of carbohydrate degradation occurring during autohydrolysis, the concentrations of furfural and 5-hydroxymethylfurfural were determined in all autohydrolyzates. Furfural is considered a representative molecule of pentose degradation and 5-hydroxymethylfurfural representative of hexose degradation. However, 5-hydroxymethylfurfural is an unstable intermediate which will further degrade into levulinic acid and formic acid.[44] For this purpose, formic acid concentrations were measured in addition to 5-hydroxymethylfurfural concentrations. Acetic acid concentrations in autohydrolyzate were measured to monitor the surviving quantity of original hydrolysis

reactant. The purpose in quantifying the four aforementioned molecules is to further describe the components of autohydrolyzate solution.

Figure 3.1 visualizes the quantities measured of the four said degradation products, in addition to the measured autohydrolyzate pH after cooling. Furfural quantities in autohydrolyzate were found to range from ~1-2 g/100g OD raw for nonwoods. For each tested hardwood, furfural quantities were ~1 g/100g OD. The approximately double quantity of furfural in nonwood autohydrolyzate supports previous speculation that pentose dehydration occurred to a greater extent during autohydrolysis of nonwoods. Still, the presence of furfural across all autohydrolyzates demonstrates that undesired dehydration reactions of pentoses cannot be avoided, even when utilizing reduced autohydrolysis intensity as afforded by subsequent mechanical refining. Similar concentrations of 5-hydroxymethylfurfural and formic acid were measured across nonwood and hardwood autohydrolyzates: formic acid, 0.3-0.4 g/100g OD raw; 5-hydroxymethylfurfural, 0.1 g/100g OD raw. The presence of minor amounts of these compounds indicate degradation of hexoses as well, although glucose (hexose), galactose (hexose), and mannose (hexose) each exist in autohydrolyzate at relatively low concentrations compared to xylose (pentose).

The pH of each autohydrolyzate was below the pKa of acetic acid (4.8), with values ranging from 3.3-4.3. Hardwood autohydrolyzate pH values ranged from 3.3-3.6 and correlate with the concentrations of acetic acid and formic acid present. Nonwood autohydrolyzate pH values also correlate with acetic acid and formic concentrations, in spite of the higher pH values observed (pH 3.6-4.3). A contributor to the elevated pH value of nonwood autohydrolyzates is likely the previously discussed ash content of raw nonwood,

which decreased by ~50-90% w/w after autohydrolysis (Table 3.2). Soluble ash is known to exhibit a buffering effect on the pH of the autohydrolyzate due to the presence of various counter-ions.[45] The buffering effect of the solubilized inorganics is especially profound when considering the additional acidity capable of being delivered by the carboxylic acids *para*-coumaric acid and ferulic acid (pKa = ~3.5-4.5). Exclusive to nonwood biomass autohydrolyzates, these carboxylic acids are produced from cleavage of native coumarate and ferulate esters (quantified in raw biomass, Table 3.1). Therefore, the observation that nonwood autohydrolyzate pH values are greater (compared to hardwood autohydrolyzates) is a profound testament to the buffering capabilities of solubilized nonwood-derived ash, especially given the extra potential for acidity.

3.3. Enzymatic hydrolysis of autohydrolyzed and mechanically-refined biomass

After autohydrolysis, all pretreated pulps were subjected to 6000 revolutions of mechanical refining using a PFI pulp refiner. The importance of refining cannot be stated enough; utilization of a refining step enables employment of a reduced autohydrolysis severity, which is shown to recover the xylans that are not recoverable at increasing autohydrolysis severities.[46] Refining is able to deliver the enzymatically-advantageous aspects of pretreatment to solid biomass without extensive pentose dehydration to furfural. Following refining, enzymatic hydrolysis was performed utilizing a cellulase cocktail spiked with cellulase and hemicellulase enzymes. The enzyme dosage applied to the autohydrolyzed and mechanically refined solids was 5 FPU / OD g solid- an intentionally low cellulase dosage selected to demonstrate the efficacy of the applied pretreatment process

for producing carbohydrates. Enzymatic digestibility of pretreated solids was used to determine enzymatic carbohydrate yields in terms of total raw carbohydrate, Figure 3.2.

Total carbohydrate digestibility was lowest in maple and switchgrass at 65% w/w and 71% w/w conversion, which represented 48% w/w and 40% w/w of raw carbohydrate, respectively. Unlike the two aforementioned biomasses, all other feedstocks demonstrated total carbohydrate digestibility between 83-87% w/w, representing between 50-64% w/w of raw carbohydrates. With regards to cellulolytic enzyme digestibility of hardwoods, a similar range of digestibilities were observed across several hardwood feedstocks, albeit using a Kraft (alkaline) pretreatment.[47] The digestibility-associated yields produced by the investigated process reveals that most of the tested biomasses exhibit an encouraging alacrity to yield carbohydrates via enzymatic hydrolysis following the employed pretreatment and refining operations.

Total percent carbohydrate recovery, which couples yields of soluble carbohydrates both after autohydrolysis and after enzymatic hydrolysis, can be summated within Figure 3.2. The purpose of the summated metric is to gauge the efficacy of the entire process for producing carbohydrates from raw biomass. For nonwood biomasses, total carbohydrate recovery was similar despite differences in carbohydrate recovery in autohydrolyzate relative to enzymatic hydrolyzate. It can be seen that greater carbohydrate recovery in nonwood autohydrolyzate correlates with lower carbohydrate recovery using enzymatic hydrolysis. This correlation suggests that the autohydrolysis conditions utilized may be too severe for the tested nonwoods, leading to solubility of a significant portion amount of hemicellulose-derived carbohydrates which tend to degrade and render lower total carbohydrate recovery.

Across the hardwood biomasses, an increasing amount of sugar can be recovered during enzymatic hydrolysis, seemingly unrelated to recovery in autohydrolyzate. The cause of this observation is hypothesized to be due to differences in autohydrolyzed solid lignin composition and structural features, specifically its degree of inter-lignin condensation. It has been concluded that a more-condensed lignin exerts greater non-productive binding affinity towards cellulase enzyme systems, leading to lower solid digestibility and carbohydrate yields.[46, 48-51] To begin elucidation of how lignin's condensed chemical structures affects carbohydrate balances, nitrobenzene oxidation of the raw and autohydrolyzed biomasses was performed to determine the quantity of non-condensed lignin structures within each raw and autohydrolyzed solid lignin.

3.4. Lignocentric analysis of carbohydrate production

3.4.1 Nitrobenzene oxidation of raw and autohydrolyzed solid biomasses

The yields of non-condensed lignin structures from nitrobenzene oxidation of raw biomass and autohydrolyzed solid biomass are displayed in Table 3.3. When comparing the lignins of hardwood and nonwood, two overarching differences can be noted. Primarily, reaction products from non-condensed *para*-hydroxyphenyl (H) lignin structures are only produced by the nonwood biomasses. Secondarily, the molar sum of non-condensed structures from nonwood protolignin ranged from 1.61-2.32 mmol/g lignin, lower than the total yield from hardwood protolignin, 2.62–2.82 mmol/g lignin. This demonstrates that nonwood protolignin bears a greater amount of native condensed inter-lignin structures. The S/V ratio of non-condensed lignin structures in hardwood protolignin varied (2.5-5.1) despite

similarities in total yield. This finding reveals that the chemical structure of hardwood protolignin is dissimilar across the tested hardwoods. Said variance is likely provided by differences during lignin biosynthesis. Specifically, a hardwood species which exhibits preference for utilizing syringyl monolignols during lignin biosynthesis will inevitably have a lesser degree of C5 aromatic ring inter-lignin bonds due to the presence of the additional methoxyl group. Concerning variations across the hardwood biomasses, it was observed that maple protolignin has a significantly larger quantity of non-condensed guaicyl lignin structures (measured as V+Va) compared to sweetgum and nitens. This is especially interesting given that maple's total non-condensed structural yield was similar to the other biomasses (2.71 vs. 2.62 & 2.82 mmol/g lignin) - in spite of having the lowest S/V ratio (2.5 vs. 3.6 & 5.1). At the other end, nitens is found to have the greatest quantity of non-condensed syringyl lignin structures (measured as S+Sa) in addition to the greatest total yield and highest S/V. In all, there is evidence that the absolute quantity of non-condensed protolignin structures does not significantly vary across the hardwood species tested, however, the distribution of said structures (according to monolignol origin) absolutely depends on raw biomass.

For the raw nonwood biomasses, it is clear that the protolignin in each nonwood is unique to one another. This claim is based upon the distinguishable differences across nonwoods involving quantified non-condensed H, V, and S structures, S/V and H/V ratios, and finally, the differing total yields. Specifically, sugarcane bagasse was shown to bear the least condensed protolignin, followed by switchgrass, and pulling up the rear, wheat straw. Supporting sugarcane bagasse protolignin's non-condensed enrichment is the greatest

quantity of both non-condensed syringyl structures and non-condensed *para*-hydroxylphenyl structures (measured as H) seen in the tested nonwoods. Switchgrass protolignin, found to be more condensed that sugarcane bagasse's, uniquely bears the greatest abundance of non-condensed guaiacyl structures across the tested nonwoods. Wheat straw, housing the most condensed nonwood protolignin examined, still bears quantities of guaiacyl and syringyl non-condensed structures comparable to the other nonwood biomasses. However, the quantity of non-condensed *para*-hydroxyphenyl structures is lowest in wheat straw. This could reveal that the *para*-hydroxyphenyl monolignol structures produced in wheat straw protolignin are further condensed into the lignin structure prior to the completion of biosynthesis. Or, quite oppositely, wheat straw may simply not utilize significant quantities *para*-hydroxylphenyl monolignols during biosynthesis, rendering a protolignin structure lacking notable quantities of native non-condensed structures.

The total yield of non-condensed lignin structures from autohydrolyzed biomasses is known to decrease relative to raw, based upon the two established reaction pathways for lignin in acidic environments: 1) inter-lignin condensation , or 2) preferential hydrolysis of specific lignin moieties followed by solubilization into autohydrolyzate.[52] While both obviously decrease the quantity of non-condensed lignin structures in solid lignin, the manner by which this decrease occurs is fundamentally opposite in that one pathway "builds up" lignin and the other breaks it down. Inter-lignin condensation, previously identified as counterproductive to cellulolytic hydrolysis, is known to occur more readily with strong acid catalysts. The key advantage to autohydrolysis is that our active ingredient, natively-spawned acetic acid, is a weak acid. The presence of weak acid at elevated temperatures and

pressures is hypothesized to partially shift lignin reactivity towards preferential hydrolysis as opposed to inter-lignin condensation. While the occurrence of inter-lignin condensation likely cannot be avoided outright due to the acidic conditions employed, the extent is predicted to be reduced by utilizing sequential low-intensity autohydrolysis and mechanical refining compared to greater intensity autohydrolysis alone. Additionally, if the presence of solubilized lignin is proven in the future to negatively impact downstream carbohydrate production by cellulolytic hydrolysis, then this penalty can be overcome by washing autohydrolyzed biomass to relieve autohydrolyzate from the autohydrolyzed solid pulp.

When considering the changes to solid lignin's S/V and H/V seen after autohydrolysis, no clear trends can be extrapolated. For example, S/V of maple increases from 2.5 to 2.7, however, sweetgum sees its S/V decrease from 3.6 to 3.0 following autohydrolysis. For nonwoods, H/V increased, decreased, and remained constant- dependent upon biomass. The S/V did increase slightly across all the nonwoods after autohydrolysis, but drastically only in the case of sugarcane bagasse (1.5 to 2.1). All the aforementioned discrepancies are believed to be a product of the opposite reaction pathways for lignin to follow during autohydrolysis. If a structurally-specific reactivity trend were to have been found in the data, then it would have been possible to hypothesize upon which non-condensed lignin structure is most affected by autohydrolysis and by what reaction it is predominantly affected by. However, based upon the vague changes to S/V and H/V presented, it can only be said that both reaction pathways are being followed to an unspecified degree in each biomass. Disadvantageously, we are without an essential third element of data that could potentially illuminate unseen trends: S/V and H/V of the lignin soluble in autohydrolyzate.

As expected, a decrease in total yield of non-condensed structures (relative to the quantity in raw biomass) was observed, especially in the tested hardwoods. Inter-lignin condensation can be clearly observed yet not exactly quantified. Take into consideration the significant decrease in non-condensed guaicyl structures witnessed in maple, previously observed to bear the greatest quantity of non-condensed guaicyl structures within its protolignin. With a relative decrease of 41% following autohydrolysis (sweetgum: 21%, nitens: 46%), this loss is also seen to slightly shift the S/V of maple lignin upward (2.5 to 2.7). Furthermore, maple also produced the lowest sugar recovery from cellulolytic enzyme hydrolysis (48% w/w carbohydrate recovery versus 60% & 64% w/w carbohydrate recovery). This finding is in support of extended guaiacyl inter-lignin condensation occurring in maple, effectively penalizing carbohydrate yields using cellulolytic hydrolysis. In the case of nitens, there is actually greater relative decrease in non-condensed guaicyl structures witnessed, but at a lesser absolute loss (nitens: 0.21 mmol V+Va/g lignin lost; maple: 0.32 mmol V+Va/g lignin lost). Unique to nitens, we also observed significant decrease in the quantity of non-condensed syringyl structures, which showed a relative decrease of 43% (maple: 36%; sweetgum: 35%). This indicates that there is preferential reactivity occurring within the non-condensed syringyl portion of nitens lignin during autohydrolysis. Furthermore, when considering solid delignification by autohydrolysis, nitens is again a slight outlier (19% w/w delignification vs. 15% & 16% w/w delignification). This finding suggests that autohydrolysis of hardwoods with higher S/V ratios will lead to preferential hydrolysis of a greater quantity of non-condensed syringyl structures, resulting in a greater degree of syringyl lignin solubilized.

For nonwoods, it can be inferred that the protolignin structure of each of the tested nonwood facilitates a feedstock-specific reactivity during autohydrolysis. We base this inference upon the unique total yields of non-condensed structures in both raw and autohydrolyzed nonwoods. For example, sugarcane bagasse starts with the most native non-condensed lignin structures (2.32 mmol/g lignin) but also retains the greatest quantity following autohydrolysis (1.87 mmol/g lignin). Wheat straw displayed the opposite of this behavior, beginning with the lowest quantity of native non-condensed lignin structures (1.61 mmol/g lignin) and concluding with the least (1.17 mmol/g lignin). For all nonwoods, the relative decrease in total yield was only 19-27%, demonstrating reduced reactivity of non-condensed structures compared to hardwoods. This observation came as a surprise as it was expected that the additional presence of *para*-hydroxylphenyl (H) lignin would facilitate a greater degree of inter-lignin condensation, therefore retention of non-condensed H structures would be low. Retention of non-condensed H structures after autohydrolysis ranged from 60-74%, displaying adequate resistance to modification (condensation or hydrolysis) in spite of their naked aromatic C3 and C5 positions. For V+Va, the percent of structures preserved after autohydrolysis ranged from 65-70%, and for S+Sa, 82-94%. From this, it can be seen that the most-affected non-condensed lignin structures in nonwood biomasses are those that yield H and V+Va during nitrobenzene oxidation. As discussed in section 3.2.3, the pH of the nonwood autohydrolyzate was observed to be slightly higher compared to the hardwoods due to greater presence of raw ash (HW: pH 3.3-3.6; NW: pH 3.6-4.3). This pH difference may contribute to the lesser extent of non-condensed lignin reactions observed across the nonwood biomasses, indicating that autohydrolyzed nonwood lignin is of greater similarity to

its protolignin compared to autohydrolyzed hardwood lignin's similarity to its own protolignin.

While it cannot be debated that the quantity of non-condensed lignin structures is lesser in all biomasses after autohydrolysis, exact delineation between inter-lignin condensation and structurally preferential hydrolysis cannot be provided from the data produced. With evidence of both reactions having occurred across all biomasses, it will be extremely important in future work to quantify non-condensed lignin structures within the lignin solubilized into autohydrolyzate. Quantification of these structures found in solubilized lignin will afford a complete molar balance of non-condensed lignin structures after autohydrolysis, enabling calculation of the recovery of non-condensed structures in autohydrolyzate. The difference between the quantity of raw non-condensed lignin structures and those measured in both solubilized and insoluble non-condensed lignin structures will indicate the extent of non-condensed inter-lignin condensation taking place at the tested mild autohydrolysis condition. The challenges associated with describing autohydrolyzate lignin lies in the current dearth of effective protocols for isolating lignin from hemicellulose-rich hydrolyzates. A standardized protocol using a solvent-free soluble lignin isolation method is currently under development after the outstanding results recently published by colleagues concerning soluble lignin segregation from carbohydrates within a biorefinery process [43].

3.4.2. Lignin properties influencing the enzymatic digestibility of autohydrolyzed and refined solids

The lignin content of autohydrolyzed solids and each respective percent delignification (as it relates to raw lignin content) can generally be considered as the most-appropriate metric of a pretreatment's efficiency when the goal is downstream carbohydrate production. Delignification is believed to be beneficial towards carbohydrate production based upon the simple explanation that lignin removal increases carbohydrate accessibility for cellulolytic enzymes. As previously noted, the tested autohydrolysis condition was shown to produce different degrees of delignification across the tested biomasses. For the tested hardwoods, raw delignification ranged from 15-19% w/w, lower than the 30-36% w/w delignification seen from the nonwood biomasses utilized. This observation is particularly interesting when considering the lower lignin content in the tested raw nonwoods (~21-22% w/w) versus the tested hardwoods (~24-25% w/w). It was hypothesized originally that the magnitude of lignin removal after autohydrolysis will positively trend with a given autohydrolyzed and refined biomass' enzymatic digestibility, again along the lines that there should be less lignin present in the biomass to inhibit enzymatic saccharification. The results revealed this hypothesis to be invalid due to no visible trend amongst hardwoods and the inverse trend being true amongst nonwoods. Unexpectedly, a slight negative trend was discovered between enzymatic digestibility of the pretreated nonwood solids and the percent delignification obtained from autohydrolysis (not shown). A possible explanation for said observation is that percent delignification fails to address changes to residual solid lignin in biomass: so for a nonwood biomass with a greater extent of autohydrolysis-borne

delignification, a greater penalty to carbohydrate yield occurs during downstream enzymatic saccharification. This indicates that the absolute quantity of lignin present is not significantly affecting carbohydrate production. Instead, it must be the properties of the lignin that exert influences.

To further probe the observed differences between hardwood and nonwood delignification, the relationship between *para*-coumaric acid and ferulic acid in raw nonwood lignin and nonwood delignification after autohydrolysis was gauged. The lability of the aforementioned structures during acidic pretreatments is well known, leading to the hypothesis that these structures are contributing to the increased delignification in nonwoods after autohydrolysis. The total quantity of *para*-coumarate and ferulate moieties within the raw nonwood biomasses showed good linear correlation with delignification (R^2=0.83). However, the presence of these moieties only exhibited a minor effect on delignification of nonwoods, despite being originally hypothesized as a major contributor to differences in delignification. This claim is based upon the limited ranges of both linearly-related metrics (30-36% w/w delignification; 2-7% w/w coumarate and ferulate in lignin).

Next, we probed for a relationship between a given autohydrolyzed lignin's non-condensed structural composition and the ability of cellulase systems to depolymerize said biomass' structural carbohydrates. The total molar yield from non-condensed lignin structures roughly correlates to the extent of inter-lignin condensation, therefore this metric is applied to symbolize a given lignin's extent of inter-lignin condensation. It was originally hypothesized that pretreated biomasses wielding greater quantities of non-condensed lignin structures would positively correlate with cellulolytic digestibility of the pretreated solids.

Upon correlation analysis, a surprising absence of discernable trends was noted for both the hardwoods and nonwoods tested. For hardwoods, it is possible to suggest that a greater abundance of non-condensed syringyl (S) structures enhances enzymatic digestibility of the pretreated solids, however, supporting such a claim is hindered by limited scatter (1.23, 1.34, 1.34 mmol S+Sa/g lignin vs. 65%, 83%, 84% w/w carbohydrate digestibility). It is extremely important to note that our utilization of total molar yields to represent the extent of inter-lignin condensation is not without its flaws. For example, we are limited to describing the molar quantities of non-condensed structures on the basis of per-gram of lignin. This is due to not knowing the average formula weight of each biomass lignin's C_9 units. Estimating degree of condensation on a completely molar basis would afford significantly greater accuracy, potentially elucidating unseen trends in the data analyzed.

In totality, our findings suggest that the enzymatic digestibility of autohydrolyzed solids is affected by far greater influences than lignin content, delignification, and total quantity of residual non-condensed lignin structures. While lignin is a major component of lignocellulosic biomasses, its influence upon enzymatic digestibility is likely most pronounced when considering lignin's effects upon the physical and morphological properties of the biomass itself.

3.4.3. Lignin properties influencing total sugar recovery from the biorefinery process

As previously defined, total percent sugar recovery is a metric which combines yields of soluble sugars present in the separate hydrolyzates produced after autohydrolysis and after enzymatic hydrolysis. The yield at which a biorefinery process can produce carbohydrates is

affected not only by the process itself, but also by incoming feedstock's tendency to fractionate. When considering separating carbohydrates from the solid matrix of lignocellulosic biomasses, it was originally hypothesized that a given raw biomass whose lignin bears a greater amount of non-condensed lignin structures would result in greater total carbohydrate recovery. This hypothesis is based on the fact that highly-condensed protolignins are proven opponents to carbohydrate recovery, due in part to their previously-noted inhibitory effects upon enzymatic saccharification, as well as their expected stability at the mild autohydrolysis conditions applied.

Concerning the hardwood biomasses, Figure 3.3 displays the trends between total sugar recoveries and the non-condensed lignin structures present in each respective raw lignin. The inverse trend between S+Sa and V+Va yields plotted against total sugar recovery exposes the effect that hardwood protolignin's non-condensed S/V has on carbohydrate recovery using the tested biorefinery process. Since the total yield of non-condensed structures in hardwoods was similar, ~2.6-2.8 mmol/g lignin, it is difficult to conclude that the absolute amount of non-condensed lignin structures plays much of a role in total sugar recovery. Instead, this suggests that the monolignol constituents, their abundances, and the manner in which they are coupled during lignin biosynthesis each contribute to total sugar recovery from hardwood biomasses using the investigated biorefinery process.

Unfortunately for the tested nonwood biomasses, the limited range of total sugar recovery (72-76% w/w) hinders the extrapolation of meaningful trends. Said range produces scatter plots which do not generate any useful relationships for interpretation. The lack of trends does suggest that such a process could merit further optimization for nonwood biomasses.

Based upon the negative relationship between delignification and enzymatic digestibility of pretreated solids (discussed in section 3.4.2), it can be suggested that greater yields of carbohydrates will be obtained by reducing the intensity of autohydrolysis upon nonwoods. This finding was surprising, as it was originally thought that the chosen conditions for autohydrolysis (180 °C, 40 min, 10:1 liquid:solid) were sufficiently mild, no matter the raw biomass.

4. Conclusions

Results reveal that the investigated biorefinery process for carbohydrate production is robust to incoming biomasses and susceptible to protolignin structural differences. Solid lignin recovery after autohydrolysis was doubly lower in nonwoods compared to hardwoods, showcasing differences between hardwood and nonwood protolignin structure. Across all biomasses, total yield of non-condensed lignin structures in autohydrolyzed pulps did not clearly affect enzymatic digestibility. Delineation between inter-lignin condensation and preferential hydrolysis mandates analysis of solubilized lignin. The S/V of hardwood non-condensed lignin structures positively affected total sugar recovery, revealing that hardwood protolignin's chemical structure can greatly affect a biologically-based biorefinery process' efficiency.

REFERENCES

1. Budzinski, M.; Nitzsche, R. Comparative economic and environmental assesment of four beech wood based biorefinery concepts. *Bioresource Technology* **2016,** *216*, 613-621.

2. Mitchell, R. B.; Schmer, M. R.; Anderson, W. F.; Jin, V.; Balkcom, K. S.; Kiniry, J.; Coffin, A.; White, P. Dedicated energy crops and crop residues for bioenergy feedstocks in the central and eastern USA. *BioEnergy Research* **2016,** *9* (2), 384-398.

3. Venkata Mohan, S.; Nikhil, G. N.; Chiranjeevi, P.; Nagendranatha Reddy, C.; Rohit, M. V.; Naresh Kumar, A.; Sarkar, O. Waste biorefinery models towards sustainable circular bioeconomy: Critical review and future perspectives. *Bioresource Technology* **2016,** *2015*, 2-12.

4. Kim, H. J. Beyond Asian Biotechnology: Collaboration between Asia and Europe. *Biotechnology Journal* **2015,** *10* (12), 1838-1843.

5. Bozell, J. J.; Astner, A.; Baker, D.; Biannic, B.; Cedeno, D.; Elder, T.; Hosseinaei, O.; Delbeck, L.; Kim, J.-W.; O'Lenick, C. J.; Young, T. Integrating separation and conversion—conversion of biorefinery process streams to biobased chemicals and fuels. *BioEnergy Research* **2014,** *7* (3), 856-866.

6. Piltan, M.; Sowlati, T. A review of partnership studies in the forest products value chain: with a focus on developed countries (United States, Canada, and Western Europe). *Forest Products Journal* **2014,** *64* (1-2), 4-10.

7. Phillips, R. B.; Jameel, H.; Chang, H.-m. Integration of pulp and paper technology with bioethanol production. *Biotechnology for Biofuels* **2013,** *6* (13).

8. Wang, Y.; Chen, C.; Cai, D.; Wang, Z.; Qin, P.; Tan, T. The optimizatin of L-lactic acid production from sweet sorghum juice by mixed fermentation of *Bacillus coagulans* and *Lactobacillus rhamnosus* under unsterile conditions. *Bioresource Technology* **2016,** *218*, 1098-1105.

9. Geng, W.; Jin, Y.; Jameel, H.; Park, S. Strategies to achieve high-solids enzymatic hydrolysis of dilute-acid pretreated corn stover. *Bioresource Technology* **2015,** *187*, 43-48.

10. Leskinen, T.; Kelley, S. S.; Argyropoulos, D. S. Refining of ethanol biorefinery residues to isolate value added lignins. *ACS Sustainable Chemistry & Engineerng* **2015,** *3* (7), 1632-1641.

11. Mariano, A. P. Due diligence for sugar platform biorefinery projects: technology risk. *Journal of Science & Technology for Forest Products and Processess* **2014,** *4* (5), 12.

12. Treasure, T.; Gonzalez, R.; Jameel, H.; Phillips, R. B.; Park, S.; Kelley, S. Integrated conversion, financial, and risk modeling of cellulosic ethanol from woody and non-woody biomass via dilute acid pre-treatment. *Biofuels, Bioproducts, & Biorefining* **2014,** *8,* 755-769.

13. Beckham, G. T.; Johnson, C. W.; Karp, E. M.; Salvachua, D.; Vardon, D. R. Opportunities and challenges in biological lignin valorization. *Current Opinion in Biotechnology* **2016,** *42*, 40-53.

14. Ma, X.; Cui, K.; Hao, W.; Ma, R.; Tian, Y.; Li, Y. Alumnia supported molybdenum catalyst for lignin valorization: Effect of reduction temperature. *Bioresource Technology* **2015,** *192,* 17-22.

15. Mahmood, N.; Yuan, Z.; Schmidt, J.; Xu, C. Hydrolytic deploymerization of hydrolysis lignin: Effects of catalysts and solvents. *Bioresource Technology* **2015,** *190,* 416-419.

16. Ragauskas, A. J.; Beckham, G. T.; Biddy, M. J.; Chandra, R.; Chen, F.; Davis, M. F.; Davison, B. H.; Dixon, R. A.; Gilna, P.; Keller, M.; Langan, P.; Naskar, A. K.; Saddler, J. N.; Tschaplinski, T. J.; Tuskan, G. A.; Wyman, C. E. Lignin valorization: improving lignin processing in the biorefinery. *Science* **2014,** *344* (6185), 709-720.

17. Batalha, L. A. R.; Han, Q.; Jameel, H.; Chang, H.-m.; Colodette, J. L.; Gomes, F. J. B. Production of fermentable sugars from sugarcane bagasse by enzymatic hydrolysis after autohydrolysis and mechanical refining. *Bioresource Technology* **2015,** *180*, 97-105.

18. Han, Q.; Jin, Y.; Jameel, H.; Chang, H.-m.; Phillips, R.; Park, S. Autohydrolysis pretreatment of waste wheat straw for cellulosic ethanol production in a co-located straw pulp mill. *Applied Biochemistry and Biotechnology* **2015,** *175* (2), 1193-1210.

19. Vargas, F.; Dominguez, E.; Vila, C.; Rodriguez, A.; Garrote, G. Agricultural residue valorization using a hydrothermal process for second generation bioethanol and oligosaccharides production. *Bioresource Technology* **2015,** *191* (263-270).

20. Jiang, W.; Chang, S.; Qu, Y.; Zhang, Z.; Xu, J. Changes on structural properties of biomass pretreated by combined deacetylation with liquid hot water and its effect on enzymatic hydrolysis. *Bioresource Technology* **2016,** *220*, 448-456.

21. Rissanen, J. V.; Murzin, D. Y.; Salmi, T.; Grenman, H. Aqueous extraction of hemicelluloses from spruce - From hot to warm. *Bioresource Technology* **2016,** *199*, 279-282.

22. Santucci, B. S.; Maziero, P.; Rabelo, S. C.; Curvelo, A. A. S.; Pimenta, M. T. B. Autohydrolysis of hemicelluloses from sugarcane bagasse during hydrothermal pretreatment: a kinetic assessment. *BioEnergy Research* **2015,** *8* (4), 1778-1787.

23. Li, H.-Q.; Jiang, W.; Jia, J.-X.; Xu, J. pH pre-corrected liquid hot water pretreatment on corn stover with high hemicellulose recovery and low inhibitors formation. *Bioresource Technology* **2014,** *153*, 292-299.

24. Koo, B.-w.; Treasure, T. H.; Jameel, H.; Philips, R. B.; Chang, H.-m.; Park, S. Reduction of enzyme dosage by oxygen delignification and mechanical refining for enzymatic hydrolysis of green liquor-pretreated hardwood. *Applied Biochemistry and Biotechnology* **2011,** *165* (3), 832-844.

25. Wu, S.; Chang, H.-m.; Jameel, H.; Philips, R. B. Novel green liquor pretreatment of loblolly pine chips to facilitate enzymatic hydrolysis into fermentable sugars for ethanol production. *Journal of Wood Chemistry and Technology* **2010,** *30* (3), 205-218.

26. Park, J.; Jones, B.; Koo, B.; Chen, X.; Tucker, M.; Yu, J.-H.; Pschorn, T.; Venditti, R.; Park, S. Use of mechanical refining to improve the production of low-cost sugars from lignocellulosic biomass. *Bioresource Technology* **2016,** *199*, 59-67.

27. Jones, B. W.; Venditti, R.; Park, S.; Jameel, H.; Koo, B. Enhancement in enzymatic hydrolysis by mechanical refining for pretreated hardwood lignocellulosics. *Bioresource Technology* **2013,** *147*, 353-360.

28. Narron, R. H.; Kim, H.; Chang, H.-m.; Jameel, H.; Park, S. Biomass pretreatments capable of enabling lignin valorization in a biorefinery process. *Current Opinion in Biotechnology* **2016,** *38*, 39-46.

29. Sluiter, J. B.; Ruiz, R. O.; Scarlata, C. J.; Sluiter, A. D.; Templeton, D. W. Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. *Journal of Agricultural and Food Chemistry* **2010,** *58* (16), 9043-9045.

30. Sluiter, A.; Hames, B.; Ruiz, R.; Scarlata, C.; Sluiter, J.; Templeton, D. Determination of sugars, byproducts, and degradation products in liquid fraction process samples. In *National Renewable Energy Laboratory Technical Report NREL/TP-510-42623*, 2008.

31. Ghose, T. K. Measurement of cellulase activities. *Pure & Applied Chemistry* **1987,** *59*, 257-268.

32. Shen, H.; He, X.; Poovaiah, C. R.; Wuddineh, W. A.; Ma, J.; Mann, D. G. J.; Wang, H.; Jackson, L.; Tang, Y.; Stewart Jr., N. S.; Chen, F.; Dixon, R. A. Functional

characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. *New Phytologist* **2012,** *193* (1), 121-136.

33. Santos, R. B.; Capanema, E. A.; Balakshin, M. Y.; Chang, H.-m.; Jameel, H. Lignin structural variation in hardwood species. *Journal of Agricultural and Food Chemistry* **2012,** *60* (19), 4923-4930.

34. Min, D.; Xiang, Z.; Liu, J.; Jameel, H.; Chiang, V.; Jin, Y.; Chang, H.-m. Improved protocol for alkaline nitrobenzene oxidation of woody and non-woody biomass. *Journal of Wood Chemistry and Technology* **2014,** *35* (1), 52-61.

35. Lupoi, J. S.; Seema, S.; Simmons, B. A.; Henry, R. J. Assesment of lignocellulosic biomass using analytical spectroscopy: an evolution to high-throughput techniques. *BioEnergy Research* **2014,** *7* (1), 1-23.

36. Vassilev, S. V.; Baxter, D.; Andersen, L. K.; Vassileva, C. G. An overview of the composition and application of biomass ash. Part 1. Phase-mineral and chemical composition and classification. *Fuel* **2013,** *105*, 40-76.

37. del Rio, J. C.; Rencorett, J.; Prinsen, P.; Martinez, A. T.; Ralph, J.; Gutierrez, A. Structural characterization of wheat straw lignin as revealed by analytical pyrolysis, 2D-NMR, and reductive cleavage methods. *Journal of Agricultural and Food Chemistry* **2012,** *60* (23), 5922-5935.

38. del Rio, J. C.; Lino, A. G.; Colodette, J. L.; Lima, C. F.; Gutierrez, A.; Martinez, A. T.; Lu, F.; Ralph, J.; Rencoret, J. Differences in the chemical structure of the lignins from sugarcane bagasse and straw. *Biomass and Bioenergy* **2015,** *81*, 322-338.

39. Pihlajaniemi, V.; Sipponen, M. H.; Pastinen, O.; Lehtomaki, I.; Laakso, S. Yield optimization and rational function modelling of enzymatic hydrolysis of wheat straw pretreated by NaOH-delignification, autohydrolysis and their combination. *Green Chemistry* **2015,** *17* (3), 1683-1691.

40. Vallejos, M. E.; Zambon, M. D.; Area, M. C.; da Silva Curvelo, A. A. Low liquid-solid ratio fractionation of sugarcane bagasse by hot water autohydrolysis and organosolv delignification. *Industrial Crops and Products* **2015,** *65*, 349-353.

41. Romani, A.; Garrote, G.; Lopez, F.; Parajo, J. C. Eucalyptus globulus wood fractionation by autohydrolysis and organosolv delignification. *Bioresource Technology* **2011,** *102* (10), 5896-5904.

42. Ruiz, H. A.; Ruzene, D. S.; Silva, D. P.; da Silva, F. F. M.; Vicente, A. A.; Teixeira, J. A. Development and characterization of an environmentally friendly process sequence (autohydrolysis and organosolv) for wheat straw delignification. *Applied Biochemistry and Biotechnology* **2011,** *164* (5), 629-641.

43. Huang, C.; Jeuck, B.; Du, J.; Yong, Q.; Chang, H.-m.; Jameel, H.; Phillips, R. B. Novel process for the coproduction of xylo-olgiosaccharides, fermentable sugars, and lignosulfonates from hardwood. *Bioresource Technology* **2016,** *219*, 600-607.

44. Wang, T., Nolte, M. W., Shanks, B. H. Catalytic dehydration of C6 carbohydrates for the production of hydroxymethylfurfural (HMF) as a versatile platform chemical. *Green Chemistry* **2014,** *16*, 548-572.

45. Huang, C.; Wu, X.; Huang, Y.; Lai, C.; Li, X.; Yong, Q. Prewashing enhances the liquid hot water pretreatment efficiency of waste wheat straw with high free ash content. *Bioresource Technology* **2016,** *219*, 583-588.

46. Ko, J. K.; Ximenes, E.; Kim, Y.; Ladisch, M. R. Effect of liquid hot water pretreatment severity on properties of hardwood lignin and enzymatic hydrolysis of cellulose. *Biotechnology and Bioengineering* **2015,** *112* (2), 252-262.

47. Santos, R. B.; Treasure, T.; Gonzalez, R.; Phillips, R.; Lee, J. M.; Jameel, H.; Chang, H.-m. Impact of hardwood species on production cost of second generation ethanol. *Bioresource Technology* **2012,** *117*, 193-200.

48. Ko, J. K.; Ximenes, E.; Kim, Y.; Ladisch, M. R. Adsorption of enzyme onto lignins of liquid hot water pretreated hardwoods. *Biotechnology and Bioengineering* **2015,** *112* (3), 447-456.

49. Yu, Z.; Gwak, K.-S.; Treasure, T.; Jameel, H.; Chang, H.-m.; Park, S. Effect of lignin chemistry on the enzymatic hydrolysis of woody biomass. *ChemSusChem* **2014,** *7* (7), 1942-1950.

50. Duarte, G. C.; Moreira, L. R. S.; Jaramillo, P. M. D.; Filho, E. X. F. Biomass-derived inhibitors of holocellulases. *BioEnergy Research* **2012,** *5* (3), 768-777.

51. Tejirian, A.; Xu, F. Inhibition of enzymatic cellulolysis by phenolic compounds. *Enzyme and Microbial Technology* **2011,** *48* (3), 239-247.

52. Yokoyama, T. Revisiting the mechanism of β-O-4 bond cleavage during acidolysis of lignin. Part 6: a review. *Journal of Wood Chemistry and Technology* **2014,** *35* (1), 165-174.

TABLES

Table 1. Chemical composition of raw material (g / 100 g oven-dry raw material)

Biomass	Glucan	Xylan	Galactan	A+M [a]	AIL [b]	ASL [c]	Ash	Ext [d]	Unidentified	*para*-Coumarate [e]	Ferulate [e]
Sugarcane bagasse	42.3 ± 0.4	20.9 ± 0.3	0.9 ± 0.1	2.6 ± 0.4	18.2 ± 0.1	2.3 ± 0.0	2.7 ± 0.1	3.9 ± 0.3	6.2	5.2% ± 0.3%	0.5% ± 0.2%
Wheat straw	39.0 ± 0.5	18.8 ± 0.4	1.2 ± 0.0	3.1 ± 0.0	19.4 ± 0.1	2.9 ± 0.2	5.6 ± 0.4	1.9 ± 0.1	8.1	0.7% ± 0.1%	1.4% ± 0.3%
Switchgrass	34.3 ± 0.4	24.1 ± 0.4	0.9 ± 0.1	3.1 ± 0.1	18.4 ± 0.3	2.9 ± 0.4	3.0 ± 0.1	2.1 ± 0.1	11.2	5.2% ± 0.5%	2.0% ± 0.3%
Maple	43.2 ± 0.5	13.0 ± 0.0	0.7 ± 0.1	3.4 ± 0.2	22.1 ± 0.2	3.1 ± 0.1	0.3 ± 0.0	2.0 ± 0.1	12.2	-	-
Sweetgum	40.2 ± 1.0	15.7 ± 0.5	0.8 ± 0.1	2.2 ± 0.0	21.5 ± 0.5	3.7 ± 0.0	0.6 ± 0.1	1.2 ± 0.3	14.1	-	-
Nitens	44.4 ± 0.8	14.3 ± 0.1	1.1 ± 0.0	1.4 ± 0.0	19.9 ± 0.5	4.6 ± 0.0	0.2 ± 0.1	1.1 ± 0.1	13.0	*nd*	*nd*

[a] Arabinan + mannan (co-eluting monosaccharides)
[b] Acid-insoluble lignin (Klason lignin)
[c] Acid-soluble lignin
[d] Extractives
[e] %w/w of total raw lignin
nd = not detected

Table 2. Chemical composition of pretreated biomass and total carbohydrate content in autohydrolyzate (g / 100 g oven-dry raw material)

Biomass	Solid yield[a]	Glucan		Xylan		Galactan[b]		A&M[c]		AIL[d]	ASL[e]	Ash	Unidentified
		Solid	Liquid	Solid	Liquid	Solid	Liquid	Solid	Liquid				
Sugarcane bagasse	59.0 ± 0.3	37.3 ± 0.1	1.9 ± 0.0	3.5 ± 0.0	10.4 ± 0.2	*nd*	0.7 ± 0.1	0.4 ± 0.1	1.2 ± 0.3	13.2 ± 0.0	0.8 ± 0.0	1.4 ± 0.1	4.1
Wheat straw	54.9 ± 0.5	30.9 ± 0.3	3.3 ± 0.1	4.3 ± 0.1	9.3 ± 0.1	*nd*	1.0 ± 0.1	0.3 ± 0.1	1.6 ± 0.3	14.9 ± 0.2	0.7 ± 0.0	0.7 ± 0.4	5.6
Switchgrass	53.1 ± 0.5	30.0 ± 0.1	2.1 ± 0.1	5.1 ± 0.0	14.0 ± 0.2	*nd*	0.8 ± 0.2	0.3 ± 0.1	2.0 ± 0.2	13.1 ± 0.1	0.6 ± 0.0	0.4 ± 0.1	6.8
Maple	69.2 ± 0.4	40.1 ± 0.3	1.0 ± 0.0	3.0 ± 0.1	9.2 ± 0.1	*nd*	0.6 ± 0.0	1.5 ± 0.4	1.5 ± 0.1	20.4 ± 0.0	0.9 ± 0.0	*nd*	4.8
Sweetgum	66.5 ± 0.2	37.5 ± 0.0	2.3 ± 0.2	3.7 ± 0.0	10.4 ± 0.1	*nd*	0.5 ± 0.1	1.6 ± 0.0	0.5 ± 0.0	20.3 ± 0.0	1.1 ± 0.2	0.1 ± 0.0	3.3
Nitens	68.8 ± 0.2	43.3 ± 0.2	0.6 ± 0.1	2.5 ± 0.1	10.7 ± 0.1	*nd*	1.1 ± 0.1	0.6 ± 0.1	0.7 ± 0.1	18.7 ± 0.2	1.1 ± 0.1	*nd*	3.8

[a] Weight of extensively-washed solids after autohydrolysis
[b] Galactan not detected in autohydrolyzed solids
[c] Arabinan + mannan (co-eluting monosaccharides)
[d] Acid-insoluble lignin (Klason lignin)
[e] Acid-soluble lignin
nd = not detected

Table 3. Non-condensed lignin yields from nitrobenzene oxidation of raw and autohydrolyzed biomass

Biomass	mmol / g lignin				S/V	H/V
	V+Va [a]	S+Sa [b]	H [c]	Total		
Sugarcane bagasse	0.58 ± .03	0.87 ± .02	0.87 ± .02	2.32	1.5	1.5
Wheat straw	0.69 ± .02	0.67 ± .06	0.25 ± .01	1.61	1.0	0.4
Switchgrass	0.84 ± .05	0.61 ± .03	0.53 ± .01	1.98	0.7	0.6
Maple	0.78 ± .01	1.93 ± .06	-	2.71	2.5	-
Sweetgum	0.57 ± .02	2.05 ± .05	-	2.62	3.6	-
Nitens	0.46 ± .01	2.36 ± .08	-	2.82	5.1	-
Autohydrolyzed biomass	mmol / g lignin				S/V	H/V
	V+Va [a]	S+Sa [b]	H [c]	Total		
Sugarcane bagasse	0.40 ± .02	0.82 ± .04	0.65 ± .01	1.87	2.1	1.6
Wheat straw	0.45 ± .04	0.55 ± .05	0.17 ± .02	1.17	1.2	0.4
Switchgrass	0.60 ± .01	0.54 ± .02	0.32 ± .02	1.46	0.9	0.5
Maple	0.46 ± .01	1.23 ± .05	-	1.69	2.7	-
Sweetgum	0.45 ± .04	1.34 ± .04	-	1.79	3.0	-
Nitens	0.25 ± .06	1.34 ± .05	-	1.59	5.4	-

[a] Vanillin + vanillic acid (170 °C, 2.5 hours)
[b] Syringaldehyde + syringic acid (170 °C, 2.5 hours)
[c] 4-hydroxybenzaldehyde (190 °C, 4 hours)

FIGURES

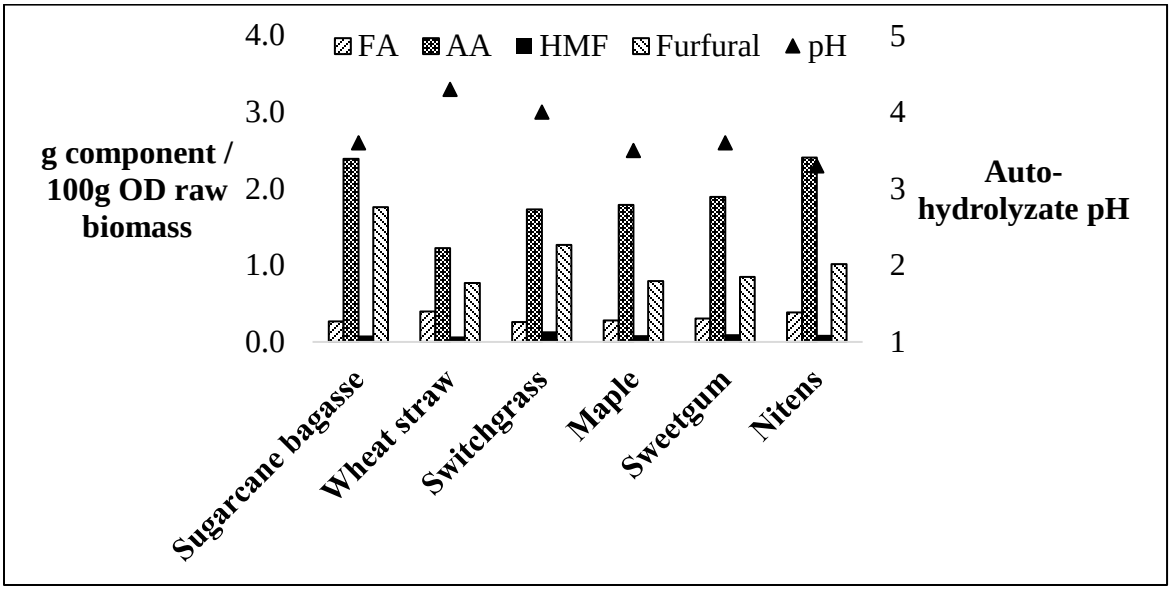

Figure 1. Organic byproducts obtained in autohydrolyzate and autohydrolyzate pH.

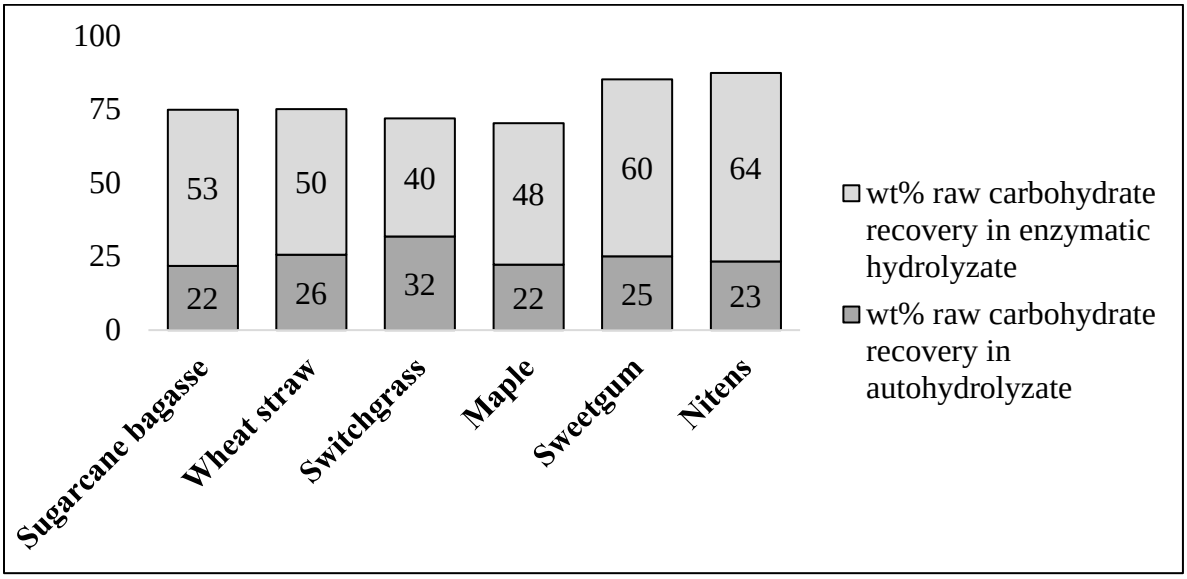

Figure 2. Total sugar recovery by autohydrolysis, mechanical refining, and enzymatic hydrolysis at mild enzyme dosages (5 FPU/ oven-dry g biomass).

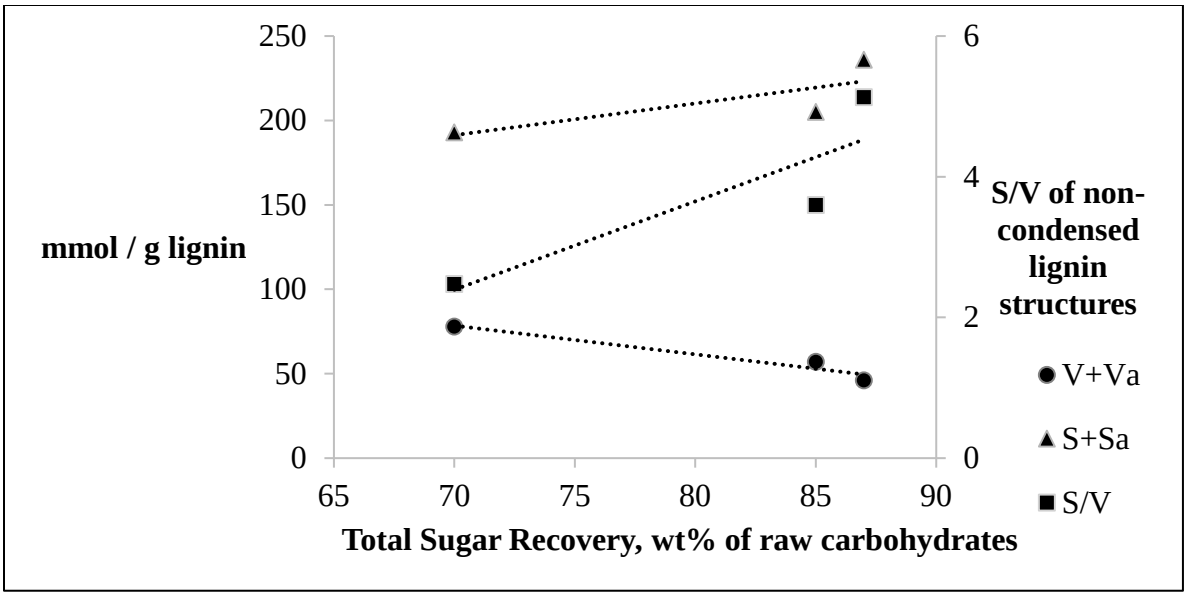

Figure 3. Total sugar recovery from hardwoods as a function of non-condensed lignin structures.

CHAPTER 3. Soluble lignin recovered from biorefinery pretreatment hydrolyzate characterized by lignin-carbohydrate complexes

Robert H. Narron, Hou-min Chang, Hasan Jameel, Sunkyu Park

ACS Sustainable Chemistry & Engineering, (in press)

Abstract

The lignin rendered soluble by lignocellulosic biorefinery pretreatment remains insufficiently understood along the lines of molecular properties and chemical composition. To procure a representative soluble lignin preparation, aromatic-selective adsorptive resin was utilized. ~90% of soluble lignin could be recovered from autohydrolysis pretreatment hydrolyzate (autohydrolyzate) produced from a hardwood and a non-wood biomass. Adsorbate compositional characterization revealed a befuddling magnitude of carbohydrate in selectively isolated lignin adsorbates. Quantitative structural analysis of the lignin by NMR suggested lignin-carbohydrate complexes (LCCs) as the cause behind the pronounced carbohydrate contents. Analyzed spectra revealed both hardwood and non-wood soluble lignin features ~10 total LCC per 100 aromatic rings, with each lignin bearing unique LCC profiles. In addition, native structures remained in large quantities. The improved understanding of hydrolyzate-soluble lignin granted from this work will aid biorefinery development by improving discourse around a biorefinery lignin source.

1. Introduction

Efforts to sustainably produce fuels, chemicals, and materials have led to formulation of the lignocellulosic biorefinery. Biologically-based biorefineries, which by definition employ biotransformative unit operations (enzymatic, fermentative) are currently poised to produce high yields of biosugars for further upgrading into higher-value products. One of many key requirements for biorefineries at the industrial scale is successful employment of an economically-feasible pretreatment process, required to fracture native recalcitrance towards biotransformation. The pretreatment arsenal which one can select from to dismantle biomass' intrinsic recalcitrance is endless.[1-4] Unfortunately, when considering both pretreatment cost and effectiveness, the two tend to be intimately inseparable.[5-6] Unique amongst pretreatment technologies is autohydrolysis, which boasts several features that afford excellent techno-economic advantages when integrated into a biorefinery process.[7-9] In addition to affordably decomposing biomass in a manner that is favorable for downstream cellulolytic enzymes, autohydrolysis also yields a liquid phase ("autohydrolyzate") that is rich in soluble hemicellulosic and lignin derivatives. Autohydrolyzate commands understanding due to the presence of a significant quantity of sugars (~25 wt% of raw carbohydrates)[10] and a variety of biotransformative inhibitors.[11]

One of the aforementioned inhibitors in autohydrolyzate is solublized lignin. Total quantities of lignin solubilized by pretreatment are conventionally back-calculated using both biomass weight loss and diminuation in lignin content (g/100 g oven-dry "o.d." lignocellulosic solid). This quantification is useful for quantitatively describing the extent of delignification, yet fails to describe chemical and molecular properties of the soluble lignin. Because of this,

understanding of the molecular and chemical properties of lignin rendered soluble by pretreatment is generally incomplete. Pretreatment hydrolyzate-soluble lignin has been characterized based upon chemical behavior (free-radical scavenging activity, redox reagent response) rather than chemical structure.[12-14] One report successfully performed exact quantification of an enormous variety of soluble monomeric lignin fragments found in dilute acid pretreatment hydrolyzate from a breadth of differing biomasses.[15] Importantly, previous work found this method able to characterize only minor portions (~1-2% w/w) of the soluble lignin in autohydrolyzate.[16] This revealed autohydrolyzate-soluble lignin to be predominantly non-monomeric.

The non-monomeric nature of autohydrolyzate-soluble lignin suggested adsorptive materials capable of engaging aromatic chemical structures are an appropriate means of lignin sequestration from autohydrolyzate. Investigations and reviews of adsorption treatments of biomass hydrolyzates are available, providing useful information concerning hydrolyzate response to various adsorptive materials. Known for adsorptive potency, activated carbon was demonstrated as an effective adsorbent for soluble lignin, but lacked selectivity due to carbohydrate adsorbance.[17-18] In pursuit of lignin-selective adsorptive material, Amberlite® XAD resins presented an opportunity to semi-selectively adsorb soluble lignin plus furans produced from carbohydrate dehydration (furfural & 5-hydroxymethylfurfural ("HMF")).[18-23] The semi-selectivity is a function of the polymeric adsorbent's chemical composition- a novel formulation of styrene-divinylbenzene copolymer, which engages aqueous-soluble hydrophobic molecules through π-π' and van der Waal interactions. One publication suggested XAD-adsorbed lignin from biomass hydrolyzate featured a strong presence of covalently

linked carbohydrate segments (lignin-carbohydrate complex ("LCC") structures).[22] Despite logical speculation, explicit proof of LCC connectivities was not provided due an absence of spectroscopic analysis. Spectroscopic elucidation of LCCs in XAD adsorbates was succesful in one instance, however, the adsorbate originated from a soluble fraction of ball-milled raw biomass rather than biorefinery pretreatment hydrolyzate.[24]

To spectroscopically measure the extent of lignin-carbohydrate covalent connectivity within the soluble lignin in a biorefinery hydrolyzate, first, XAD resin adsorption was performed to isolate said lignin. After adsorbate procurement, chemical and molecular characterization was performed. Lignin adsorbates from unlike lignocellulosic biomasses (hardwood & non-wood) are compared throughout. The long-term goal of performing characterization upon the investigated lignin fraction is to improve academic and industrial discourse around hydrolyzate-soluble lignin's permeating effects upon biorefinery processing, either as biotransformation inhibitors or valorizable co-products.

2. Materials and methods

2.1. Raw materials and autohydrolysis pretreatment

Hardwood chips from maple (*Acer rubrum*) were acquired from the Tree Improvement Program of North Carolina State University (Raleigh, NC). Sugarcane bagasse was provided by a cane sugar manufacturer located in the State of Minas Gerais, Brazil. Prior to autohydrolysis, non-structural extractives were removed from each biomass to nullify non-lignin result contamination. For extraction, raw biomasses were fluxed with benzene-ethanol solution (2:1 v:v) overnight in soxhlet extractors. Also, the quantity of extractives removed

was measured gravimetrically. Autohydrolysis pretreatment was performed within a 1 L alloy reactor (Parr Model C-276, Parr Instrument Company, USA). ~50 air-dry g of biomass was massed and loaded into the reactor, followed by addition of the appropriate amount of deionized water to establish starting water-to-solid ratio (g:g) of 10:1 (always accounting for biomass moisture content). Pretreatment temperature was set to 180°C with ramp-up of ~30 min. Target temperature was maintained for 40 minutes, followed by rapid thermal quenching via ice water bath. After cooling, the reactor was emptied, and solid was separated from liquid using vacuum filtration. After isolation of bulk autohydrolyzate, the autohydrolyzed solids were washed with deionized water until resultant wash filtrate reached neutrality. Bulk autohydrolyzate was passed through Gooch crucibles (fine grade) prior to experimentation to ensure the absence of suspended solids.

2.2. Resin adsorption protocol

Amberlite® XAD16N was used for lignin adsorption from autohydrolyzate. To prepare resin, an extensive washing procedure was conducted: 1) water-washing to remove preservative salts, 2) methanol washing to remove adsorbed impurity, 3) water washing to remove methanol. Prepared resins were sealed in plastic bags and kept under refrigeration prior to use. Before resin adsorption experiments, moisture content of prepared resin was checked using an infrared moisture balance (Sartorius MA160-1, Cole Parmer, USA) to account for its dilution effect upon autohydrolyzate during adsorption. For adsorption optimization, resin weight (o.d. basis) was kept constant while varying the volume of

autohydrolyzate. After 30 minutes of low-speed mixing, the autohydrolyzate and resin suspension were separated by vacuum filtration.

Following separation of resin and autohydrolyzate, adsorbate-bearing resin was washed with deionized water (~10x autohydrolyzate volume) to remove un-adsorbed autohydrolyzate solutes. Next, methanol was used to desorb adhered molecules, and the methanol solution was collected. Desorption was considered complete once methanol filtrate reached total visual transparency. The amount of methanol required per adsorption was ~3x the volume of starting autohydrolyzate. Next, methanol was evaporated by rotary evaporator (Rotavapor R-303, Buchi, USA), and the obtained solids were re-suspended into deionized water and freeze-dried (Freezone 6, Labconco, USA) to produce handleable solid adsorbates. Solids were stored under vacuum at 35°C over solid P_2O_5 until constant mass prior to analysis.

2.3. UV spectrophotometry

The initial concentration of lignin in autohydrolyzate was measured as spectrophotometric absorbance at 280 nm, accounting for dilution (Lambda XLS, PerkinElmer, USA). Following mixing, the resin-autohydrolyzate suspension was separated and resultant filtrate was tested for UV absorption. The decrease in UV response represents the extent of adsorptive lignin removal from autohydrolyzate. To measure UV responses free of furanic contribution, an aliquot of autohydrolyzate was mixed with excess sodium borohydride overnight to reduce all solutes. After reduction, UV response was recorded. All absorbance values reported are average of duplicate analyses.

2.4. Compositional analysis of lignin adsorbates

The chemical composition of adsorbates was performed using the standard protocol provided by the National Renewable Energy Laboratory.[25] ~100 mg of carefully weighed o.d. adsorbate was subjected to 72% (w/w) sulfuric acid digestion for 1 hr, followed by dilution to 3% (w/w) H_2SO_4 and autoclaving for 1.5 hr. Modifications were made to acid digestion methodology to decrease lignin adsorbate's exposure to hot sulfuric acid. An aliquot of acid digestate was taken to determine acid-soluble lignin quantities spectrophotometrically and monosaccharides by liquid chromatography. For acid-soluble lignin, UV absorbance was measured at 205 nm, accounting for dilution. The molar absorptivity coefficient 110 L g^{-1} cm^{-1} was applied to convert absorbance to concentration. For monosaccharides, an HPLC system (Agilent 1200 series, Agilent, USA) was used. In this setup, a Shodex SP-0810 column is utilized (8x300 mm, Showa Denko, Japan) and maintained at 80 °C. Mobile phase was Milli-Q water, flowing 0.5 mL/min for 60 minutes per sample. A refractive index detector monitored eluents. Before injection, acid digestate samples were neutralized to pH ~6 with $CaCO_3$ and filtered through 0.2 μm nylon filters. Calibration was performed with commercial standards of glucose, xylose, galactose, mannose, and arabinose. Due to observed co-elution of mannose and arabinose, quantification of either saccharide is reported as the sum of both. All values represent the average of duplicate analyses, and average standard deviation was 1.5%.

2.5. Autohydrolyzate saccharide quantification around resin adsorption

Autohydrolyzate carbohydrate concentrations (monomeric and oligomeric) before and after resin adsorption was measured utilizing the same HPLC system used to measure

monosaccharides in acid digestates. To describe the presence of oligomeric carbohydrates[26], autohydrolyzates were subjected to mild sulfuric acid (3% w/w) hydrolysis in autoclave for 60 minutes to produce quantifiable oligomer-derived monosaccharides. Oligosaccharide concentrations were then back-calculated using the difference in monosaccharide concentrations before and after acid hydrolysis. All values represent the average of duplicate analyses, and average standard deviation was 0.3%.

2.6. NMR experiments

All 2D-HSQC and ^{13}C NMR spectra were acquired at 25°C using dimethyl sulfoxide-d6 (99.9%, VWR, USA) solvent. Adsorbate concentration was ~20% w/v. 2D-HSQC was performed first, followed by ^{13}C upon the same sample. 2D-HSQC NMR spectra were acquired on a 500 MHz NMR spectrometer equipped with 5 mm BBI probe (Avance, Bruker, USA). Coupling constant ($^{1}J_{C-H}$) of 147 Hz was applied. ^{13}C NMR spectra were acquired on the same 500 MHz spectrometer with a different probe, 5 mm BBO. An inverse gated proton decoupling sequence was employed for ^{13}C observation. Chromium (III) acetylacetonate was added to the NMR tube (0.01 M) following 2D-HSQC analysis to assist relaxation of carbon nuclei during ^{13}C observation. Each spectra was acquired once, therefore ~5% error is assumed.

3. Results and discussion

3.1. Autohydrolysis of raw biomass to produce autohydrolyzate-soluble lignin

Autohydrolysis pretreatment is known to effectively hydrolyze hemicelluloses and a portion of lignin. For lignin, there exists two known modes of degradation in acidic mediums: acidolysis and condensation.[3, 27-28] Regarding the former, lignin acidolysis is accompanied by loss of native chemical structures and decreases in molecular weight. Concerning condensation, electrophilic aromatic substitution around reactive benzylic carbocations occurs, resulting in increased lignin molecular weight through formation of new aryl carbon-carbon bonds. Currently, exact delineation between hydrolysis and condensation has been hindered by the absence of characterization of soluble lignin's molecular properties. This reason can be viewed as an additional motivator for better understanding autohydrolyzate-soluble lignin, with the conundrum having been discussed in detail in our previous publication.[10]

As can be seen from the chemical composition changes induced by autohydrolysis (Table 1), hemicellulosic xylan is the carbohydrate most affected by autohydrolysis, with the soluble portion existing in autohydrolyzate solution as xylooligosaccharides, xylose, and furfural (depending upon pretreatment severity).[10] Concerning lignin, minor decreases are observable for both acid-soluble lignin and acid-insoluble lignin contents, suggesting that enough hydrolysis has occurred to grant solubility to different portions of the original lignin (maple: 15% delignification, sugarcane bagasse: 32% delignification). The differences between hardwood and non-wood protolignin monomeric constitution and assembly are likely driving forces behind the differing amounts of lignin removal, especially considering

that the biomass with more raw lignin (maple) features a lesser degree of delignification. In all, the absolute quantity of soluble lignin produced and the percentage of raw lignin it embodies demands its characterization.

3.2. UV spectrophotometry to quantify lignin adsorbance

UV spectrophotometry was used to measure soluble lignin's UV response and the extent of adsorption. Table 2 displays back-calculated quantities of soluble lignin in autohydrolyzate (from Table 1), solute UV response, and residual solute UV response post adsorption. Furfural (dehydrated pentose) and HMF (dehydrated hexose), both of which generate through carbohydrate dehydration during autohydrolysis, problematically exhibit UV absorption at the absorption wavelength used to monitor soluble lignin (λ=280 nm, approximate λ_{max}).[29] To combat lignin overestimation, autohydrolyzate solutes before and after resin treatment were subjected to borohydride reduction to render UV-silent furfuryl alcohols, amongst other reduced compounds. This approach was demonstrated viable in a previous publication concerning spectrophotometric differentiation between lignin and furans.[30]

Overall, removal of lignins and furans was found to be 92% for maple and 90% for sugarcane bagasse, Table 2. Silencing furan response through reduction revealed 93% and 92% removal of lignin from maple and sugarcane bagasse autohydrolyzates, respectively. Comparing lignin adsorption from hardwood versus non-wood autohydrolyzate, a similar percentage of lignin is adsorbed from both autohydrolyzates despite variances in starting soluble lignin quantities and raw biomass. At the maximum resin loading tested, a slightly

greater quantity of un-adsorbed lignin is present in the sugarcane bagasse autohydrolyzate (~0.5 g) over maple autohydrolyzate (~0.3 g/100g o.d. raw biomass). This suggests further increasing resin loading may be required to achieve greater extents of lignin adsorption. Alternatively, un-adsorbed lignin could be unobtainable regardless of resin loading due to low extent of interaction between the remaining lignin and resin. Future work to resolve the aforementioned speculations is to be performed, involving extensive adsorption curves and manipulation of adsorption methodology in order to better understand the chemical properties of the presently unadsorbed solutes. Nevertheless, adsorption yields above 90% represents extensive soluble lignin isolation yields, especially considering the extent of soluble lignin characterization we previously reported ($\leq$ 2%).[15-16] The maximum resin loading tested (10:1) was maintained for all further lignin adsorptions.

3.3. Chemical composition of lignin adsorbates using standard and modified acid digestion

After laboratory processing of each adsorbate into handlable lignocellulosic solids, the first analytical assay performed was determination of adsorbate chemical composition to quantify lignin purity. Standard compositional analysis of lignocellulosic applies two-stage (concentrated to dilute) sulfuric acid digestion to degrade polymeric biomass fractions into soluble and insoluble fractions for quantification.[25] As seen in the first columns of Table 3, standard compositional analysis of lignin adsorbates rendered subpar mass balances. This value is considered as the extent of chemical composition characterization (maple: 77.3 g; sugarcane bagasse: 65.0 g/100g o.d. adsorbate). From these underwhelming results, it was

hypothesized that standard strong acid digestion caused excessive degradation of adsorbate constituents, resulting in lowered mass balances. To test, the acid digestion protocol was modified to only employ dilute acid digestion at conventional (1.5 hr) and reduced (1 hr) times. Continuing on Table 3, it is seen that modification of acid digestion protocol slightly enhanced mass balances (maple: 77.3-82.8 g/; sugarcane bagasse: 65.0-71.9 g/100g o.d. adsorbate). Mass balance enhancements were mostly a function of improvement in carbohydrate quantification, which was still relatively minor (e.g. maple: 46.5 g total carbohydrate at 1.5 hrs; 48.3 g total carbohydrate at 1 hr). Retention of carbohydrates with lessened acid digestion can be attributed to decreased loss of monosaccharides to dehydration. The magnitude of carbohydrates quantified and the ramifications of their pronounced presence in lignin adsorbates is extensively discussed in the latter portion of this manuscript.

Apart from carbohydrates, total lignin quantification was not greatly enhanced through acid digestion modification (33.9-35.7 & 23.6-26.1 g total lignin /100g o.d. adsorbate). Instead, the lignin fractions quantified shifted states from solid (acid-insoluble lignin) to liquid (acid-soluble lignin) with decreasing digestion severity. This is due to the increase in inter-lignin condensation catalyzed by greater exposure of lignin to hot sulfuric acid. Considering extent of characterization, the lowered mass balances afforded by conventional and modified acid digestion protocols suggest future work aimed at quantifying biomass components not specifically quantified during standard compositional analysis. We hypothesize that the key components eluding quantitation and hampering mass balance are likely hemicellulosic uronic acid substituents, hemicellulosic acetic acid substituents (if any), as well as hydroxycinnamic acids (*para*-coumaric acid, ferulic acid) found in non-woods only. The presence of

hydroxycinnamic acids in sugarcane bagasse adsorbates may also explain why its mass balance was ~11% lower than maple's. While furfural and HMF quantitation is also important, it is unlikely that the quantities of hexose-derived HMF and pentose-derived furfural will substantially aid analytical mass balance. This statement is supported by the minor relative percent increase in carbohydrate recovery observed when lowering dilute acid digestion times (maple: 4% improvement; sugarcane bagasse: 8% improvement).

3.4. Autohydrolyzate carbohydrate concentrations around resin adsorption

The chemical composition of the adsorbates was found to be surprisngly carbohydrate-enriched (maple: 48.3 g; sugarcane bagasse: 48.0 g total carbohydrate/100g o.d. adsorbate). This came in spite of the selective adsorptive medium. Given that original intent was to selectively separate soluble hydrophobic lignin fragments from hydrophilic autohydrolyzate solutes, oligosaccharide adsorption would indicate resin engagement of solutes not bound to lignin. To monitor the effect of resin application upon autohydrolyzate oligosaccharide concentrations, quantities of monosaccharides and oligosaccharides were measured around adsorption (Table 4).

In agreement with the pronounced presence of xylan in both adsorbates, the quantity of xylooligomers in both autohydrolyzates was found to decrease after resin application (sugarcane bagasse: 9.0 to 5.2 g (42% decrease); maple: 6.4 to 4.3 g/100g o.d. raw biomass (33% decrease)). Beyond xylan, a slight decrease in oligomeric glucan and arabinan + mannan concentrations was also observed. Monosaccharide concentrations in either autohydrolyzate was not affected by resin treatment. This was expected, as monomeric carbohydrates do not

bear the molecular properties suitable for adherence to XAD resin. Regardless, the findings verified that the starting autohydrolyzate solution is experiencing decreases in total carbohydrate concentration after resin treatment. This suggests either selective adsorbance of lignin that is enriched with chemically-attached carbohydrate domains, or, non-selective adsorption that engages oligosaccharides. It is important to note that the methodology applied to quantify oligosaccharide loss is indiscriminate of the aforementioned possiblities (monosaccharide concentrations before and after acid digestion). In attempt to distinguish between LCC adsorption and free oligosaccharide adsorption, spectroscopic analysis using NMR was performed in effort to identify and quantify LCC structures in lignin adsorbates.

3.5. Lignin-carbohydrate complex and native inter-lignin structural quantification by NMR analysis

NMR quantification of LCC structures was performed to provide explanation for carbohydrate-enrichment in lignin adsorbates. The technique of combining information-rich 2D-HSQC NMR spectra with informationally-coarse ^{13}C NMR spectra was utilized.[31-32] In this method, individual signals elucidated by 2D-HSQC and their integrals are interconverted into ^{13}C base units of number per 100 aromatic rings (#/100 Ar) to quantify structural abundance. The following equation was used for quantification of lignin moieties:

$$\frac{X}{100\ Ar} = \frac{2D_x}{2D_{IS}} * \frac{13C_{IS}}{13C_{163-103}} * 600$$

where X is the structure quantified, $2D_X$ is the integral of the 2D-HSQC signal utilized for quantification, IS is the integral of the internal standard cluster and range utilized in 2D-HSQC

and ^{13}C (respectively), $^{13}C_{163\text{-}103}$ is the integral of the aromatic region in ^{13}C spectras, and 600 is derived from 600 aromatic carbons being present per 100 aromatic rings. Each 2D-HSQC internal standard cluster and accompanying ^{13}C range utilized for quantification is explicitly stated in the discussion around each structure quantified. For native structures, integrals originating from C_α-H_α were utilized for the equation's "X", due to acknowledged absence of inappropriately overlapping signals. For LCC structures, the chemical shift integral employed is stated within each respective discussion (carbohydrate C_1 or side chain C_γ). All structurally-representative signals and their corresponding chemical shifts in 2D-HSQC and ^{13}C spectra were identified using reputable publications.[32-37] Figure 1 displays the chemical structures of each LCC and native structures quantified along with labeling acronyms, and Table 5 lists each structure's individual chemical shift.

Identification and quantification of LCC structures, in addition to several native lignin structures, was successful (Figure 2 and Table 6). Primarily, it can be observed that there is indeed a strong presence of LCC structures in both lignin adsorbates, visible from strong resonance of LCC signals in 2D-HSQC spectra. This provides reasonable explanation for the significant carbohydrate quantities of both adsorbates, and confirms the postulation by Westerberg et al.[22]

The quantified LCC structures of lowest abundance were benzyl ethers. There are two types of native benzyl ether LCC linkages: 1) benzyl ethers connecting Cα of lignin to carbohydrate primary hydroxyl groups, and 2) benzyl ethers connecting the Cα of lignin to carbohydrate secondary hydroxyl groups. The chemical shift integrated to quantify Cα-1° OH benzyl ether LCCs originates from the structure's C_α-H_α resonance, found at δ 81.3/4.75

ppm.[32, 38-39] For quantification, the internal standard cluster and range utilized was: δ 88-82/5.6-3.9 ppm (2D-HSQC); δ 88-82 ppm (^{13}C). The quantities of Cα-1° OH benzyl ethers present in maple and sugarcane bagasse adsorbates were measured to be 2.0 and 0.5 per 100 Ar, respectively. A representative signal belonging to Cα-2° OH benzyl ethers is known to overlap with a signal belonging to (non-LCC) spirodienone moieties at δ 81.2/5.1 ppm.[40] Fortunately, inaccuracies caused by this overlap are not problematic due to no observable signal at δ 81.2/5.1 ppm. From this, it can be concluded that neither native Cα-2° OH benzyl ethers nor spirodienone structures are present in adsorbates. The relatively minor quantities of Cα-1° OH benzyl ether LCC structures quantified indicates that the structure is either labile during autohydrolysis, not found in great abundance within either biomass‘ protolignin, or both.

Another form of lignin-carbohydrate complex is the uronosil ester. This ester linkage exists between hemicellulosic uronic acid substituents and lignin's side chain Cγ. The spectral region containing this LCC structure (δ 65-62/4.0-4.5 ppm) is also populated by signals pertaining to non-LCC esters, e.g. *para*-coumarates and ferulates. Due to the existence of these structures in non-wood lignin and their general absence from hardwood lignin (except reported *para*-benzoate moieties in *Populus*[41]), we first analyzed the region in the maple spectra to unbiasedly assign uronosil ester's chemical shift within the acquired spectra. A single signal was observed (δ 63.0/4.22 ppm), therefore this chemical shift was assigned to uronosil esters. This chemical shift is in agreement with the C_γ-H_γ chemical shifts noted in previous publications characterizing LCC structures using the same NMR technique and lignin solvent.[32, 38-39] Once the appropriate signal was identified in hardwood, the non-wood spectra

was overlaid and a matching signal was identified (Figure 2(b)). Flanking the uronosil ester signal in sugarcane bagasse's 2D-HSQC spectra are several signals attributable to non-LCC esters. These signals were not integrated, however, they likely somewhat contribute to non-wood uronosil ester's signal integral due to chemical shift proximity. In addition, a separate signal was also observed at δ 61.6/4.08 ppm, pertaining to Cγ acyl structures.[42] Observation of said signal is significant because it was previously reported that the two signals (uronosil ester and Cγ acyl) could not be resolved with 300 MHz NMR, but were visible with 950 MHz NMR.[32] As a reminder, 500 MHz NMR was utilized for analysis in this work. In total, we are able to accurately report uronosil ester structural quantities due to the signal's resonance being free from known interferants.

To perform quantiation of uronosil esters, the following internal standard cluster and range was utilized: δ 65-58/5.0-2.5 ppm (2D-HSQC); δ 65-58 ppm (^{13}C). The quantity of uronosil esters was found to be greater in maple adsorbate (7.6) compared to sugarcane bagasse (5.2/100 Ar). In agreement with the presence of uronosil esters in both adsorbates was observation of chemical shifts pertaining to non-esterified and esterified 4-O-methyl-D-glucuronic acid moieties, found at δ 97.2/5.18 ppm and δ 98.9/4.70 ppm, respectively (Figure 2(c))[32]. Signal intensity of the chemical shift pertaining to esterified glucuronic acid residues is visually more intense than the signal pertaining to non-esterified glucuronic acid residues, possibly indicating that the majority of these structures are indeed esterified to lignin. Finally, the low mass balances achieved during chemical compositional analysis of either adsorbates was hypothesized to be due in part to the presence of quantitation-eluding uronic acids. The strong presence of uronosil ester LCC structures quantified is in agreement with this belief.

Quantification of phenyl glycoside LCC structures relies upon integration over a small and sparsely populated chemical shift range (δ 104-99/4.8-5.2 ppm).[32, 38-39] A range is scoured for these signals (originating from C_1 of the attached carbohydrate) due to variance in phenyl glycoside resonance when lignin bonds to different carbohydrates. As seen in Figure 2(c), three signals were found in the phenyl glycoside chemical shift region, with differences in chemical shift when comparing adsorbates. For quantitation, the internal standard cluster and region utilized was δ 103-96/5.5-3.8 ppm (2D-HSQC), δ 103-96 ppm (^{13}C). The quantity of phenyl glycosides in either adsorbate was dissimilar, with 3.8 measured for sugarcane bagasse and 0.6/100 Ar for maple. This wide difference could indicate that sugarcane bagasse protolignin bears an enrichment of phenyl glycoside LCC structures. Alternatively, this may be evidence that native hardwood phenyl glycoside structures are autohydrolysis-labile, and those in non-wood are generally not.

In addition to quantifying the presence of lignin-carbohydrate complexes within both adsorbates, the following native lignin structures were also observed and quantified: β-O-4', β-β', β-5', and dibenzodioxicin (DBDO) (Figure 2(a) and Table 6).[35] Native structural quantification relied upon utilization of the following internal standard cluster and range: δ 88-82/5.6-3.9 ppm (2D-HSQC); δ 88-82 ppm (^{13}C). Signals pertaining to each structure's Cα-Hα resonance were integrated for quantification (Table 5). To quantify β-O-4', the Cα signal (δ 72.6/5.00 ppm) was integrated. Maple adsorbate was found to contain a higher quantity of β-O-4' lignin structures, 33.6, compared to 13.4/100 Ar in sugarcane bagasse adsorbate. The magnitude of β-O-4' recovered in autohydrolyzate-soluble lignin indicates that lignin solubility is not solely a function of decreased molecular weight through native aryl-ether

structural hydrolysis.[43-44] Instead, it appears that native structure fragmentation coupled with lignin-carbohydrate chemical connectivity within fragmented lignins provides a more-detailed explanation as to why this lignin fraction was rendered. By this logic, it is also reasonable to assume that prolonging autohydrolysis to induce major LCC cleavage could result in an amount of soluble lignin that is lesser than what is reported at the current condition, depending upon the extent of inter-lignin condensation versus native structure hydrolysis.[3, 27-28] Maple was also found to be more enriched in β-β' and β-5' structures compared to sugarcane bagasse. Cα signals used for quantifying β-β' and β-5' structures were δ 84.9/4.69 ppm and δ 86.8/5.49, respectively. Another significant observation was the large amount of dibenzodioxicin ("DBDO") structures quantified exclusively in sugarcane bagasse adsorbate (8.4/100 Ar). The signal integrated to quantify DBDO (C_α-H_α) was only observable in sugarcane bagasse 2D-HSQC spectra, at δ 83.6/4.86 ppm.[35] The pronounced presence of these structures indicates DBDO is resistant to degradation by autohydrolysis to an unknown extent. Overall, these findings highlight the significant quantity of native inter-lignin connections remaining in autohydrolyzate-soluble lignin as well as the connectivity differences between autohydrolyzate-soluble hardwood and non-wood lignins.

4. Conclusions

The fraction of lignin solubilized in autohydrolysis pretreatment hydrolyzate was obtainable through application of aromatic-selective resin to hardwood and non-wood autohydrolyzates. In the adsorbates, the strong presence of carbohydrates (~48 g/100g o.d. adsorbate) merited further investigation. Further analysis revealed the isolation of

carbohydrates to be a function of lignin-carbohydrate complexes, totaling ~10 total LCC per 100 Ar in both adsorbates. In addition to LCC structures, a variety of native inter-lignin linkages were also quantified (e.g. 33.6 & 13.4 β-O-4' /100 aromatic rings). This indicates molecular weight-lowering degredation of native lignin structures is not the only contributor to lignin solubilization, but so too is lignin-carbohydrate connectivity within autohydrolysis-degraded lignin. These findings elucidate new structural characteristics of the lignin fraction solubilized by autohydrolysis pretreatment.

References

1. Silveira, M. H. L.; Morais, A. R. C.; Lopes, A. M. d. C.; Olekszyszen, D. N.; Lukasik, R. B.; Andreaus, J.; Ramos, L. P. Current pretreatment technologies for the development of cellulosic ethanol and biorefineries. *ChemSusChem* **2015,** *8* (20), 3366-3390.

2. Yuan, X.; Singh, S.; Simmons, B. A.; Cheng, G. Biomass pretreatment using dilute aqueous ionic liquid (IL) solutions with dynamically varying IL concentration and its impact on IL recycling. *ACS Sustainable Chemistry & Engineering* **2017,** *5* (5), 4408-4413.

3. Narron, R. H.; Kim, H.; Chang, H.-m.; Jameel, H.; Park, S. Biomass pretreatments capable of enabling lignin valorization in a biorefinery process. *Current Opinion in Biotechnology* **2016,** *38*, 39-46.

4. Renders, T.; Van den Bosch, S.; Koelewijn, S.-.-F.; Schutyser, W.; Sels, B. F. Lignin-first biomass fractionation: the advent of active stabilisation strategies. *Energy & Environmental Science* **2017,** *10*, 1551-1557.

5. Tao, L.; Aden, A.; Elander, R. T.; Pallapolu, V. R.; Lee, Y. Y.; Garlock, R. J.; Balan, V.; Dale, B. E.; Kim, Y.; Mosier, N. S.; Ladisch, M. R.; Falls, M.; Holtzapple, M. T.; Sierra, R.; Shi, J.; Ebrik, M. A.; Redmond, T.; Yang, B.; Wyman, C. E.; Hames, B.; Thomas, S.; Warner, R. E. Process and technoeconomic analysis of leading pretreatment technologies for lignocellulosic ethanol production using switchgrass. *Bioresource Technology* **2011,** *102* (24), 11105-11114.

6. Bozell, J. J.; Astner, A.; Baker, D.; Biannic, B.; Cedeno, D.; Elder, T.; Hosseinaei, O.; Delbeck, L.; Kim, J.-W.; O'Lenick, C. J.; Young, T. Integrating separation and conversion—conversion of biorefinery process streams to biobased chemicals and fuels. *BioEnergy Research* **2014,** *7* (3), 856-866.

7. Han, Q.; Jin, Y.; Jameel, H.; Chang, H.-m.; Phillips, R.; Park, S. Autohydrolysis pretreatment of waste wheat straw for cellulosic ethanol production in a co-located straw pulp mill. *Applied Biochemistry and Biotechnology* **2015,** *175* (2), 1193-1210.

8. Nitsos, C. K.; Matis, K. A.; Triantafyllidis, K. S. Optimization of hydrothermal pretreatment of lignocellulosic biomass in the bioethanol production process. *ChemSusChem* **2012,** *6* (1), 110-122.

9. Bozell, J. J.; Petersen, G. R. Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy's "Top 10" revisited. *Green Chemistry* **2010,** *12*, 539-554.

10. Narron, R. H.; Han, Q.; Park, S.; Chang, H.-m.; Jameel, H. Lignocentric analysis of a carbohydrate-producing lignocellulosic biorefinery process. *Bioresource Technology* **2017,** *241*, 857-867.

11. Rasmussen, H.; Tanner, D.; Sorensen, H. R.; Meyer, A. S. New degradation compounds from lignocellulosic biomass pretreatment: routes for formation of potent oligophenolic enzyme inhibitors. *Green Chemistry* **2017,** *19*, 464-473.

12. Akpinar, O.; Sabanci, S.; Levent, O.; Sayaslan, A. Evaluation of antioxidant activity of dilute acid hydrolysate of wheat straw during xylose production. *Industrial Crops and Products* **2012,** *40*, 39-44.

13. Lee, K. M.; Min, K.; Choi, O.; Kim, K.-Y.; Woo, H. M.; Kim, Y.; Han, S. O.; Um, Y. Electrochemical detoxification of phenolic compounds in lignocellulosic hydrolysate for *Clostridium* fermentation. *Bioresource Technology* **2015,** *187*, 228-234.

14. Faustino, H.; Gil, N.; Baptista, C.; Duarte, A. P. Antioxidant activity of lignin phenolic compounds extracted from Kraft and sulphite black liquors. *Molecules* **2010,** *15* (12), 9308-9322.

15. Mitchell, V. D.; Taylor, C. M.; Bauer, S. Comprehensive analysis of monomeric phenolics in dilute acid plant hydrolysates. *BioEnergy Research* **2014,** *7*, 654-669.

16. Boes, K. S.; Narron, R. H.; Chen, Y.; Park, S.; Vinueza, N. R. Characterization of biofuel refinery byproduct via selective electrospray ionization tandem mass spectrometry. *Fuel* **2017,** *188*, 190-196.

17. Montane, D.; Nabarlatz, D.; Martorell, A.; Torne-Fernandez, V.; Fierro, V. Removal of lignin and associated impurities from xylo-oligosaccharides by activated carbon adsorption. *Industrial Engineering & Chemistry Research* **2006,** *45*, 2294-2302.

18. Sainio, T.; Turku, I.; Heinonen, J. Adsorptive removal of fermentation inhibitors from concentrated acid hydrolyzates of lignocellulosic biomass. *Bioresource Technology* **2011,** *102* (10), 6048-6057.

19. Huang, C.; Jeuck, B.; Du, J.; Yong, Q.; Chang, H.-m.; Jameel, H.; Phillips, R. B. Novel process for the coproduction of xylo-olgiosaccharides, fermentable sugars, and lignosulfonates from hardwood. *Bioresource Technology* **2016,** *219*, 600-607.

20. Schwartz, T. J.; Lawoko, M. Removal of acid-soluble lignin from biomass extracts using Amberlite XAD-4 resin. *BioResources* **2010,** *5* (4), 2337-2347.

21. Soto, M. L.; Moure, A.; Dominguez, H.; Parajo, J. C. Recovery, concentration and purification of phenolic compounds by adsorption: a review. *Journal of Food Engineering* **2011,** *105*, 1-27.

22. Westerberg, N.; Sunner, H.; Helander, M.; Henriksson, G.; Lawoko, M.; Rasmuson, A. Separation of galactoglucomannans, lignin, and lignin-carbohydrate complexes from hot-water-extracted norway spruce by cross-flow filtration and adsorption chromatography. *BioResouces* **2012,** *7* (4), 4501-4516.

23. Koivula, E.; Kallioinen, M.; Sainio, T.; Anton, E.; Luque, S.; Manttari, M. Enhanced membrane filtration of wood hydrolysates for hemicelluloses recovery by pretreatment with polymeric adsorbents. *Bioresource Technology* **2013,** *143*, 275-281.

24. Giummarella, N.; Zhang, L.; Henriksson, G.; Lawoko, M. Structural features of mildly fractionated lignin carbohydrate complexes (LCC) from spruce. *RSC Advances* **2016,** *6*, 42120-42131.

25. Sluiter, J. B.; Ruiz, R. O.; Scarlata, C. J.; Sluiter, A. D.; Templeton, D. W. Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. *Journal of Agricultural and Food Chemistry* **2010,** *58* (16), 9043-9045.

26. Sluiter, A.; Hames, B.; Ruiz, R.; Scarlata, C.; Sluiter, J.; Templeton, D. Determination of sugars, byproducts, and degradation products in liquid fraction process samples. In *National Renewable Energy Laboratory Technical Report NREL/TP-510-42623*, 2008.

27. Yokoyama, T. Revisiting the mechanism of β-O-4 bond cleavage during acidolysis of lignin. Part 6: a review. *Journal of Wood Chemistry and Technology* **2014,** *35* (1), 165-174.

28. Matsumoto, Y.; Yokoyama, T.; Akiyama, T.; Shioya, T. Formation rate of benzyl cation intermediate from p-Hydroxyphenyl, guaiacyl, or syringyl nucleus in acidolysis of lignin. *Journal of Wood Chemistry and Technology* **2017,** *37* (2), 75-86.

29. Calvo-Flores, F. G.; Dobado, J. A.; Isac-Garcia, J.; Martin-Martinez, F. J. *Lignin and lignans as renewable raw materials: chemistry, technology and applications*. John Wiley & Sons, Ltd: West Sussex, U.K., 2015.

30. Chi, C.; Zhang, Z.; Chang, H.-m.; Jameel, H. Determination of furfural and hydroxymethylfurfural formed from biomass under acidic conditions. *Journal of Wood Chemistry and Technology* **2009,** *29* (4), 265-276.

31. Zhang, L.; Gellerstedt, G. Quantitative 2D-HSQC NMR determination of polymer structures by selecting suitable internal standard references. *Magnetic Resonance in Chemistry* **2006,** *45* (1), 37-45.

32. Balakshin, M.; Capanema, E.; Gracz, H.; Chang, H.-m.; Jameel, H. Quantification of lignin-carbohydrate linkages with high-resolution NMR spectroscopy. *Planta* **2011,** *233* (6), 1097-1110.

33. Capanema, E.; Balakshin, M.; Kadla, J. A comprehensive approach for quantitative lignin characterization by NMR spectroscopy. *Journal of Agricultural and Food Chemistry* **2004,** *52* (7), 1850-1860.

34. Ralph, J.; Akiyama, T.; Kim, H.; Lu, F.; Schatz, P.; Marita, J.; Ralph, S.; Reddy, M.; Chen, F.; Dixon, R. Effects of coumarate 3-hydroxylase down-regulation on lignin structure. *Journal of Biological Chemistry* **2006,** *281* (13), 8843-8853.

35. Yelle, D. J.; Ralph, J.; Frihart, C. R. Characterization of nonderivatized plant cell walls using high-resolution solution-state NMR spectroscopy. *Magnetic Resonance in Chemistry* **2008,** *46* (6), 508-517.

36. Yelle, D. J.; Kaparaju, P.; Hunt, C. G.; Hirth, K.; Kim, H.; Ralph, J.; Felby, C. Two-dimensional NMR evidence for cleavage of lignin and xylan substituents in wheat straw through hydrothermal pretreatment and enzymatic hydrolysis. *BioEnergy Research* **2013,** *6* (1), 211-221.

37. Kim, H.; Ralph, J.; Akiyama, T. Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d_6. *BioEnergy Research* **2008,** *1*, 56-66.

38. Du, X.; Gellerstedt, G.; Li, J. Universal fractionation of lignin-carbohydrate complexes (LCCs) from lignocellulosic biomass: an example using spruce wood. *The Plant Journal* **2013,** *74* (2), 328-338.

39. Du, X.; Perez-Boada, M.; Fernandez, C.; Rencoret, J.; del Rio, J.; Jimenez-Barbero, J.; Li, J.; Gutierrez, A.; Martinez, A. Analysis of lignin-carbohydrate and lignin-lignin linkages after hydrolase treatment of xylan-lignin, glucomannan-lignin and glucan-lignin complexes from sprucew ood. *Planta* **2014,** *239* (5), 1079-1090.

40. Zhang, L.; Gellerstedt, G.; Lu, F.; Ralph, J. NMR studies on the occurence of spirodienone structures in lignins. *Journal of Wood Chemistry and Technology* **2006,** *26*, 65-79.

41. Ralph, J.; Lundquist, K.; Brunow, G.; Lu, F.; Kim, H.; Schatz, P.; Marita, J.; Hatfield, R.; Ralph, S.; Christensen, J.; Boerjan, W. Lignins: natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochemistry Reviews* **2004,** *3*, 29-60.

42. Li, K.; Helm, R. Synthesis and rearrangement reactions of ester-linked lignin-carbohydrate model compounds. *Journal of Agricultural and Food Chemistry* **1995,** *43* (8), 2098-2103.

43. Leschinsky, M.; Zuckerstatter, G.; Weber, H. K.; Patt, R.; Sixta, H. Effect of autohydrolysis of *Eucalyptus globulus* wood on lignin structure. Part 1: Comparison of different lignin fractions formed during water prehydrolysis. *Holzforschung* **2008,** *62*, 645-652.

44. Leschinsky, M.; Zuckerstatter, G.; Weber, H. K.; Patt, R.; Sixta, H. Effect of autohydrolysis of *Eucalyptus globulus* wood on lignin structure. Part 2: Influence of autohydrolysis intensity. *Holzforschung* **2008,** *62*, 653-658.

TABLES

Table 1. Raw and autohydrolyzed biomass chemical compositions (g/100g o.d. raw biomass)

Biomass	Maple		Sugarcane bagasse	
	Raw	Autohydrolyzed	Raw	Autohydrolyzed
Gravimetric solid yield[a]	100	69.2 ± 0.4	100	59.0 ± 0.3
Glucan	43.2	40.1	42.3	37.3
Xylan	13.0	3.0	20.9	3.5
Galactan	0.7	*nd*[e]	0.9	*nd*
Arabinan + mannan[b]	3.4	1.5	2.6	0.4
Total carbohydrate	60.3	44.6	66.7	41.2
Acid insoluble lignin	22.1	20.4	18.2	13.2
Acid-soluble lignin	3.1	0.9	2.3	0.8
Total lignin	25.1	21.3	20.5	14.0
Ash	0.3	*Nd*	2.7	1.4
Extractives[c]	2.0	-	3.9	-
Unidentified	12.3	4.8	6.2	4.1
Autohydrolyzate-soluble lignin[d]	3.8, 15%[f]		6.5, 32%	

[a] Mass of water-washed solids after autohydrolysis
[b] Co-eluting monosaccharides
[c] Raw benzene-ethanol extractives measured gravimetrically, but fully removed prior to raw compositional analysis and autohydrolysis
[d] Back-calculated based on gravimetric yield and decrease to lignin content
[e] Not detected
[f] Degree of delignification (wt% of raw lignin)

Table 2. Quantification around soluble lignin adsorption from autohydrolyzates

Autohydrolyzate biomass origin	Maple		Sugarcane bagasse	
Soluble lignin abundance[a]	3.8		6.5	
Status	*reduced*	*as is*	*reduced*	*as is*
UV active components	Reduced lignin	Lignin, furfural, HMF	Reduced lignin	Lignin, furfural, HMF
Autohydrolyzate UV response[b]	42 ± 2	180 ± 5	60 ± 3	320 ± 10
After 20:1 resin adsorption[c]	5 ± 0	32 ± 2	7 ± 0	66 ± 3
After 10:1 resin adsorption[c]	3 ± 0	14 ± 1	5 ± 0	33 ± 2
% A removal at 10:1	93%	92%	92%	90%

[a] From Table 1
[b] Absorbance units * dilution factor
[c] X mL autohydrolyzate: 1 g o.d. resin

Table 3. Chemical composition (g/100g) of lignin adsorbates around standard protocol and lowered acid hydrolysis conditions

Adsorbate biomass origin	Maple			Sugarcane bagasse		
hr 72% H_2SO_4 hydrolysis	1	0	0	1	0	0
hr 3% H_2SO_4 hydrolysis	1.5	1.5	1	1.5	1.5	1
Acid-insoluble lignin	15.8	10.1	7.8	16.6	9.8	8.7
Acid-soluble lignin	19.9	23.8	26.7	9.5	13.8	15.2
Total lignin	35.7	33.9	34.5	26.1	23.6	23.9
Glucan	2.1	2.4	2.4	5.0	5.6	5.8
Xylan	34.6	37.8	39.1	31.2	35.3	38.6
Galactan	0.2	0.4	0.5	0.1	0.2	0.2
Arabinan + Mannan	4.7	5.9	6.3	2.6	3.2	3.4
Total carbohydrate	41.6	46.5	48.3	38.9	44.3	48.0
Mass balance	77.3	80.4	82.8	65.0	67.9	71.9

Table 4. Autohydrolyzate saccharide abundances around resin adsorption (g/100g o.d. raw biomass).

Biomass		Autohydrolyzate			After adsorption		
		Monomer	Oligomer[b]	Total	Monomer	Oligomer	Total
Maple	Glucan	0.6	0.5	1.1	0.6	0.4	1.0
	Xylan	2.9	6.5	9.4	2.9	4.3	7.2
	Galactan	0.4	0.3	0.7	0.4	0.4	0.8
	A + M[a]	1.0	0.8	1.8	1.0	0.6	1.6
Sugarcane bagasse	Glucan	0.4	1.5	1.9	0.4	1.4	1.8
	Xylan	1.4	9.0	10.4	1.4	5.2	6.6
	Galactan	0.3	0.5	0.8	0.3	0.5	0.8
	A + M	1.3	0.9	2.2	1.3	0.3	1.6

[a] Arabinan + mannan (co-eluting monosaccharides)
[b] Different of total and monomer

Table 5. Signal assignments for lignin structures in 2D-HSQC NMR spectra

Label	δ_C/δ_H	Assignment
C_β	53.1/3.49	C_β-H_β of phenylcoumaran structures (C)
B_β	53.7/3.05	C_β-H_β of resinol structures (B)
-OCH_3	55.9/3.73	C-H of methoxyls
A_γ	59.6-60.8/3.37-3.72	C_γ-H_γ of β-O-4 structures(A)
$Acyl_\gamma$	61.6/4.08	C_γ-H_γ of Cγ acylated side chains
UrE_γ	63.0/4.22	C_γ-H_γ of uronosil ester LCC (UrE)
B_γ	71.3/4.18	C_γ-H_γ of resinol structures (B)
A_α	72.6/5.00	C_α-H_α of β-O-4 structures (A)
$BE_{\alpha\text{-}1^\circ}$	81.3/4.75	C_α-H_α of benzyl ether (primary alcohol) LCC (BE)
$A_{\beta\ (G/H)}$	82.5/4.27	C_β-H_β of β-O-4 structures linked to G unit (A)
D_α	83.6/4.86	C_α-H_α of dibenzodioxicin (DBDO) (D)
B_α	84.9/4.69	C_α-H_α of resinol structures (B)
D_β	85.8/3.90	C_β-H_β of dibenzodioxicin (DBDO) (D)
$A_{\beta\ (S)}$	86.0/4.11	C_β-H_β of β-O-4 structures linked to S unit (A)
C_α	86.8/5.49	C_α-H_α of phenylcoumaran structures (C)
$GlcA_1$	97.2/5.18	C_1 of non-esterified 4-O-methylglucuruonic acid
$EGlcA_1$	98.9/4.70	C_1 of esterified 4-O-methylglucuruonic acid
PhG_1	100.0-103.0/4.89-5.20	C_1 of phenyl glycoside LCC (PhG)

Table 6. Quantitation of native chemical structures in lignin adsorbates (# / 100 aromatic rings)

Biomass Origin	Maple	Sugarcane bagasse
Uronosil ester (UrE)	7.6	5.2
Benzyl ether[a] (BE)	2.0	0.5
Phenyl glycoside (PhG)	0.6	3.8
Total LCC	10.2	9.5
Cγ acyl ($Acyl_\gamma$)	1.3	*nq*[b]
β-O-4' (A)	33.6	13.4
β-β' (B)	2.8	0.3
β-5' (C)	1.5	0.7
DBDO (D)	*nd*[c]	8.4

[a] Cα-1° OH only
[b] Not quantified
[c] Not detected

FIGURES

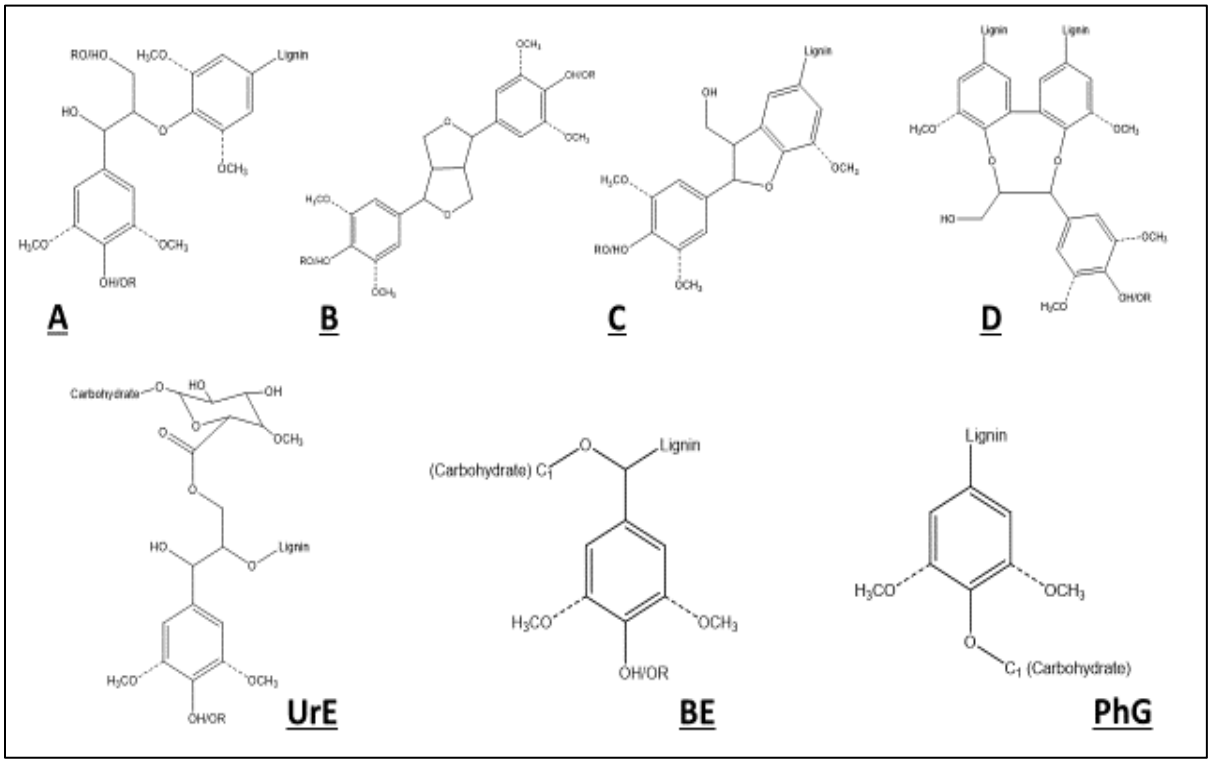

Figure 1. Native and LCC structures identified and quantified with NMR spectroscopy
A: β-O-4‘
B: β-β‘
C: β-5‘
D: Dibenzodioxicin (DBDO)
UrE: Uronosil ester LCC
BE: Benzyl ether LCC
PhG: Phenylglycoside LCC

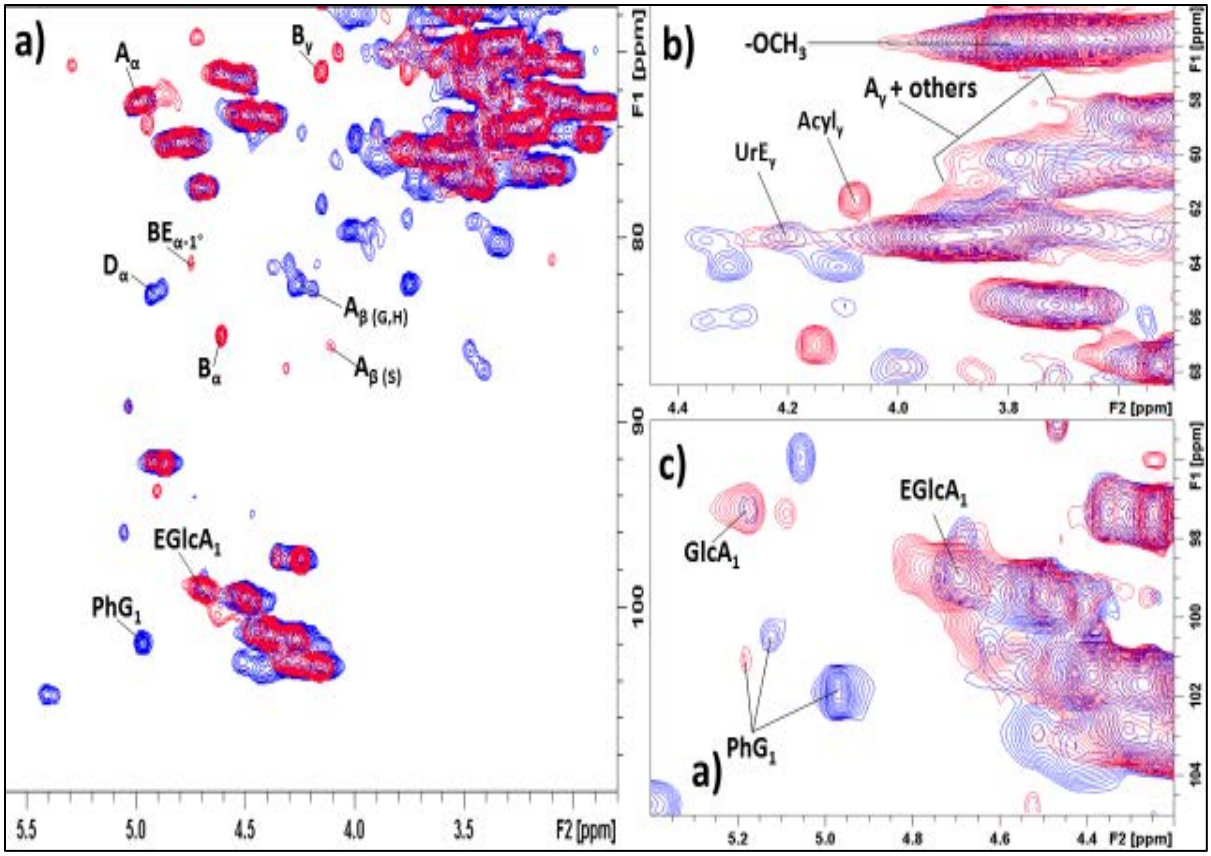

Figure 2. 2D-HSQC NMR spectra from sugarcane bagasse and maple adsorbates
Sugarcane bagasse: blue
Maple: red
a) Expanded aliphatic region displaying the quantified signals belonging to LCC and native lignin structures in adsorbates
b) Zoom-in of separated Cγ signals: uronosil ester and acyl ester
c) Zoom-in of phenyl glycoside signals and both esterified and non-esterified glucuronic acid moieties.

Chapter 4. Following a hardwood protolignin through biorefinery processing

Abstract

The effects of a biological-based biorefinery process upon protolignin properties are described in this work. Lignin preparations were procured from a hardwood's protolignin as well as from the resultant lignin fractions of biorefining (pretreatment-soluble (~15%) and cellulolytic hydrolysis-insoluble (~85%)). Two protolignin preparations were produced, representing ~50% of protolignin. Two accompanying biorefinery lignin preparations were also obtained, with the soluble lignin preparation being higher yield (98%) than the insoluble lignin preparation (30%). Native substructures and phenolic hydroxyls were quantified in all preparations using various NMR techniques. Soluble lignin was found to be heavily degraded but contained lignin-carbohydrate complexes ("LCCs", ~11/100 aromatic rings) and β-O-4' substructures (43/100 aromatic rings). Insoluble lignin retained few LCCs (~1/100 aromatic rings) and some β-O-4' (24/100 aromatic rings). Incorporation of lignin characterization into the biorefinery process' mass balance revealed ~50% β-O-4' recovery, and the majority of LCCs quantified in protolignin (3/100g o.d. raw biomass) were recovered in autohydrolyzate-soluble lignin (2/100g o.d. raw biomass). These findings demonstrate the specific effects of biorefining upon the studied hardwood's protolignin.

1. Introduction

Lignocellulosic biorefinery processes continue to warrant investigation in spite of a general absence of implementation at the industrial scale. This is in part due to the bleak economic prospects derived from technoeconomic analyses of hypothetical industrial-scale biorefinery processes.[1-3] Economic complications arise from biorefinery process' current dependence upon carbohydrate-derived products as the main source of revenue. A decline in demand for cellulosic (2nd generation) ethanol plus a lack of well-developed alternative valorization pathways for carbohydrates is to blame. To overcome this financial hindrance, focus has shifted towards upgrading non-carbohydrate biorefinery process streams into co-products in order to generate additional revenue streams.[4-6] It is believed that co-product valorization will provide significant security to a given industrial biorefinery's financial positioning.

A prime candidate for co-product valorization is biorefinery lignin, which has garnered interest due to its relative abundance and enticing chemical and molecular properties. However, monetizing biorefinery lignin presents an entirely new series of challenges to biorefinery scheming due to its significant heterogeneity. Our recent findings were able to relate the quantity and variety of non-condensed aromatic ring structures in hardwood lignins to the carbohydrate output of a biological biorefinery process (pretreatment and cellulolytic hydrolysis).[7] Said findings demonstrated that different protolignins subjected to identical biorefinery processing render lignin-rich process streams with varying chemical composition (g carbohydrate and g lignin /100 g oven-dry "o.d." raw biomass) and

disparate molecular properties. This variability adds significant financial risk to biorefinery investment when considering feedstock availability over prolonged time periods.

Prior to having the ability to address the effects of varying biomass feedstocks (i.e. varying protolignins) upon biorefinery lignin valorization, it is first necessary to systematically analyze how such biorefinery processing transforms the properties of a single protolignin. Sweetgum (*L. styraciflua*) was chosen due to its exhibition of median results in our previous cross-species work.[7] The process investigated is a technically feasible process which employs sequential autohydrolysis pretreatment and mechanical refining prior to cellulolytic hydrolysis.[7-8] Said process ultimately yields both a soluble (minor) and an insoluble (major) lignin stream. Until recently, complete analysis of biorefinery lignin was hindered by the difficulty in isolating the soluble lignin fraction in pretreatment hydrolyzate ("autohydrolyzate") for characterization. However, our recent research effort successfully isolated and characterized the lignin soluble in autohydrolyzate with both selectively and strong isolation yields (~90%).[9-10] Apart from the portion of lignin solubilized during pretreatment is that which remains insoluble throughout pretreatment and cellulolytic biotransformation. This lignin fraction represents the bulk of lignin produced by the biorefinery process, therefore its efficient utilization is imperative to maximize lignin contribution to economics. Previous characterization of this lignin fraction indicated an unknown extent of inter-lignin condensation. Additional characterization was not performed, leaving room for further elucidative insoluble lignin characterization

In order to comprehensively understand the effects of biorefinery processing upon a given biomass' protolignin, we began with isolation and characterization of representative

lignin preparations from the raw hardwood biomass. With protolignin characterization in hand, representative lignin preparations were procured from autohydrolyzate (soluble lignin) and cellulolytic hydrolysis residue (insoluble lignin). Both biorefinery lignin preparations were characterized, and informed comparisons were made between the properties of protolignin and biorefinery lignins, including framing characterization results within the process' mass balance. All lignin preparations were characterized for preparation purity, preparation yield (sample representativeness), inter-lignin chemical linkages, and hydroxyl functionalities. The goal of this work is to definitively understand the properties of biorefinery lignin fractions as they relate to a given protolignin source to assist lignin valorization efforts.

2. Materials and methods

2.1. Raw biomass and protolignin isolation

Hardwood chips from sweetgum (*Liquidambar styraciflua*) were obtained from the Tree Improvement Program at North Carolina State University (Raleigh, NC, USA). Wood chips were milled by Wiley mill to produce woodmeal. Additional wood chips were retained for biorefinery processing. Woodmeal and wood chips were extracted overnight in a soxhlet extractor with 2:1 (V) benzene-ethanol solution. The quantity of extractives in woodmeal was measured gravimetrically around extraction. Only extractive-free wood chips and woodmeal were used for experimentation.

A planetary ball mill (Pulverisette 7 premium line, Fritsch, Germany) was used to enable protolignin extraction. 2g of air-dried extractive-free woodmeal was added to each

ball mill capsule in addition to nine ZrO_2 balls. Samples were milled at 600 rpm for 6 hrs, with 30 minutes of instrument cool-down per 15 minutes of milling. After milling, woodmeal was dried in a vacuum oven at 35 °C over solid P_2O_5 until constant mass. Next, dried woodmeal was subjected to extraction using 96% 1,4-dioxane/4% water solvent (volume %). Solid and extract were separated using vacuum filtration and Gooch crucibles (fine grade). Extraction was considered complete once resultant solvent filtrate reached complete visual transparency. Next, extracted solids were oven-dried at 105 °C until constant mass to remove residual solvent. Once dried, the material was subjected to high-dosage (20 FPU / o.d. g) cellulolytic hydrolysis at 50 °C. An enzyme cocktail of Cellic CTec2 (Novozymes, Denmark) and Cellic Htec2 (1/9 volume of Ctec2) was applied in an acetate buffer system (pH ~5). ~180 rpm of system agitation was utilized. Following 72 hr of cellulolytic hydrolysis, remaining solids were extensively washed with deionized water to remove free protein and buffer. Washing was considered complete once wash water filtrate reached neutrality. Washed solids were then dried in vacuum oven as before. Dried solids were again extracted with the same solvent and methodology as was applied for the first extraction. Both extract solutions were evaporated by rotary evaporator (Rotavapor® R-300, Buchi, USA), re-suspended into deionized water then frozen, and finally freeze-dried (FreeZone 6, Labconco, USA) to produce solids. All crude protolignin extracts were stored in vacuum oven until constant mass prior to purification or characterization. To measure gravimetric yield of crude protolignin preparations, the mass of extracted solids relative to the un-extracted mass was recorded (i.e. weight loss).

2.2. Protolignin purification

Crude protolignin preparations were purified using a dissolution-precipitation method. First, samples were dissolved into 90% acetic acid/10% water solution. After complete dissolution, lignin-containing solution was precipitated drop-wise into excess deionized water (20x). Precipitates were collected using vacuum filtration, and washed with deionized water until wash water filtrate reached neutrality. Precipitates were allowed to air dry, and next were stored in vacuum oven until constant mass prior to characterization. To measure loss of preparation yield induced by purification, the mass of purified solids was measured relative to original crude solid mass.

2.3. Biorefinery processing

Autohydrolysis pretreatment was performed upon extractive-free woodchips in a 1L heated Parr reactor (Parr Instrument Company, USA). Pretreatment temperature was set to 180 °C with ~30 minutes of temperature ramp-up. Pretreatment lasted for 40 mins at temperature, followed by rapid thermal quenching in an ice water bath. After sufficient cooling, autohydrolyzed solids were separated from autohydrolyzate using vacuum filtration. Bulk autohydrolyzate obtained was filtered again using Gooch crucibles (fine grade) to ensure removal of suspended solids. Dually-filtered autohydrolyzates were bottled and kept in a refrigerator at 4 °C prior to further handling. Autohydrolyzed solids were washed with deionized water to remove residual autohydrolyzate until wash water filtrate reached neutrality. The mass of washed solids (accounting for moisture) relative to the mass of the original woodchips was recorded to obtain gravimetric solid yield.

Washed autohydrolyzed solids were next mechanically refined for 6000 revolutions using a laboratory PFI mill. After refining, the solids were subjected to the same cellulolytic hydrolysis protocol applied to woodmeal. After 72 hrs of incubation with enzymes, the solids were extensively washed with deionized water to remove free protein and buffer solution until wash water filtrate reached neutrality. The residual solids were air-dried and then stored in vacuum oven until constant mass prior to lignin extraction.

2.4. Obtaining biorefinery lignin preparations

Autohydrolyzate-soluble lignin was obtained from autohydrolyzate using Amberlite® XAD16N adsorptive resin (Dow Chemical Company, USA). Autohydrolyzate (free of suspended solids) was mixed with a known amount of resin to create autohydrolyzate-to-resin ratio of 10:1 (o.d. g resin:mL autohydrolyzate). After 30 minutes of low speed mixing, resin and solution were separated by vacuum filtration. Adsorbate-containing resin was subjected to water washing to remove un-adsorbed autohydrolyzate solutes. Next, methanol was poured over the captured resin until methanol filtrate reached complete visual transparency. Methanol consumption was ~3x the amount of starting autohydrolyzate. Methanol solution with adsorbates was evaporated and freeze-dried in the same manner as protolignin preparations. Solid adsorbates were kept in vacuum oven until constant mass before characterization. To measure extent of lignin adsorption, an aliquot of autohydrolyzate and resin-treated autohydrolyzate were reduced using sodium borohydride to gather a lignin-only UV response. UV absorption at 280 nm was measured (Lambda XLS,

PerkinElmer, USA), and the difference in UV response was taken as the extent of lignin adsorption.

Dried cellulolytic hydrolysis-insoluble solids were extracted identically to protolignin preparations. Lignin-containing solvent was then handled in the same manner as previously noted. In addition, the preparation was purified by the same methodology used with both protolignin preparations.

2.5. Compositional analysis

The chemical composition of woodmeal, autohydrolyzed solids, and all lignin preparations was performed using the standard protocol provided by the National Renewable Energy Laboratory.[11] ~100 mg of carefully weighed o.d. adsorbate was subjected to 72% (w/w) sulfuric acid digestion for 1 hr, followed by dilution to 3% (w/w) H_2SO_4 and autoclaving for 1.5 hr. Per previous work, lignin from autohydrolyzate was only subject to 1 hr of 3% H_2SO_4 digestion to preserve dehydration-prone carbohydrates.[10] An aliquot of acid digestate was taken to determine acid-soluble lignin quantities spectrophotometrically and monosaccharides by liquid chromatography. For acid-soluble lignin, UV absorbance was measured at 205 nm, accounting for dilution. The molar absorptivity coefficient 110 L g^{-1} cm^{-1} was applied to convert absorbance to concentration. For monosaccharides, an HPLC system (Agilent 1200 series, Agilent, USA) was used. In this setup, a Shodex SP-0810 column is utilized (8x300 mm, Showa Denko, Japan) and maintained at 80 °C. Mobile phase was Milli-Q water, flowing 0.5 mL/min for 60 minutes per sample. A refractive index detector monitored eluents. Before injection, acid digestate samples were neutralized to pH ~6 with $CaCO_3$ and

filtered through 0.2 μm nylon filters. Calibration was performed with commercial standards of glucose, xylose, galactose, mannose, and arabinose. Due to observed co-elution of mannose and arabinose, quantification of either saccharide is reported as the sum of both. All values represent the average of duplicate analyses, and average standard deviation was 2.3%.

2.6. Lignin preparation characterization

To estimate molecular weight of the biorefinery lignin preparations, a published method utilizing size-exclusion chromatography (Shimadzu, Japan) calibrated to polystyrene standards was applied.[12] The system was equipped with HR1 and HR5E Styragel columns (Waters, USA) and dual detectors, which are a SPD-20A UV/Vis and a RID-10A refractive index. The oven maintained 35°C and the wavelength for the UV detector was set to 280 nm. Samples were modified by acetylation prior to injection to ensure total solubility in the mobile phase, tetrahydrofuran. Acetylation involves dissolution of lignin preparation into pyridine, addition of excess acetic anhydride, and 72 hrs of reaction time in darkness. The acetylation slurry was poured over deionized ice and left to settle overnight. Solid precipitate was collected via vacuum filtration and washed excessively with deionized water. After washing, samples were stored in a vacuum oven until constant mass. Prior to injection, the acetylated lignins were dissolved into tetrahydrofuran at an approximate concentration of 0.5 mg acetylated lignin per mL tetrahydrofuran and finally filtered through 0.2 μm nylon filters. The mobile phase flow rate was set to 0.7 mL/min. The calibration curve utilized for estimating molecular weight was based on twelve polystyrene dispersions (PSS ReadyCal-Kit, Polymer Standards Service, Germany).

Quantification of hydroxyl groups utilizing ^{31}P NMR was performed utilizing an established method.[13-14] ~40 mg of lignin sample was massed directly inside of a 5 mm NMR tube. After dissolution and internal standard addition, 100 μL of organophosphitylating agent (2-chloro-4,4,5,5-tetramethyl-1,2,3-dioxaphospholane; 95%, Sigma-Aldrich, USA) was introduced and the sample immediately subjected to observation by ^{31}P NMR. The chemical shifts generated from organophosphitylation were recorded using a 300 MHz NMR spectrometer at 25 °C equipped with a 5mm QNP probe (Avance, Bruker, USA).

All 2D-HSQC and ^{13}C NMR spectra were acquired at 25°C using dimethyl sulfoxide-d6 (99.9%, VWR, USA) as a solvent. The concentration of lignin in each sample was ~20% w/v. 2D-HSQC was performed first, followed by ^{13}C upon the same sample. Chromium (III) acetylacetonate (99.99%, Sigma-Aldrich, USA) was added to the NMR tube (0.01 M) following 2D-HSQC analysis to assist in relaxation of carbon nuclei during ^{13}C observation. 2D-HSQC NMR spectra were acquired on a 500 MHz NMR spectrometer equipped with 5 mm BBI probe with Z-axis gradient (Avance, Bruker, USA). A coupling constant ($^{1}J_{C-H}$) of 147 Hz was applied. Quantitative ^{13}C NMR spectra were acquired on the same 500 MHz spectrometer with a different probe than before: 5 mm BBO probe with Z-axis gradient. An inverse gated proton decoupling sequence was applied.

3. Results and discussion

3.1. Protolignin isolation

The isolation scheme for procurement of protolignin preparations is visualized in Figure 1. As can be seen, two protolignin isolation techniques were employed to enable

addressment of preparation characteristics attributable to isolation methodology. A necessary step in protolignin isolation, planetary ball milling of lignocellulosic material enables lignin solubility in lignin-friendly solvents through degradation of native lignin to the point of solubility. Following milled wood lignin ("MWL") extraction, the remaining insoluble material was next subjected to high-dosage cellulolytic hydrolysis to render an additional lignin fraction extractable via carbohydrate scission.[15] The newly extractable lignin (cellulolytic enzyme lignin, "CEL"), while still having been subjected to ball milling, bears a lessened extent of ball mill-borne degradation. For this reason, results from the two protolignin preparations are subject to more informed characterization and discussion compared to analysis of a single protolignin preparation isolated following multiple steps of solubility-rendering operations.[16]

3.2. Raw hardwood and protolignin preparation chemical composition

Displayed in Table 1 is the chemical composition of raw hardwood alongside the chemical compositions of each protolignin preparation. Preparation yield (or preparation's protolignin representativeness) was calculated based upon the gravimetric mass of each isolate procured (g preparation/100g total raw lignin) and isolate lignin purity (wt% total lignin). First, it can be seen that the employed ball milling time (6 hr) enabled extraction of a 38% representative MWLc preparation. This value is well over 20%, previously identified as a benchmark yield based upon another work's conclusion that ~20% preparations are sufficiently representative of raw lignin (however, only based on wood and protolignin preparation syringyl-to-guaiacyl ratio).[16] CELc preparation yield was found to be 19%, also

in proximity to the previously described protolignin preparation benchmark. Additional ball milling time would certainly have enabled higher preparation yield values for both MWLc and CELc, but at the expense of additional degradation. For this reason, ball milling time was not extended. It is possible that <6 hr of ball milling could also have satisfied the yield benchmark, however, such experiments were un-performed based upon the observed yield loss induced by purification.

Purification of each crude preparation was performed due to carbohydrates (either free or chemically-attached to lignin) also being extractable by the employed lignin solvent (96% 1,4-dioxane/4% water solvent, volume %). Separation of unbound carbohydrates enables stronger lignin resonance on NMR spectra, of significant benefit to substructure quantification. While the purification step was successful in lowering preparation carbohydrate contents (MWLc: 16.2 g total carbohydrate, MWLp: 5.9g total carbohydrate/100g o.d. solid), it also resulted in decreases to preparation yield. The decrease was most pronounced for MWL (38% to 23%), but was also experienced by CEL (19% to 16%). We suspect this phenomenon to be a function of lignin-carbohydrate complexes ("LCCs") present in both crude preparations. It is likely that the purification methodology resulted in some LCC-bearing lignins avoiding precipitation during the applied purification process. To confirm this hypothesis, LCC substructure quantification was performed on both MWLc and MWLp, due to MWLc being the most negatively affected by the employed purification operation.

It is also important to address the extent of uncharacterized components within each characterized solid. The amount of uncharacterized components in raw woodmeal, MWLc,

and CELc was 14.1 wt%, 15.4 wt%, and 11.4 wt% (respectively). This value is disappointing yet reproducible, suggesting notable quantities of components not specifically enumerated during conventional compositional analysis. It is our hypothesis that additional quantitation of hemicellulosic uronic acid and acetic acid substituents would result in lowered quantities of uncharacterized components. In agreement with our hypothesis that the aforementioned carbohydrate-linked moieties are leading to the high quantities of uncharacterized components, it can be observed that the quantity of un-characterized components decreased following carbohydrate-targeting purification. Specifically, the amount of uncharacterized components in MWLp and CELp was 7.2 wt% and 3.1 wt% (respectively), much lower than 15.4 wt% (MWLc) and 11.4 wt% (CELc).

3.3. Quantitation of protolignin preparation hydroxyl groups

Table 2 displays the quantity and variety of hydroxyl functional groups in both purified lignin preparations. No significant difference was observable between aliphatic hydroxyl quantities in MWLp and CELp. In addition, the presented aliphatic hydroxyl quantities are inflated by the presence of carbohydrates in the purified preparations. For this reason, Table 2 also contains aliphatic hydroxyl quantities in units of mmol/g solid, in acknowledgement of carbohydrate infiltration. It can be speculated that carbohydrates remaining in purified protolignin preparations (MWLp: 5.9 wt% total carbohydrate; CELp: 6.5 wt% total carbohydrate) are directly or indirectly linked to lignin through LCC substructures. This speculation will be evaluated through substructure quantification with NMR.

Concerning phenolic hydroxyls, it is first important to explain the logic behind the labels listed for each type of phenolic hydroxyl. The chemical shift of organophosphoryl-derivatized phenolic hydroxyls is most affected by adjacent substituents (aromatic C3 or C5). It was recently concluded that the spectral region pertaining to syringyl ("S") phenolic hydroxyls, which feature dual methoxyl groups, can also contain chemical shifts produced from condensed guaiacyl ("G") phenolic hydroxyls.[14] Condensed G lignin structures feature dual substitution as well, however, the substituents are one methoxyl group and one carbon-carbon bond. For this reason, phenolic hydroxyls are reported based upon the defining characteristics of their chemical shift range, i.e. bonding at aromatic C3 and C5.

The most notable difference between MWLp and CELp is the difference in total phenolic hydroxyl groups, measured as 1.8 and 1.1 mmol/g total lignin for MWLp and CELp, respectively. Phenolic hydroxyl groups are generally produced through degradation of native ether-containing substructures, in addition to those formed during biosynthesis. In agreement with previous discussion concerning the degradative effects of ball milling, MWLp features ~40% more phenolic hydroxyl groups compared to CELp. This supports previous discussion around the extent of ball mill-induced degradation within the MWL preparation versus the CEL preparation. Nevertheless, the amounts of phenolic hydroxyls measured in CELp are also likely artifacts of ball milling.

3.4. Native substructure quantification with combined 2D-HSQC and ^{13}C NMR analysis

Native substructures were identified with 2D-HSQC and quantified utilizing both 2D-HSQC and ^{13}C NMR spectra from the protolignin preparations, as described in other

publications.[17-19] Within the quantification method, signals separated by the 2D-HSQC observation are individually integrated and interconverted into ^{13}C base units of #/100 aromatic rings (#/100 Ar). Table 3 is a list of 2D-HSQC chemical shifts related to individual positions within native lignin substructures, as well as those produced by aromatic ring structures. Chemical shifts were identified from reputable publications concerning substructure chemical shift elucidation and substructure quantification.[18-22]

Figure 2 displays the 2D-HSQC (with substructure identification) and ^{13}C spectra obtained from MWLp and CELp. Most substructures denoted in Table 3 were identifiable, with the exception of LCC structures and hemicellulosic uronic acids (non-esterified and esterified). The absence of these signals is in agreement with previously speculated LCC removal during the applied purification protocol. For this reason, MWLc 2D-HSQC and ^{13}C were also acquired to confirm speculation about LCC's fate around purification. The absence of significant LCC resonance from MWLp and CELp suggests the majority of carbohydrates quantified in either preparation (~6 g total carbohydrate/100 g o.d. solid) are indirectly bound to lignin through a minor amount of LCC substructures, which fail to strongly resonate with 500 MHz NMR due to inherently low quantities.

Figure 3 displays LCC signal-containing regions from 2D-HSQC spectra, with MWLc signals overlaid upon MWLp. All three LCC substructures not distinctively resonating on MWLp 2D-HSQC spectra are more distinguishable on the 2D-HSQC spectra acquired from MWLc. It is important to note that the signal resonating at δ_C/δ_H 81.2/5.13 is known to be an overlapped signal containing resonance from Cα of spirodienone substructures and Cα of benzyl ether LCCs with attachment to secondary hydroxyls of carbohydrate.[18] Due to this

overlap, the signal was also visible in MWLp spectra due to the presence of spirodienones. To account for the overlap, spirodienone substructures must be quantified using resonance from Cβ, allowing the aforementioned benzyl ether LCC substructures to be semi-quantified using the overlapped signal integral minus the spirodienone Cβ resonance integral. In conclusion, the presence of LCC signals in MWLc supports the hypothesis that the purification methodology utilized resulted in inadvertent LCC removal, resulting in preparation yield loss. This conclusion was also drawn in another work which sought to maximize protolignin characterization using protolignin preparations.[18] To obtain protolignin LCC quantities, quantification was also performed upon MWLc.

3.5. Native substructure quantification using combined 2D-HSQC and ^{13}C spectra

The following equation was used for quantification of lignin substructures:

$$\frac{X}{100\ Ar} = \frac{2D_x}{2D_{IS}} * \frac{13C_{IS}}{13C_{163-103}} * 600$$

where X is the substructure being quantified, $2D_x$ is the integral of the 2D-HSQC signal used for structural quantification, IS the integral cluster or region for 2D-HSQC and ^{13}C (respectively), $^{13}C_{163\text{-}103}$ is the integral of the aromatic region of ^{13}C spectra, and 600 is derived from 600 aromatic carbons per 100 aromatic rings. Computed values are viewed as the quantity/percentage of substructures per lignin unit. Unless otherwise noted in discussion, Cα substructure signals are integrated for quantitation. In addition, non-condensed S/G ratio was measured utilizing the following equation:

$$S/G = \frac{\left(\frac{2D_{S(2,6)} + 2D_{S'(2,6)}}{2}\right)}{\left(2D_{G(6)} + 2D_{G'(6)}\right)}$$

where $2D_{S \text{ or } G}$ denotes the aromatic signals integrated for evaluation of the ratio between S and G aromatic ring structures. Finally, substructure quantities denoted as not detected (*nd*) failed to distinctively resonate in separated fashion at any spectral contour. For the substructures labeled as not quantified (*nq*), their computed values were less than 0.5/100 Ar. All integrals relating in substructure values of <0.5/100 Ar were rejected based upon instrument insensitivity. While the 500 MHz NMR used for this work is powerful, we do not feel able to accurately quantify structures whose signal integrals are so minor, especially considering that analysis was not performed in duplicate. In addition, reported substructure quantities have been rounded to the ones place to reflect this. Quantitation of non-condensed S/G and native substructures is shown in Table 4.

Substructure quantification in protolignin preparations revealed several distinguishable differences between protolignin preparations attributable to differences in procurement. First, it can be seen that purification of MWLc mostly resulted in losses to uronosil ester LCC substructures, as well as preparation non-condensed S/G. This confirms prior discussion regarding the loss of preparation yield associated with the applied purification methodology. The decrease in non-condensed S/G (MWLc: 2.7; MWLp: 2.0) also indicates that the (predominantly uronosil ester) LCCs removed through purification are associated with S lignin structures. This finding is significant in that it attributes uronosil ester LCC substructures to a specific monolignol constituent (S). Surprisingly, total β-O-4' abundance was found to not differ between MWLc and MWLp, suggesting that removal of

LCC structures does not result in significant re-distribution of β-O-4' structures, despite the loss in preparation yield. Minor native substructures (β-β', β-5', spirodienone) were each slightly higher following purification, confirming that the carbohydrate-removing purification did indeed render MWLp with higher protolignin representativeness, based upon elevated native substructure quantities.

Comparing MWLp to CELp, first it can be seen that the non-condensed S/G of CELp is the same value as was represented in MWLc. This indicates that CELp is more representative of protolignin non-condensed S/G. Next, β-O-4' abundance was much higher in CELp compared to either MWL preparation. This is in agreement with previous discussion concerning the extent of ball milling degradation imparted upon each respective lignin preparation. Clearly, ball milling exerts great influence upon β-O-4' abundance. However, as noted before, the increased β-O-4' values in CELp are still likely underestimations of their abundance in protolignin due to the application of planetary ball milling. Regardless, the results prove that independently isolated CEL preparations are more representative of protolignin β-O-4' content compared to MWL preparations. Beyond β-O-4', no significant difference in substructure abundance could be observed.

3.6. Autohydrolysis mass balance of sweetgum's chemical components

Autohydrolysis is recognized as a favorable pretreatment for biorefinery processing due to its low-cost implementation and enabling effects upon downstream cellulolytic saccharification.[7-8, 23-24] The lignocellulose constituent most affected by the water-only pretreatment is hemicellulose, in addition to a portion of lignin. At elevated temperatures,

cleavage of labile hemicellulosic acetyl substituents is incited, doping the starting water with acetic acid to create a mildly acidic medium. The ensuing mildly acidic solution facilitates hydrolysis of hemicellulosic glycosidic bonds. Native ether- and ester-containing lignin substructures are also subject to hydrolysis (ether hydrolysis commonly referred to as acidolysis[25-26]), depending on pretreatment severity. The autohydrolysis conditions utilized (180 °C, 40 min, 10:1 liquid:solid) are seen as moderate based upon works focusing upon time-temperature relationship of biomass and autohydrolysis outcome.[23, 27-28] Additionally, these conditions enabled high cellulolytic saccharification yields at low enzyme dosages (5 FPU/o.d. g solid), the focus of a previous work.[7] Table 5 contains the chemical compositions of raw sweetgum woodmeal (from Table 1), the resultant residue from autohydrolysis, and quantities of both soluble sugars and lignin in autohydrolyzate.

Results from Table 5 demonstrate that hydrolysis of hemicellulosic carbohydrates (mostly xylan) is greatly induced by autohydrolysis pretreatment (raw: 15.7 g xylan; residue: 3.7 g xylan/100g o.d. raw biomass). In addition, a minor portion of lignin is rendered soluble, found to be 3.8g soluble lignin/100g o.d. raw biomass, or ~15 wt% of total raw lignin. As mentioned previously, the majority of the investigated biorefinery's lignin remains solid throughout biorefinery processing, yet a minor fraction is rendered soluble in autohydrolyzate. To understand lignin's properties in both autohydrolyzate and post-enzymatic hydrolysis residue, two lignin preparations were procured- one from autohydrolyzate and the other from the solid residue remaining after cellulolytic hydrolysis. Both biorefinery lignin isolation protocols are graphically displayed in Figure 4.

3.7. Isolation of biorefinery lignin fractions

Autohydrolyzate-soluble lignin (~15 wt% total raw lignin) was selectively adsorbed utilizing an aromatic-selective resin, Amberlite® XAD16N. This resin engages hydrophobic solutes in aqueous medium through π-π' and van der Waal's intermolecular forces. Said intermolecular interactions enable a selective separation methodology that renders adsorbates free from unbound mono- and oligosaccharides. Development of the adsorption protocol and characterization of lignin adsorbates obtained from different biomasses is the focus of a previous work[10], intended to obtain representative autohydrolyzate-soluble lignin preparations for characterization. For this work, the same adsorption protocol and yield determination methodology are applied to sweetgum autohydrolyzate. Lignin adsorbate is appropriately labeled as "XADL", reflecting its mode of isolation.

Autohydrolyzed solids (containing ~85 wt% total raw lignin) were subjected to further biorefinery processing to produce biorefinery residual solids (insoluble material from sequential pretreatment and cellulolytic hydrolysis). A key difference between the biorefinery processing applied in this work and that which is described in our previous work[7] is a pronounced increase in cellulase cocktail dosage (20 FPU vs. 5 FPU/o.d. g solid). The reasoning behind the enzyme charge adjustment was to create an extractable lignin fraction that can be treated as a CEL preparation. For this reason, lignin preparations from cellulolytic hydrolysis residues are labeled as "AHCEL". Chemical composition of XADL and AHCEL lignin preparations, as well as preparation yields, are displayed in Table 6 and Table 7.

The first item to address is the approach towards quantifying soluble lignin isolation through adsorption to XAD16N resin (Table 6). Reduction of autohydrolyzate solutes with borohydride ($NaBH_4$) enables acquisition of a lignin-only UV response at 280 nm. This is made possible by reduction of UV-active furan aldehydes in autohydrolyzate to UV-silent furan alcohols, leaving the remaining absorbance attributable to lignin. This technique was proven successful in another work which monitored lignin and furan aldehyde quantities in a similar solution.[29] From the results shown in Table 6, significant lignin adsorption occurred, rendering nearly lignin-free autohydrolyzate. This means the autohydrolyzate-soluble lignin preparation (XADL) possesses almost complete representativeness. Continuing to Table 7, it can be seen that the chemical composition of XADL is remarkably carbohydrate-enriched (47.2 g xylan, 57.1 g total carbohydrate/100g o.d. solid adsorbate). This result is in agreement with our previous work observing carbohydrate-enrichment in XAD16N adsorbates from autohydrolyzate. The conclusion of said work revealed carbohydrate enrichment to be a function of high quantities of LCC substructures. Similar abundancy of LCC substructures is expected to be quantifiable in the autohydrolyzate-soluble lignin preparation generated in this effort. Also, the amount of uncharacterized components in XADL (14.0g/100g o.d. solid adsorbate) suggests notable quantities of quantitation-eluding hemicellulosic substituents are present within the adsorbate. This is consistent with previous observation and discussion around the extent of unidentified chemical components in protolignin preparations.

The approach to quantifying preparation yield from residual solid biorefinery lignin is the same as it was for protolignin preparation procurement. Unlike protolignin preparations

which have been subjected to planetary ball milling, the solids which AHCEL is procured from have only undergone biorefinery processing. Because of this, the obtained lignin preparations are solely representative of changes induced to lignin from biorefining. AHCEL preparation yields were determined using the amount of lignin in starting autohydrolyzed solids (21.4 g/100g o.d. solid) and the preparation's gravimetric yield, with correction for carbohydrates. We have assumed that the quantity of lignin in autohydrolyzed solids remains constant throughout mechanical refining and cellulolytic hydrolysis. Preparation yields of AHCELc and AHCELp were found to be 34% and 30% (respectively) of autohydrolysis-insoluble lignin. As with CELp, only minimal yield loss was observable during AHCEL purification (4%). For this reason, AHCELp is the only AHCEL preparation subject to characterization. Unfortunately, there is currently no published research aimed at relating preparation yield from pretreated material to preparation representativeness. Because of this, the large amount of un-extractable lignin (~65% of autohydrolyzed solid lignin) necessitates future work to maximize preparation yield from autohydrolyzed solid lignin. However, this is likely only achievable through implementation of degradative planetary ball milling or mild acidolysis to produce enzyme mild acidolysis lignin preparations ("EMAL").[30-31] Concerning AHCEL chemical compositions, the high total lignin content of AHCELc and AHCELp (82.7 & 90.4 wt% total lignin) are the result of the material having undergone hemicellulose-affecting pretreatment and cellulolytic hydrolysis. For this reason, a quantifiable presence of LCC substructures is not expected. Finally, accompanying purification is the observation of decreasing amounts of uncharacterized

components (AHCELc: 6.0% unidentified; AHCELp: 3.1% unidentified). This observation is congruent with similar observations made during protolignin preparation purification.

3.8. Hydroxyl group quantitation and molecular weight estimation of biorefinery lignin preparations

Table 8 displays the quantities of hydroxyl functional groups measured in both biorefinery lignin fractions and approximate molecular mass of each. First, it can be seen that the amount of aliphatic hydroxyls significantly varies between XADL and AHCELp. As discussed during protolignin preparation hydroxyl quantification, XADL's carbohydrate content (57.1 g total carbohydrate/100g o.d. adsorbate) exerts strong influence on the quantity of aliphatic hydroxyl groups measured (28.5 mmol/g total lignin in adsorbate). When quantification is re-computed in terms of mmol per g solid, aliphatic hydroxyl quantities in XADL drastically decrease, attributable to the preparation's relatively low lignin content (28.9 g/100g o.d. adsorbate). Such computation has only a minor effect on AHCELp aliphatic hydroxyl quantities due to its minimal carbohydrate constitution (6.5 g total carbohydrate/100g o.d. solid). Comparing phenolic hydroxyl groups, XADL possesses ~40% more phenolic hydroxyls than AHCELp. Because formation of phenolic hydroxyl groups is associated with native substructure acidolysis, the magnitude of phenolic hydroxyls measured in XADL suggests that autohydrolyzate-soluble lignin is significantly more degraded than the lignin that remains insoluble throughout biorefinery processing. However, the lower phenolic hydroxyl quantities in AHCELp are not a full assessment of residual solid lignin's phenolic functionality. This is because said preparation represents 30% of the

original autohydrolyzed lignin. It is possible that the un-extractable lignin possesses more phenolic functionality, but it was not isolatable by the applied methodology. Possible explanations for non-extractability could be minimal substructure degradation induced by autohydrolysis, inter-lignin condensation induced by autohydrolysis, or most likely, a combination of both. Regarding phenolic hydroxyl profiles, both preparations are dominated by C3 & C5 bound phenolic hydroxyls. As mentioned earlier, this can only be taken to indicate that the most abundant phenolic hydroxyl group in either preparation hails from both S lignin structures and condensed G lignin structures. However, measurement of non-condensed S/G by NMR should enable some differentiation between which phenolic functionality is more prominent in either biorefinery lignin preparation.

Molecular mass was only estimated for biorefinery lignin preparations because protolignin preparation molecular masses are exclusively a function of isolation methodology. However, as stated, biorefinery lignin preparations are solely altered by biorefinery processing itself. This enables informative evaluation of the molecular mass for considering the extent of molecular degradation. Prior to discussion, however, it is important to address the language used around the molecular mass estimation method applied, specifically the word “estimation”. The size-exclusion chromatograph utilized in this work requires sample derivatization (hydroxyl acetylation) to enable mobile phase (tetrahydrofuran) solubility, resulting in derivatized molecules that are of higher mass compared to the originals. Secondly, the instrument is calibrated to polystyrene standards. While polystyrene does bear aromatic moieties akin to lignin, it is otherwise decidedly non-lignin when factoring in its linearity and homogeneity. Therefore usage of polystyrene as a

calibration standard can be appropriately challenged. These two reasons cause us to address the molecular mass values measured as estimates, and the resulting molecular mass values have been rounded to the hundreds place to reflect estimation.

The approximate molecular mass of biorefinery lignin preparations is also included in Table 9. The low molecular mass estimated for XADL (Mw_n: 400, Mw_w: 800 g/mol) agrees with previous discussion around its extent of degradation. In combination with phenolic hydroxyl quantities, it is clear that XADL (98% representative) is a strongly degraded lignin fraction that mostly exists in the form of dimers, trimers, and minor oligomers. Of great significance, monomeric lignin fragments were found to only comprise ~1-2% of autohydrolyzate-soluble lignin in our previous work.[32] For AHCELp, its molecular mass (Mw_n: 1600, Mw_w: 8700) is estimated to be much higher than XADL. As discussed, the estimated mass of AHCELp does not sufficiently embody the original lignin's molecular mass (~30% representative). However, this data point does provide a "high water mark" for residual lignin's molecular mass, meaning that the un-extractable lignin (~70%) can be considered to have Mw_w of $\geq$ 8700 g/mol. This also indicates that the un-extracted lignin is largely in-tact, or it contains notable amounts of condensed lignin structures (molecular mass magnifiers). As before, it is most likely a combination of both aforementioned possibilities.

3.9. Native substructure quantification in biorefinery lignin preparations using 2D-HSQC and ^{13}C NMR

Figure 5 displays the substructure-labeled 2D-HSQC spectra (with ^{13}C spectra embedded) of both biorefinery lignin preparations, and Table 9 contains substructure

quantification. Structural identification and quantification was performed in an identical manner to protolignin substructure quantification. Greater differences can be visualized when comparing spectra from both biorefinery lignin preparations. First, several signals attributable to carbohydrates are only visible in XADL 2D-HSQC spectra.[22] These signals heavy presence in XADL and absence from AHCELp is best explained by the carbohydrate content of each respective biorefinery lignin preparation. Next, most substructure signals from XADL are at a slightly difference chemical shift than those in AHCELp. The low molecular mass and relatively low polydispersity of XADL (compared to AHCELp) best explains the observable chemical shift differences. LCC substructures are distinctively resonating on XADL 2D-HSQC spectra, however, they were found at lower contours than what is displayed in Figure 5. Additional variation between XADL and AHCELp is best explained by quantification of substructure abundance and non-condensed S/G ratios, Table 9.

The non-condensed S/G ratios of XADL and AHCELp were found to be 6.8 and 3.4 (respectively), indicating that the high quantities of C3 & C5 bound phenolic hydroxyls in both preparations are mostly associated with S aromatic structures. One can also visually interpret this result, based on the strong S aromatic signal resonance versus puny G resonance. The quantified non-condensed S/G values indicate preferential reactivity of S structures during autohydrolysis or loss of G structures through inter-lignin condensation reactions. Both of the aforementioned possibilities induce higher S/G. The higher non-condensed S/G of XADL versus AHCELp does indicate that S structures are subject to the greatest amount of autohydrolysis-induced degradation when considering previous data

establishing XADL to be the more degraded biorefinery lignin fraction (in terms of molecular mass & phenolic functionality). Our suggested preferential S reactivity is in conflict with a conclusion drawn by Matsumoto et al[25], which suggested that S lignin is more resistant to acidolysis than G lignin based on model compound studies. Importantly, said work does draw a caveat that many other factors control lignin lability during acidolysis. Concerning other lability factors, it seems native uronosil ester LCC substructures are associated with S lignin units (Table 4). The strong presence of uronosil esters in XADL (8/100 Ar) could potentially explain XADL's high S/G ratio (6.8).

LCC substructures are clearly contributors to XADL solubility in autohydrolyzate (~11 total LCC/100 Ar). Support for this claim can be found in the significantly greater amount of total β-O-4' substructures in XADL (43/100 Ar) versus AHCELp (24/100 Ar). Pronounced β-O-4' abundancy in XADL suggests β-O-4' degradation resulting in lowered lignin molecular mass is not the only reason for lignin solubility.[33-34] Instead, it appears that degraded lignin's chemical connectivity to polar carbohydrates also contributes, resulting in a soluble lignin portion that is largely non-monomeric[32] yet sufficiently hydrophilic. This conclusions was also drawn in our previous work characterizing autohydrolyzate-soluble lignin.[10] For AHCELp, lowered β-O-4' abundance does indicate that acidolysis has occurred, in agreement with XADL's high phenolic functionality and lowered molecular mass. Apart from LCC and β-O-4' contents, it was found that spirodienone structures do not survive autohydrolysis due to absence of substructural resonance in both biorefinery lignin preparations. No other significant difference between XADL and AHCELp substructure quantities was identifiable.

3.10 Biorefinery mass balance of lignin characteristics

A summative dataset regarding the flow of different lignin properties around a biorefinery's mass balance is shown in Table 10, with all units defined per 100g of o.d. raw biomass. The intent behind compilation of such a table was to track the flow of lignin characteristics described in this work throughout the investigated biorefinery process. Values were formulated using preparation yield and yield's relation to lignin quantities per 100g of o.d. raw biomass. The unit weight of sweetgum lignin was assumed to be 200g/C_9 unit, to enable conversion of NMR data (#/100 Ar) to units per 100g o.d. raw biomass.

Total phenolic hydroxyl quantities obtained from the biorefinery lignin preparations were found to be ~40 mmol/100g o.d. raw biomass. This is evidence of significant substructure degradation having occurred during autohydrolysis. As discussed before, the amount of phenolic hydroxyls in protolignin preparations (~14 mmol/100g o.d. raw biomass (MWLp + CELp)) are mostly attributable to planetary ball milling. It can be seen from these results that autohydrolysis induces significantly more chemical degradation in lignin than planetary ball milling. As previously bemoaned, it will be necessary to isolate additional lignin fractions from cellulolytic hydrolysis residue to understand total phenolic functionality of cellulolytic hydrolysis-insoluble residue. However, the possible isolation steps that may render additional lignin extractability will inherently produce phenolic hydroxyl groups not generated exclusively by the biorefinery process.

Before discussion, it is important to note that values present (per 100g o.d. raw biomass) are mostly a function of preparation yield. This must be considered before claiming definitive conclusions about the entirety of protolignin or biorefinery insoluble

lignin. Absolute quantities of non-condensed S and G units were derived from each lignin preparation's S/G ratio. The protolignin preparations were found to contain 50 mmol non-condensed S units (MWLc + CELp), translating to ~65-75% non-condensed S structures. The sum of non-condensed S structures quantified in biorefinery lignin preparations was found to be 16 (XADL) and 25 (AHCELp) mmol/100g o.d. raw biomass. These summate to 41 mmol/100g o.d. raw biomass, or ~82% recovery of quantified non-condensed S structures. This suggests only minor acidolysis of S units, however this cannot be definitively claimed based upon the limited yields of all lignin preparations (with the exception of XADL). Differently, the sum of G structures quantified in biorefinery lignin preparations was only 9 mmol/100g o.d. raw biomass, or ~50% of the amount quantified in protolignin preparations (18 mmol/100g o.d. raw biomass (MWLc + CELp). The notable loss of G abundance in biorefinery lignin preparations indicates loss of these structures through inter-lignin condensation is indeed occurring. This is likely contributing to the restriction of extractable lignin from cellulolytic hydrolysis residue. To understand the exact extent of inter-lignin condensation, it will be necessary to quantify the absolute total amount of non-condensed lignin structures in raw sweetgum, autohydrolyzate-soluble lignin, and cellulolytic residue (not using lignin preparations). Our previous work[7] performed nitrobenzene oxidation (destructive quantification of non-condensed lignin structures) upon wood and cellulolytic hydrolysis residue, and the results also suggested inter-lignin condensation. Future work must seek to incorporate quantitation of non-condensed lignin structures in autohydrolyzate-soluble lignin (XADL went previously un-isolated) into a mass balance similar to what is presented in Table 10. From this, it will be possible to precisely enumerate the extent of

inter-lignin condensation around G (major) and S (minor) aromatic structures that is induced by autohydrolysis.

β-O-4' substructures are lost during autohydrolysis, however, not as significantly as predicted. The amounts measured in MWLc + CELp come to 30 β-O-4'/100 g o.d. raw biomass, but only represent ~54% of protolignin. To our surprise, the amount of β-O-4' recovered in XADL (8 β-O-4'/100g o.d. raw biomass) is nearly a third of that. The same amount can be found in AHCELp, summating to 16 β-O-4'/100g o.d. raw biomass recovered in biorefinery lignin preparations. This indicates roughly half of the original β-O-4' substructures quantified could be recovered after autohydrolysis. It is likely that additional unreacted β-O-4' exists in the un-extractable part of protolignin and cellulolytic hydrolysis residue. However, un-extractable β-O-4' substructures in cellulolytic hydrolysis residue may be condensed to other lignin aromatic rings at the Cα position, as inter-lignin condensation around reactive benzylic carbocations does not result in loss of β-O-4' substructures. The only structural loss during inter-lignin condensation is aliphatic hydroxyl groups, replaced by carbon-carbon bonds. Finally, the amount of LCC structures in protolignin preparations was not very high, ~3/100g o.d. raw biomass. Despite this amount being minimal, the majority of what was quantified in protolignin was recovered in XADL (~2/100g o.d. raw biomass). As previously mentioned, chemically-attached carbohydrates are likely solubility-promoters for the autohydrolyzate-soluble portion of sweetgum lignin.

4. Conclusions

Protolignin preparations were prepared from (hardwood) sweetgum with total protolignin representativeness of ~54%. From these preparations, the quantities of phenolic functionalities and native substructures was determined. Purification of the MWL preparation resulted in quantifiable loss of LCC substructures. The lost substructures were located in the original crude MWL, mostly as uronosil esters associated with syringyl lignin structures. Subjection of sweetgum to biorefinery processing rendered ~15% of its lignin soluble in autohydrolyzate, with the remainder insoluble throughout the process. A highly representative autohydrolyzate preparation was procured (98% representativeness), which was found to contain high amounts of directly and indirectly chemically-bound carbohydrates, as well as native substructures. About a third of the insoluble lignin could be isolated as a CEL preparation, and was found to contain lower β-O-4' character than autohydrolyzate-soluble lignin. After incorporation of lignin preparation characterization into the investigated biorefinery process' mass balance, it was found that ~50% of the originally quantified β-O-4' substructures were recovered in biorefinery lignin preparations. The sum of this effort demonstrates the properties to expect from hardwood biorefinery lignin when developing valorization methods for either autohydrolyzate-soluble lignin or cellulolytic hydrolysis residual lignin.

References

1. Gonzalez, R.; Treasure, T.; Phillips, R.; Jameel, H.; Saloni, D. Economics of cellulosic ethanol production: Green liquor pretreatment for softwood and hardwood, greenfield and repurpose scenarios. *BioResources* **2011,** *6* (3), 2551-2567.

2. Phillips, R. B.; Jameel, H.; Chang, H.-m. Integration of pulp and paper technology with bioethanol production. *Biotechnology for Biofuels* **2013,** *6* (13).

3. Treasure, T.; Gonzalez, R.; Jameel, H.; Phillips, R. B.; Park, S.; Kelley, S. Integrated conversion, financial, and risk modeling of cellulosic ethanol from woody and non-woody biomass via dilute acid pre-treatment. *Biofuels, Bioproducts, & Biorefining* **2014,** *8*, 755-769.

4. Narron, R. H.; Kim, H.; Chang, H.-m.; Jameel, H.; Park, S. Biomass pretreatments capable of enabling lignin valorization in a biorefinery process. *Current Opinion in Biotechnology* **2016,** *38*, 39-46.

5. Ragauskas, A. J.; Beckham, G. T.; Biddy, M. J.; Chandra, R.; Chen, F.; Davis, M. F.; Davison, B. H.; Dixon, R. A.; Gilna, P.; Keller, M.; Langan, P.; Naskar, A. K.; Saddler, J. N.; Tschaplinski, T. J.; Tuskan, G. A.; Wyman, C. E. Lignin valorization: improving lignin processing in the biorefinery. *Science* **2014,** *344* (6185), 709-720.

6. Cline, S. P.; Smith, P. M. Opportunities for lignin valorization: an exploratory proces. *Energy, Sustainability and Society* **2017,** *7* (26), 1-12.

7. Narron, R. H.; Han, Q.; Park, S.; Chang, H.-m.; Jameel, H. Lignocentric analysis of a carbohydrate-producing lignocellulosic biorefinery process. *Bioresource Technology* **2017,** *241*, 857-867.

8. Han, Q.; Jin, Y.; Jameel, H.; Chang, H.-m.; Phillips, R.; Park, S. Autohydrolysis pretreatment of waste wheat straw for cellulosic ethanol production in a co-located straw pulp mill. *Applied Biochemistry and Biotechnology* **2015,** *175* (2), 1193-1210.

9. Huang, C.; Jeuck, B.; Du, J.; Yong, Q.; Chang, H.-m.; Jameel, H.; Phillips, R. B. Novel process for the coproduction of xylo-olgiosaccharides, fermentable sugars, and lignosulfonates from hardwood. *Bioresource Technology* **2016,** *219*, 600-607.

10. Narron, R. H.; Chang, H.-m.; Jameel, H.; Park, S. Soluble lignin recovered from biorefinery pretreatment hydrolyzate characterized by lignin-carbohydrate complexes. 2017.

11. Sluiter, J. B.; Ruiz, R. O.; Scarlata, C. J.; Sluiter, A. D.; Templeton, D. W. Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. *Journal of Agricultural and Food Chemistry* **2010,** *58* (16), 9043-9045.

12. Li, S.; Ogunkoya, D.; Fang, T.; Willoughby, J.; Rojas, O. J. Carboxymethylated lignins with low surface tension toward low viscosity and highly stable emulsions of crude bitumen and refined oils. *Journal of Colloid and Interface Science* **2016,** *482*, 27-38.

13. Cui, C.; Sun, R.; Argyropoulos, D. S. Fractional precipitation of softwood Kraft lignin: isolation of narrow fractions common to a variety of lignins. *ACS Sustainable Chemistry & Engineerng* **2014,** *2* (4), 959-968.

14. Balakshin, M.; Capanema, E. On the quantification of lignin hydroxyl groups with ^{31}P and ^{13}C NMR spectroscopy. *Journal of Wood Chemistry and Technology* **2015,** *35* (3), 220-237.

15. Chang, H.-m.; Cowling, E. B.; Brown, W.; Adler, E.; Miksche, G. Comparative studies on cellulolytic enzyme lignin and milled wood lignin of sweetgum and spruce. *Holzforschung* **1975,** *29*, 153-159.

16. Capanema, E.; Balakshin, M.; Katahira, R.; Chang, H.-m.; Jameel, H. How well do MWL and CEL preparations represent the whole hardwood lignin? *Journal of Wood Chemistry and Technology* **2014,** *35* (1), 17-26.

17. Zhang, L.; Gellerstedt, G. Quantitative 2D-HSQC NMR determination of polymer structures by selecting suitable internal standard references. *Magnetic Resonance in Chemistry* **2006,** *45* (1), 37-45.

18. Balakshin, M.; Capanema, E.; Gracz, H.; Chang, H.-m.; Jameel, H. Quantification of lignin-carbohydrate linkages with high-resolution NMR spectroscopy. *Planta* **2011,** *233* (6), 1097-1110.

19. Yelle, D. J.; Ralph, J.; Frihart, C. R. Characterization of nonderivatized plant cell walls using high-resolution solution-state NMR spectroscopy. *Magnetic Resonance in Chemistry* **2008,** *46* (6), 508-517.

20. Zeng, J.; Helms, G. L.; Gao, X.; Chen, S. Quantification of wheat straw lignin structure by comprehensive NMR analysis. *Journal of Agricultural and Food Chemistry* **2013,** *61* (46), 10848-10857.

21. Zhang, L.; Gellerstedt, G.; Lu, F.; Ralph, J. NMR studies on the occurence of spirodienone structures in lignins. *Journal of Wood Chemistry and Technology* **2006,** *26*, 65-79.

22. Kim, H.; Ralph, J.; Akiyama, T. Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d_6. *BioEnergy Research* **2008,** *1*, 56-66.

23. Trajano, H. L.; Engle, N. L.; Foston, M.; Ragauskas, A. J.; Tschaplinski, T. J.; Wyman, C. E. The fate of lignin during hydrothermal pretreatment. *Biotechnology for Biofuels* **2013,** *6*, 110-126.

24. Nitsos, C. K.; Matis, K. A.; Triantafyllidis, K. S. Optimization of hydrothermal pretreatment of lignocellulosic biomass in the bioethanol production process. *ChemSusChem* **2013,** *6*, 110-122.

25. Matsumoto, Y.; Yokoyama, T.; Akiyama, T.; Shioya, T. Formation rate of benzyl cation intermediate from p-Hydroxyphenyl, guaiacyl, or syringyl nucleus in acidolysis of lignin. *Journal of Wood Chemistry and Technology* **2017,** *37* (2), 75-86.

26. Yokoyama, T. Revisiting the mechanism of β-O-4 bond cleavage during acidolysis of lignin. Part 6: a review. *Journal of Wood Chemistry and Technology* **2014,** *35* (1), 165-174.

27. Ko, J. K.; Ximenes, E.; Kim, Y.; Ladisch, M. R. Adsorption of enzyme onto lignins of liquid hot water pretreated hardwoods. *Biotechnology and Bioengineering* **2015,** *112* (3), 447-456.

28. Ko, J. K.; Ximenes, E.; Kim, Y.; Ladisch, M. R. Effect of liquid hot water pretreatment severity on properties of hardwood lignin and enzymatic hydrolysis of cellulose. *Biotechnology and Bioengineering* **2015,** *112* (2), 252-262.

29. Chi, C.; Zhang, Z.; Chang, H.-m.; Jameel, H. Determination of furfural and hydroxymethylfurfural formed from biomass under acidic conditions. *Journal of Wood Chemistry and Technology* **2009,** *29* (4), 265-276.

30. Guerra, A.; Filpponen, I.; Lucia, L. A.; Argyropoulos, D. S. Comparative evaluation of three lignin isolation protocols for various wood species. *Journal of Agricultural and Food Chemistry* **2006,** *54*, 9696-9705.

31. Zoia, L.; Orlandi, M.; Argyropoulos, D. S. Microwave-assisted lignin isolation using the enzymatic mild acidolysis (EMAL) protocol. *Journal of Agricultural and Food Chemistry* **2008,** *56*, 10115-10122.

32. Boes, K. S.; Narron, R. H.; Chen, Y.; Park, S.; Vinueza, N. R. Characterization of biofuel refinery byproduct via selective electrospray ionization tandem mass spectrometry. *Fuel* **2017,** *188*, 190-196.

33. Leschinsky, M.; Zuckerstatter, G.; Weber, H. K.; Patt, R.; Sixta, H. Effect of autohydrolysis of *Eucalyptus globulus* wood on lignin structure. Part 1: Comparison of different lignin fractions formed during water prehydrolysis. *Holzforschung* **2008,** *62*, 645-652.

34. Leschinsky, M.; Zuckerstatter, G.; Weber, H. K.; Patt, R.; Sixta, H. Effect of autohydrolysis of *Eucalyptus globulus* wood on lignin structure. Part 2: Influence of autohydrolysis intensity. *Holzforschung* **2008,** *62,* 653-658.

TABLES

Table 1. Chemical composition and isolation yields of raw sweetgum and protolignin preparations (g/100g o.d. solid)

Protolignin / Chemical component	Raw woodmeal	Milled wood lignin ("MWL")		Cellulolytic enzyme lignin ("CEL")	
		Crude (MWLc)	Purified (MWLp)	Crude (CELc)	Purified (CELp)
Glucan	40.2	2.7	2.0	4.3	2.5
Xylan	15.7	7.2	1.2	4.0	1.3
Galactan	0.8	2.0	0.8	3.6	1.1
Arabinan + Mannan[a]	2.2	4.3	1.9	3.2	1.6
Total carbohydrate	58.9	16.2	5.9	15.1	6.5
Acid-insoluble lignin	21.5	67.7	86.6	70.1	90.1
Acid-soluble lignin	3.7	0.7	0.3	3.2	0.3
Total lignin	25.2	68.4	86.9	73.3	90.4
Extractives[b]	1.2	-	-	-	-
Ash	0.6	*nd*[e]	*nd*	0.2	*nd*
Uncharacterized	14.1	15.4	7.2	11.4	3.1
Gravimetric solid yield[c]	-	55.2 ± 2.5	27.0 ± 0.6	25.4 ± 2.2	18.3 ± 0.3
Protolignin preparation percent yield[d]	-	38%	23%	19%	16%

[a] Co-eluting monosaccharides
[b] Measured gravimetrically and removed prior to further experimentation
[c] g isolate / 100g total raw lignin
[d] wt% of total raw lignin, based on gravimetric yield and corrected for carbohydrate content
[e] Not detected

Table 2. Sweetgum protolignin preparation hydroxyl group abundancies

Protolignin preparation / Hydroxyl groups[a]	MWLp	CELp
Aliphatic OH	7.2 ± 0.3, 5.9[b]	7.5 ± 0.2, 6.5[b]
C3 & C5 bound phenolic OH	0.9 ± 0.1	0.5 ± 0.0
C3 or C5 unbound phenolic OH	0.7 ± 0.0	0.5 ± 0.0
C3 & C5 unbound phenolic OH	0.2 ± 0.0	0.1 ± 0.0
Total phenolic OH	1.8	1.1

[a] mmol/g total lignin in solid
[b] mmol/g solid

Table 3. 2D-HSQC signal assignments for lignin substructures and aromatic rings

Label	δ_C/δ_H	Assignment
C_β	53.1/3.49	C_β-H_β of phenylcoumaran substructures (C)
B_β	53.7/3.05	C_β-H_β of resinol substructures (B)
-OCH_3	55.9/3.73	C-H of methoxyls
A_γ	59.6-60.8/3.37-3.72	C_γ-H_γ of β-O-4 substructures(A)
D_β	59.8/2.78	C_β-H_β of spirodienone substructures (D)
$Acyl_\gamma$	61.6/4.08	C_γ-H_γ of Cγ acylated side chains
UrE_γ	63.1/4.28	C_γ-H_γ of uronosil ester LCC (UrE)
B_γ	71.3/4.18	C_γ-H_γ of resinol substructures (B)
A_a	72.6/5.00	C_a-H_a of β-O-4 substructures (A)
$BE_{\alpha\text{-}1°}$	81.1/4.54	C_α-H_α of benzyl ether (primary alcohol) LCC (BE1°)
$BE_{\alpha\text{-}2°}$	81.1/5.13	C_α-H_α of benzyl ether (secondary alcohol) LCC (BE2°)
D_α	81.1/5.13	C_α-H_α of spirodienone substructures (D)
$A_{\beta\ (G/H)}$	82.5/4.27	C_β-H_β of β-O-4 substructures linked to G unit (A)
A'_a	83.1/5.25	C_a-H_a of β-O-4 substructures with α carbonyl (A')
B_a	84.9/4.69	C_a-H_a of resinol substructures (B)
$A_{\beta\ (S)}$	86.0/4.11	C_β-H_β of β-O-4 substructures linked to S unit (A)
C_a	86.8/5.49	C_a-H_a of phenylcoumaran substructures (C)
$GlcA_1$	97.2/5.18	C_1 of non-esterified 4-O-methylglucuruonic acid
$EGlcA_1$	98.9/4.70	C_1 of esterified 4-O-methylglucuruonic acid

Table 3 (continued)

PhG_1	100.0-103.0/4.89-5.20	C_1 of phenyl glycoside LCC (PhG)
$S_{2,6}$	104.1/6.74	$C_{2,6}$-$H_{2,6}$ in etherified syringyl units (S)
$S'_{2,6}$	106.6/7.32	$C_{2,6}$-$H_{2,6}$ in syringyl units with C_a=O groups (S')
G_2	110.3/6.91	C_2-H_2 in guaiacyl units (G)
G_5	114.4/6.72	C_5-H_5 in guaiacyl units (G)
G'_6	118.7/7.33	C_6-H_6 in guaiacyl units with C_a=O groups (G')
G_6	119.0/6.82	C_6-H_6 in guaiacyl units (G)
$H_{2,6}$	127.8/7.22	$C_{2,6}$-$H_{2,6}$ in *p*-hydroxyphenyl units (H)

Table 4. Native substructure quantities for sweetgum protolignin preparations (#/100 Ar)

Substructure		**MWLc**	**MWLp**	**CELp**
Non-condensed S/G		2.7	2.0	2.7
β-O-4’ (A)		40	39	47
β-O-4’ with Cα carbonyl (A’)		1	2	1
Total β-O-4’ (A + A’)		41	41	48
β-β’ (B)		4	5	4
β-5’ (C)		2	3	3
Spirodienone (D)		*nq*[b]	2	2
Cγ acyl ($Acyl_\gamma$)		1	1	1
LCC	Uronosil ester (UrE)	5	1	1
	Benzyl ether, 1° OH (BE1°)	1	*nd*[c]	*nd*
	Benzyl ether, 2° OH[a] (BE2°)	*nq*	*nd*	*nd*
	Phenyl glycoside (PhG)	*nq*	*nq*	*nq*
	Total LCC	~6	~1	~1

[a] Semi-quantitative
[b] Not quantified
[c] Not detected

Table 5. Raw and autohydrolyzed sweetgum chemical composition, autohydrolyzate carbohydrate and lignin concentrations (g/100g o.d. raw biomass)

Chemical component	Raw	Autohydrolyzed	
		Solid	Liquid
Glucan	40.2	37.5	2.3
Xylan	15.7	3.7	10.4
Galactan	0.8	*nd*	0.5
Arabinan + Mannan[a]	2.2	1.6	0.5
Total carbohydrate	58.9	42.8	13.7
Acid-insoluble lignin	21.5	20.3	-
Acid-soluble lignin	3.7	1.1	-
Total lignin	25.2	21.4	3.8[c], 15.1%[d]
Extractives[b]	1.2	-	-
Ash	0.6	0.1	-
Uncharacterized	14.1	3.3	-

[a] Co-eluting monosaccharides
[b] Removed prior to autohydrolysis
[c] Back-calculated
[d] wt% of total raw lignin

Table 6. Autohydrolyzate-soluble lignin (XADL) adsorption yield

Soluble lignin abundance	3.8 g/100g o.d. raw biomass
Reduced autohydrolyzate UV response	43 ± 0 AU at 280 nm
Reduced post-adsorption autohydrolyzate UV response	1 ± 0 AU at 280 nm
Soluble lignin adsorption yield	98% of AU, 3.7 g/100g o.d. raw biomass

Table 7. Chemical composition of XADL and AHCEL preparations (g/100g o.d. solid), AHCEL preparation yields

Biorefinery lignin preparation / Chemical component	XADL	AHCELc	AHCELp
Glucan	3.5	3.1	2.5
Xylan	47.2	2.3	1.3
Galactan	0.7	3.5	1.1
Arabinan + Mannan[a]	5.7	2.2	1.6
Total carbohydrate	57.1	11.1	6.5
Acid-insoluble lignin	7.1	80.8	90.1
Acid-soluble lignin	21.8	1.9	0.3
Total lignin	28.9	82.7	90.4
Ash	*nd*	0.2	*nd*
Uncharacterized	14.0	6.0	3.1
Gravimetric solid percent yield[b]	-	41.0 ± 0.5	32.6 ± 0.1
Preparation percent yield	98%[c]	34%[d]	30%[d]

[a] Co-eluting monosaccharides

[b] g isolate/100g total lignin in autohydrolyzed solids

[c] From Table 6, calculable from loss in autohydrolyzate-soluble lignin UV response after adsorption

[d] wt% of total lignin in autohydrolyzed solids, based on gravimetric yield and corrected for carbohydrate content

Table 8. Sweetgum biorefinery lignin preparations hydroxyl group quantities and approximate molecular masses

Biorefinery lignin preparation / Hydroxyl groups[a]	XADL			AHCELp		
Aliphatic OH	28.5 ± 1.0, 8.3[b]			4.1 ± 0.3, 3.8[b]		
C3 & C5 bound phenolic OH	3.7 ± 0.3			2.2 ± 0.1		
C3 or C5 unbound phenolic OH	1.2 ± 0.0			0.7 ± 0.0		
C3 & C5 unbound phenolic OH	0.4 ± 0.0			0.1 ± 0.0		
Total phenolic OH	5.3			3.0		
Molecular mass (g/mol)	Mw_n[c]	Mw_w[d]	PDI[e]	Mw_n	Mw_w	PDI
	400	800	2.0	1600	8700	5.4

[a] mmol/g total lignin in solid
[b] mmol/g solid
[c] Number averaged molecular mass
[d] Weight averaged molecular mass
[e] Polydispersity index (Mw_w/Mw_n)

Table 9. Native substructure quantities for sweetgum biorefinery lignin preparations (#/100 Ar)

Substructure		XADL	AHCELp
Non-condensed S/G		6.8	3.4
β-O-4' (A)		42	23
β-O-4' with Cα carbonyl (A')		1	1
Total β-O-4' (A + A')		43	24
β-β' (B)		4	3
β-5' (C)		2	3
Spirodienone (D)		*nd*[b]	*nd*
Cγ acyl ($Acyl_\gamma$)		1	*nd*
LCC	Uronosil ester (UrE)	8	*nq*
	Benzyl ether, 1° OH (BE1°)	2	1
	Benzyl ether, 2° OH[a] (BE2°)	*nq*[c]	*nd*
	Phenyl glycoside (PhG)	1	*nd*
	Total LCC	~11	~1

[a] Semi-quantitative
[b] Not detected
[c] Not quantified

Table 10. Biorefinery mass balance of sweetgum lignin properties (X / 100g o.d. raw biomass)

Lignin property	Raw				Biorefinery			
	Wood	MWLc	MWLp	CELp	Autohydrolyzate	XADL	Cellulolytic hydrolysis residue	AHCELp
Total raw lignin representativeness[a]	100%	-	-	-	15%	-	85%	-
Lignin preparation representativeness[b]	-	38%	23%	16%	-	98%	-	30%
Total lignin, g	25.2	9.6	5.8	4.0	3.8	3.7	21.4	6.4
Phenolic hydroxyl, mmol	-	*nm*[d]	10	4	-	20	-	19
Non-condensed S units, mmol[c]	-	35	19	15	-	16	-	25
Non-condensed G units, mmol[c]	-	13	10	5	-	2	-	7
Total β-O-4', #	-	20	12	10	-	8	-	8
Total LCC, #	-	3	<1	<1		2	-	<1

[a] wt% of total raw lignin
[b] wt% of lignin in starting stream
[c] Derived from preparation S/G
[d] Not measured

FIGURES

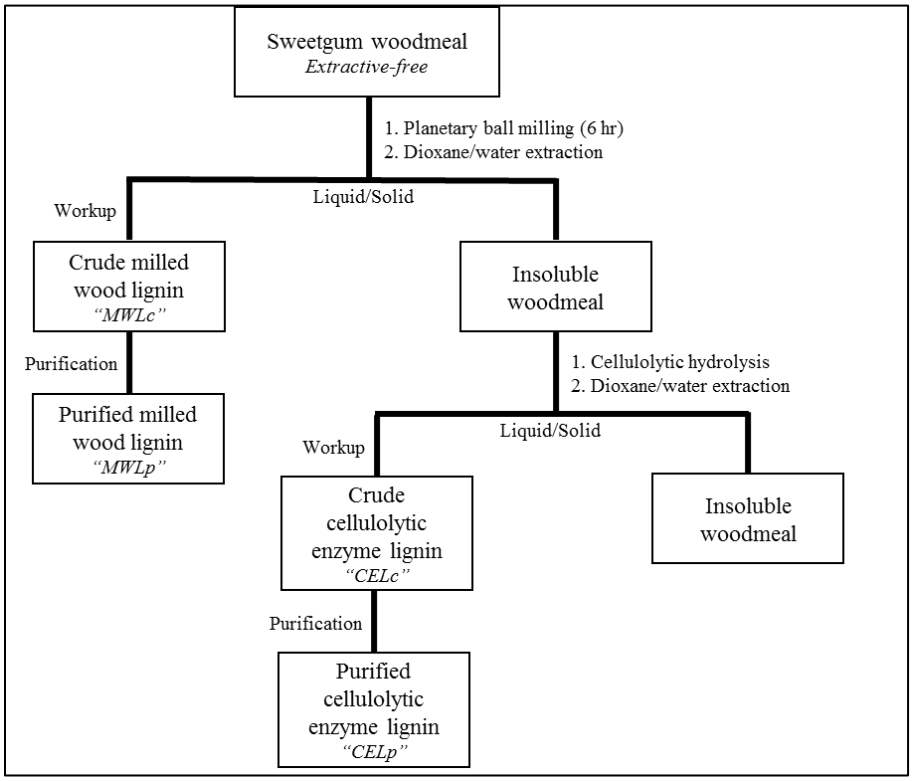

Figure 1. Protolignin preparation procurement protocol

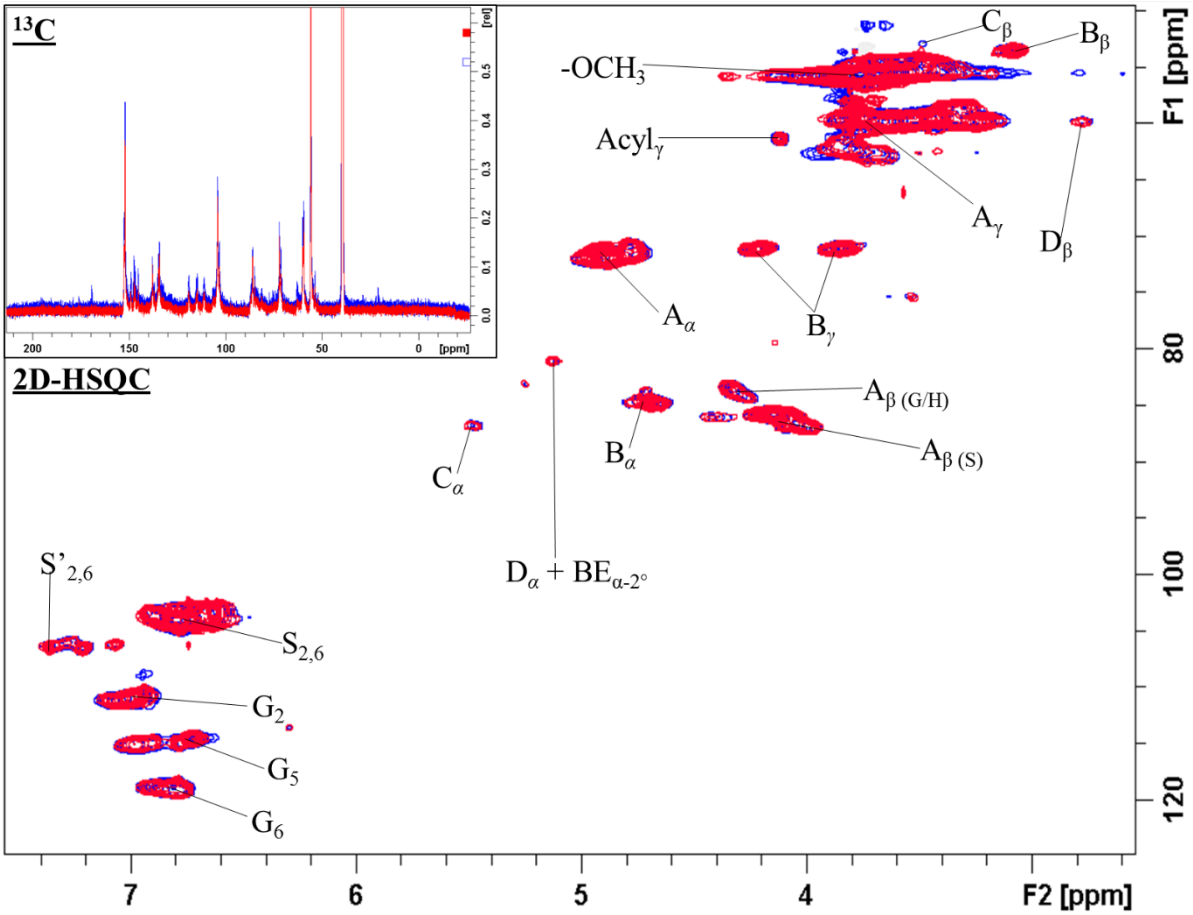

Figure 2. Substructure elucidation from 2D-HSQC and ^{13}C spectra
Red: MWLp
Blue: CELp

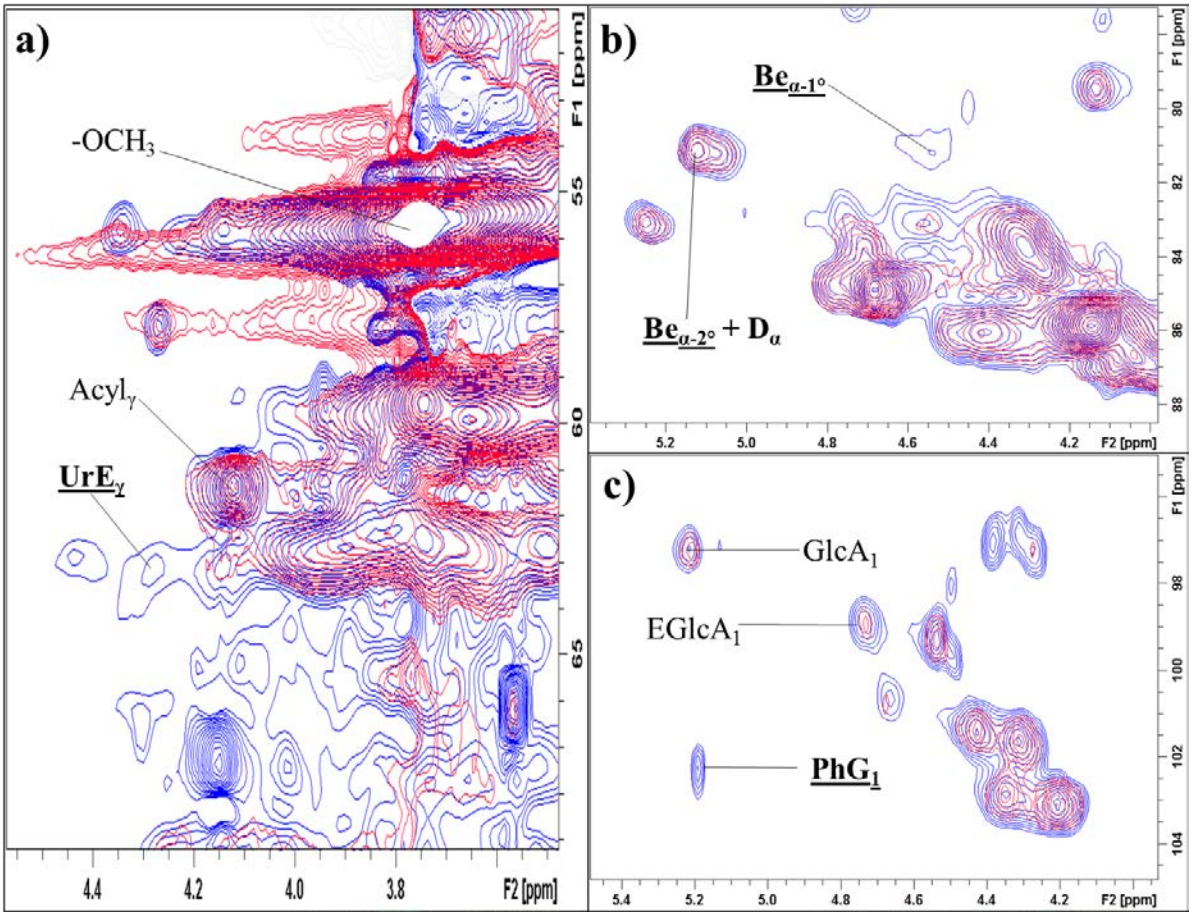

Figure 3. LCC signals from 2D-HSQC spectra of MWLc and MWLp
Blue: MWLc
Red: MWLp
a) Region containing uronosil ester LCC signals
b) Region containing benzyl ether LCC signals
c) Region containing phenyl glycoside LCC and uronic acid signals

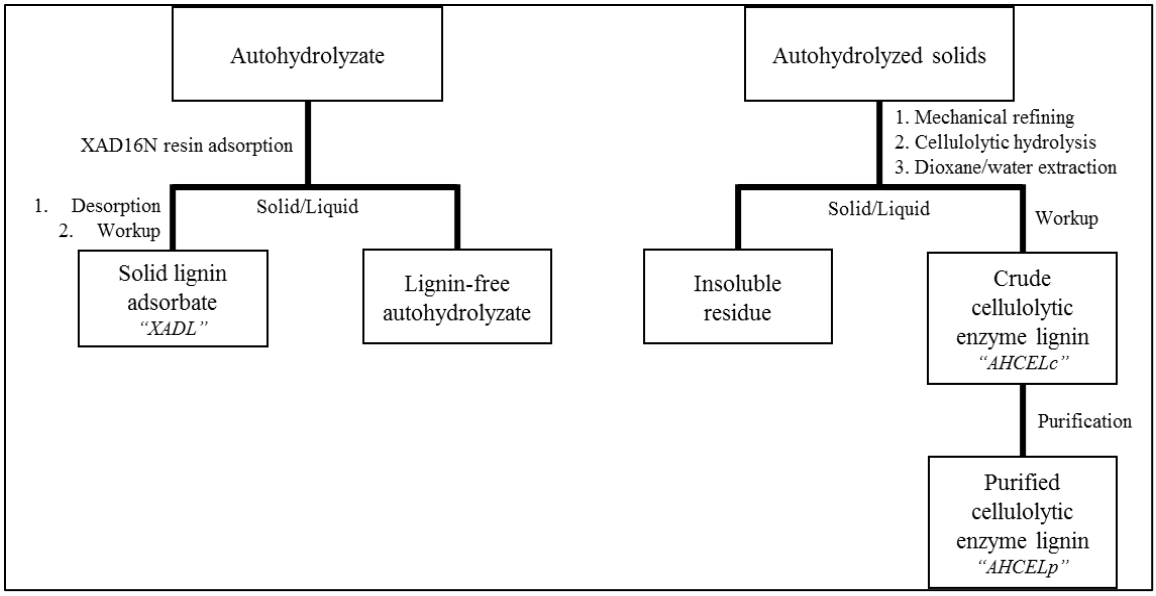

Figure 4. Isolation protocols for autohydrolyzate-soluble and cellulolytic hydrolysis-insoluble lignin

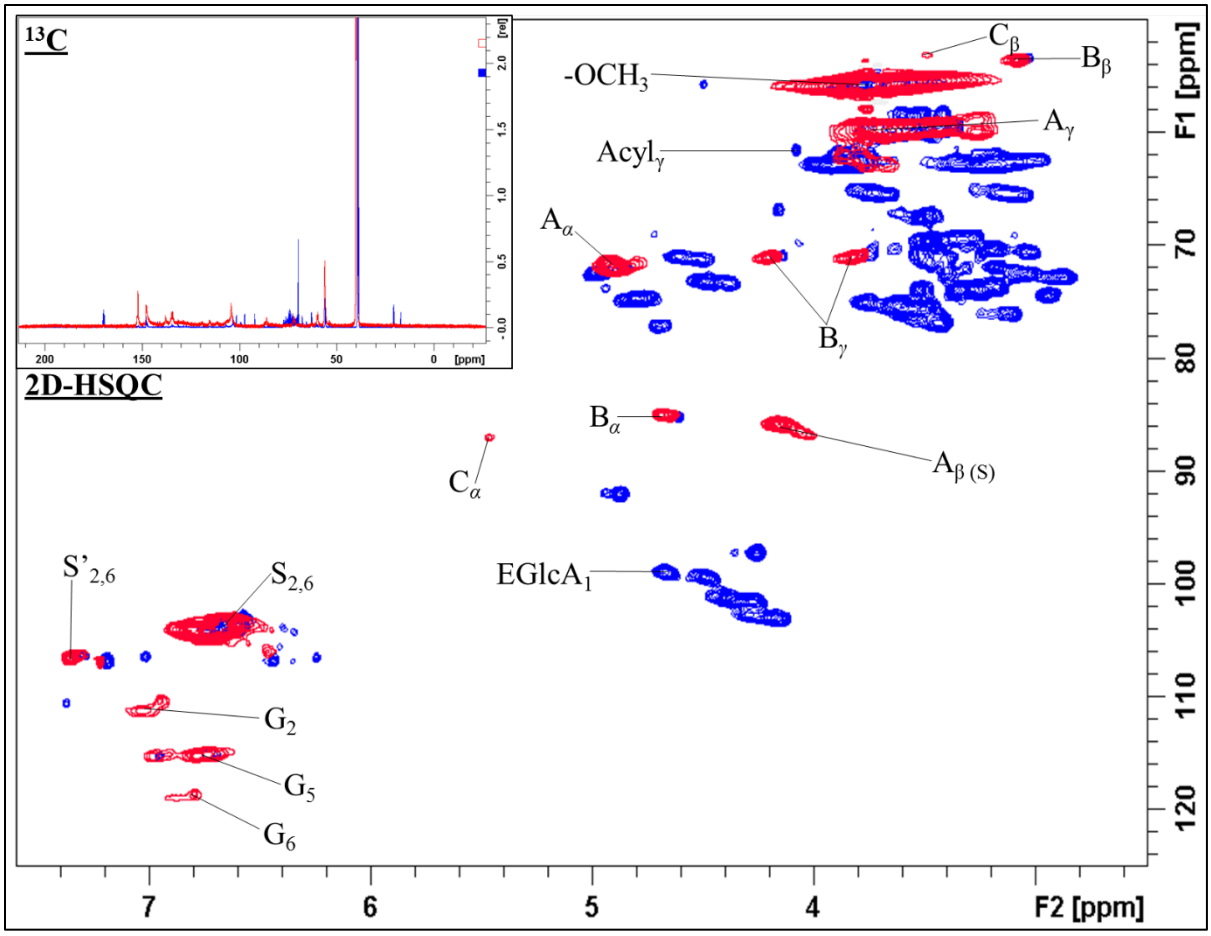

Figure 5. 2D-HSQC and ^{13}C NMR spectra from sweetgum biorefinery lignin preparations

Chapter 5. Biorefinery mass balance of lignin properties across six botanically diverse lignin sources

Abstract

A large database of protolignin and biorefinery lignin properties was compiled through isolation and characterization of lignin preparations. This information was implemented into lignin's mass balance through the investigated biorefinery process, with respect to each individual biomass' response to the operations (autohydrolysis pretreatment, cellulolytic hydrolysis). For hardwoods, clear trends were observable between protolignin S/G and the ensuing properties of biorefinery lignin. Hardwood residual solid lignin was more condensed and more acidolyzed if it contained a greater abundance of G aromatic structures. Similar analysis of non-wood behavior was less elucidative, as no clear trends between individual lignin properties and chemical behavior could be drawn. However, inter-lignin condensation was quantifiable through loss of non-condensed G aromatic structures. In addition, ~60% of non-wood β-O-4' was acidolyzed, resulting in liberation and eventual solubilization of hydroxycinammic acids and lignin molecules containing LCC and DBDO substructures. The sum of this effort provides a better understanding of how protolignin influences the properties of biorefinery lignin, and specifically denotes the properties of the two biorefinery lignin fractions produced by the investigated biorefinery process.

1. Introduction

Industrial implementation of biorefinery processes has stalled, mostly due to a decline in cellulosic ethanol's selling price and a lack of feasible alternative routes for carbohydrate valorization.[1-2] Alternative routes for carbohydrate valorization include either biocatalytic or inorganic catalytic conversion of carbohydrates into products such as succinic acid[3-4], levulinic acid[3, 5], lactic acid[6], as well as others.[7] In contrast, another approach to achieving successful industrial biorefinery implementation is valorization of biorefinery lignin to supplement the limited value output provided by ethanol sales.[8-15] Our recent research efforts[14, 16-19] have focused exclusively on achieving a more detailed understanding of the chemical and molecular properties of biorefinery lignins in effort to assist in biorefinery lignin valorization.

Valorization of biorefinery lignin is impeded by lignin's heterogeneity, which is further amplified through biorefinery processing (relative to protolignin). This includes existence of several lignin-containing process streams (process configuration dependent), with heterogeneous lignin in each respective stream. Specifically, biorefinery lignins are comprised of a distribution of lignin molecules with various molecular weights, functionality, and chemical bonding patterns. Further complicating the matter of heterogeneity is acknowledgement that highly similar feedstocks will likely be unavailable over long periods of time. It is not feasible to assume that an industrial-scale biorefinery, processing tons of biomass per day, will have consistent access to the same feedstocks year-round. For this reason, it will be imperative for biorefineries to be flexible in processing conditions with respect to various starting materials. It is important to understand the effects of feedstock

variation on biorefinery lignin properties, in hopes of eventually developing an approach to tune biorefinery processes around maximizing the output value per unit of original lignocellulosic material.

This manuscript sextuples our previous work[19], which utilized raw lignin (protolignin) preparations and biorefinery lignin preparations from a single biomass source to describe lignin's chemical and molecular properties as a result of the investigated biorefinery process. Now such analysis is expanded to include six biomasses, to address the effects of feedstock variation upon biorefinery lignin's characteristics. The biorefinery process chosen for continued investigation is a biological-based process, which utilizes sequential autohydrolysis and mechanical refining pretreatment prior to cellulolytic enzyme hydrolysis. Upon the previously studied hardwood protolignin, this process was found to render two lignin fractions: one which is solubilized by autohydrolysis (~15 wt% of total raw lignin), and the other remaining insoluble throughout the entire process (~85 wt% of total raw lignin). Autohydrolysis-soluble lignin was described as strongly degraded based upon its solubility, supported by its strong recoverability, low molecular mass, high phenolic functionality, and prevalent lignin-carbohydrate chemical connectivity. It is important to note that only a minor portion of this "strongly degraded" lignin is comprised of monomeric lignin fragments[16], i.e. lignin fragments without unit-linking chemical substructures. The insoluble lignin, which was not fully isolatable (~30% yield from insoluble lignin), bears low quantities of β-O-4' substructures, limited phenolic hydroxyl functionality, as well as a broad range of molecular masses. Implementation of the properties quantified from both protolignin and biorefinery lignin preparations into the biorefinery process' mass balance

enabled evaluation of both functional group changes and recovery of chemical structures. Similar analysis around several potential biorefinery feedstocks will enable adressment of the effect of varying starting biomasses upon biorefinery lignin's properties with respective to each individual protolignin.

Six varying lignocellulosic biomasses (3 hardwoods and 3 non-woods) were subjected to identical protolignin isolation protocols, biorefinery processing, and biorefinery lignin isolation protocols. All protolignin and biorefinery lignin preparations procured were characterized for lignin purity, preparation representativeness, hydroxyl functionality, and native inter-lignin substructures. After result compilation, all the data was framed into the biomass-specific biorefinery mass balances. From this information, trends were able to be drawn between biorefinery lignin properties as a function of different starting materials. The cumulative result of this effort will guide biorefinery development in terms of more-informed lignin valorization evaluation, process feedstock selection, and process tuning around different feedstocks.

2. Materials and methods

2.1. Raw materials

The moisture content of each sample was measured by the weight difference before and after drying in a convective oven at 105 °C until constant mass was achieved. All raw materials were air-dried at room temperature for two weeks to establish constant moisture content. Hardwood chips from maple (*Acer rubrum*) and sweetgum (*Liquidambar styraciflua*) were provided by the Tree Improvement Program at North Carolina State University (Raleigh,

NC, USA). Nitens (*Eucalyptus nitens*) woodchips were graciously supplied by Celulosa Arauco y Constitución (Arauco, Chile). All wood chips were screened according to Scandinavian Pulp, Paper, and Board Testing Committee SCAN-CM 40:01, collecting all chips below (and including) the large accept chip classification for experimentation. The non-woody biomasses: sugarcane bagasse, wheat straw, and switchgrass (*Panicum virgatum*), were each obtained from unique sources. Our supply of sugarcane bagasse was catered by a sugarcane manufacturer located in southeastern Brazil. Wheat straw was purchased at home-improvement retailer Lowe's (Raleigh, NC, USA). Finally, switchgrass was generously gifted for experimentation from a nearby farm in southern Wake County, North Carolina. All nonwoods were hand-cut into 2-3 cm size while retaining all small particles. After cutting, all feedstocks were stored in plastic bags. In addition, part of the feedstocks were ground by a Wiley Mill (Model No. 4, Thomas Scientific, USA), and the meals were screened to particle size between 20-40 mesh. The screened meals were used for compositional analysis and protolignin preparation isolation. Finally, all biomasses (including produced meals) were subject to non-structural extractives removal through soxhlet extraction using 2:1 (v) benzene-ethanol solution. The mass of extractives in raw biomass was gravimetrically recorded, however, extractives were absent from all biomasses prior to any biorefinery processing or characterization.

2.2. Protolignin preparation isolation

A planetary ball mill (Pulverisette 7 premium line, Fritsch, Germany) was used to enable protolignin extraction. 2g of air-dried extractive-free meal was added to each ball mill

capsule in addition to nine ZrO_2 balls. Samples were milled at 600 rpm for 6 hrs, with 30 minutes of instrument cool-down per 15 minutes of milling. After milling, the meal was dried in a vacuum oven at 35 °C over solid P_2O_5 until constant mass. Next, dried meal was subjected to extraction using 96% 1,4-dioxane/4% water solvent (volume %). Solid and extract were separated using vacuum filtration and Gooch crucibles (fine grade). Extraction was considered complete once resultant solvent filtrate reached complete visual transparency. Next, extracted solids were oven-dried at 105 °C until constant mass to remove residual solvent. Once dried, the material was subjected to high-dosage (20 FPU / o.d. g) cellulolytic hydrolysis at 50 °C. An enzyme cocktail of Cellic CTec2 (Novozymes, Denmark) and Cellic Htec2 (1/9 volume of Ctec2) was applied in an acetate buffer system (pH ~5). ~180 rpm of system agitation was utilized. Following 72 hr of cellulolytic hydrolysis, remaining solids were extensively washed with deionized water to remove free protein and buffer. Washing was considered complete once wash water filtrate reached neutrality. Washed solids were then dried in vacuum oven as before. Dried solids were again extracted with the same solvent and methodology as was applied for the first extraction. Both extract solutions were evaporated by rotary evaporator (Rotavapor® R-300, Buchi, USA), re-suspended into deionized water then frozen, and finally freeze-dried (FreeZone 6, Labconco, USA) to produce solids. All crude protolignin extracts were stored in vacuum oven until constant mass prior to purification or characterization. To measure gravimetric yield of crude protolignin preparations, the mass of each solid relative to the original solid's mass was recorded (i.e. weight loss). The entire protocol is graphically displayed in Figure 1.

2.3. Biorefinery processing

Autohydrolysis pretreatment was performed upon extractive-free biomass in a 1L heated Parr reactor (Parr Instrument Company, USA). Pretreatment temperature was set to 180 °C with ~30 minutes of temperature ramp-up. Pretreatment lasted for 40 mins at temperature, followed by rapid thermal quenching in an ice water bath. After sufficient cooling, autohydrolyzed solids were separated from autohydrolyzate using vacuum filtration. Bulk autohydrolyzate obtained was filtered again using Gooch crucibles (fine grade) to ensure removal of suspended solids. Dually-filtered autohydrolyzates were bottled and kept in a refrigerator at 4 °C prior to further handling. Autohydrolyzed solids were washed with deionized water to remove residual autohydrolyzate until wash water filtrate reached neutrality. The mass of washed solids (accounting for moisture) relative to the mass of the original biomass was recorded to obtain gravimetric solid yield.

Washed autohydrolyzed solids were next mechanically refined for 6000 revolutions using a laboratory PFI mill. After refining, the solids were subjected to the same cellulolytic hydrolysis protocol applied to woodmeal. After 72 hrs of incubation with enzymes, the solids were extensively washed with deionized water to remove free protein and buffer solution until wash water filtrate reached neutrality. The residual solids were air-dried and then stored in vacuum oven until constant mass prior to lignin extraction.

2.4. Obtaining biorefinery lignin preparations

Autohydrolyzate-soluble lignin was obtained from autohydrolyzate using Amberlite® XAD16N adsorptive resin (Dow Chemical Company, USA). Autohydrolyzate (free of

suspended solids) was mixed with a known amount of resin to create autohydrolyzate-to-resin ratio of 10:1 (o.d. g resin:mL autohydrolyzate). After 30 minutes of low speed mixing, resin and solution were separated by vacuum filtration. Adsorbate-containing resin was subjected to water washing to remove un-adsorbed autohydrolyzate solutes. Next, methanol was poured over the captured resin until methanol filtrate reached complete visual transparency. Methanol consumption was ~3x the amount of starting autohydrolyzate. Methanol solution with adsorbates was evaporated and freeze-dried in the same manner as protolignin preparations. Solid adsorbates were kept in vacuum oven until constant mass before characterization. To measure extent of lignin adsorption, an aliquot of autohydrolyzate and resin-treated autohydrolyzate were reduced using sodium borohydride to gather a lignin-only UV response. UV absorption at 280 nm was measured (Lambda XLS, PerkinElmer, USA), and the difference in UV response was taken as the extent of lignin adsorption.

Dried cellulolytic hydrolysis-insoluble solids were extracted identically to protolignin preparations. Lignin-containing solvent was then handled in the same manner as previously noted. In addition, the preparation was purified by the same methodology used with both protolignin preparations. A visualization of biorefinery lignin preparation procurement is shown in Figure 2.

2.5. Preparation purification

Crude lignin preparations (with the exception of XADL, which went unpurified) were purified using a dissolution-precipitation method. First, samples were dissolved into 90%

acetic acid/10% water solution. After complete dissolution, lignin-containing solution was precipitated drop-wise into excess deionized water (20x). Precipitates were collected using vacuum filtration, and washed with deionized water until wash water filtrate reached neutrality. Precipitates were allowed to air dry, and next were stored in vacuum oven until constant mass prior to characterization. To measure loss of preparation yield induced by purification, the mass of purified solids was recorded relative to original crude solid mass.

2.6. Compositional analysis

The chemical composition of each meal, autohydrolyzed solid, and all lignin preparations was performed using the standard protocol provided by the National Renewable Energy Laboratory.[20] ~100 mg of carefully weighed o.d. adsorbate was subjected to 72% (w/w) sulfuric acid digestion for 1 hr, followed by dilution to 3% (w/w) H_2SO_4 and autoclaving for 1.5 hr. Per previous work, lignin from autohydrolyzate was only subject to 1 hr of 3% H_2SO_4 digestion to preserve dehydration-prone carbohydrates.[17] An aliquot of acid digestate was taken to determine acid-soluble lignin quantities spectrophotometrically and monosaccharides by liquid chromatography. For acid-soluble lignin, UV absorbance was measured at 205 nm, accounting for dilution. The molar absorptivity coefficient 110 L g^{-1} cm^{-1} was applied to convert absorbance to concentration. For monosaccharides, an HPLC system (Agilent 1200 series, Agilent, USA) was used. In this setup, a Shodex SP-0810 column is utilized (8x300 mm, Showa Denko, Japan) and maintained at 80 °C. Mobile phase was Milli-Q water, flowing 0.5 mL/min for 60 minutes per sample. A refractive index detector monitored eluents. Before injection, acid digestate samples were neutralized to pH ~6 with $CaCO_3$ and

filtered through 0.2 μm nylon filters. Calibration was performed with commercial standards of glucose, xylose, galactose, mannose, and arabinose. Due to observed co-elution of mannose and arabinose, quantification of either saccharide is reported as the sum of both. All values represent the average of duplicate analyses, and average standard deviation was 3.4%.

2.7. Lignin preparation characterization

To estimate molecular weight of the biorefinery lignin preparations, a published method utilizing size-exclusion chromatography (Shimadzu, Japan) calibrated to polystyrene standards was applied.[21] The system was equipped with HR1 and HR5E Styragel columns (Waters, USA) and dual detectors, which are a SPD-20A UV/Vis and a RID-10A refractive index. The oven maintained 35°C and the wavelength for the UV detector was set to 280 nm. Samples were modified by acetylation prior to injection to ensure total solubility in the mobile phase, tetrahydrofuran. Acetylation involves dissolution of lignin preparation into pyridine, addition of excess acetic anhydride, and 72 hrs of reaction time in darkness. The acetylation slurry was poured over deionized ice and left to settle overnight. Solid precipitate was collected via vacuum filtration and washed excessively with deionized water. After washing, samples were stored in a vacuum oven until constant mass. Prior to injection, the acetylated lignins were dissolved into tetrahydrofuran at an approximate concentration of 0.5 mg acetylated lignin per mL tetrahydrofuran and finally filtered through 0.2 μm nylon filters. The mobile phase flow rate was set to 0.7 mL/min. The calibration curve utilized for estimating molecular weight was based on twelve polystyrene dispersions (PSS ReadyCal-Kit, Polymer Standards Service, Germany).

Quantification of hydroxyl groups utilizing ^{31}P NMR was performed utilizing an established method.[22-23] ~40 mg of lignin sample was massed directly inside of a 5 mm NMR tube. After dissolution and internal standard addition, 100 μL of organophosphitylating agent (2-chloro-4,4,5,5-tetramethyl-1,2,3-dioxaphospholane; 95%, Sigma-Aldrich, USA) was introduced and the sample immediately subjected to observation by ^{31}P NMR. The chemical shifts generated from organophosphitylation were recorded using a 300 MHz NMR spectrometer at 25 °C equipped with a 5mm QNP probe (Avance, Bruker, USA).

All 2D-HSQC and ^{13}C NMR spectra were acquired at 25°C using dimethyl sulfoxide-d6 (99.9%, VWR, USA) as a solvent. The concentration of lignin in each sample was ~20% w/v. 2D-HSQC was performed first, followed by ^{13}C upon the same sample. Chromium (III) acetylacetonate (99.99%, Sigma-Aldrich, USA) was added to the NMR tube (0.01 M) following 2D-HSQC analysis to assist in relaxation of carbon nuclei during ^{13}C observation. 2D-HSQC NMR spectra were acquired on a 500 MHz NMR spectrometer equipped with 5 mm BBI probe with Z-axis gradient (Avance, Bruker, USA). A coupling constant ($^{1}J_{C-H}$) of 147 Hz was applied. Quantitative ^{13}C NMR spectra were acquired on the same 500 MHz spectrometer with a different probe than before: 5 mm BBO probe with Z-axis gradient. An inverse gated proton decoupling sequence was applied.

All structurally-representative signals and their corresponding chemical shifts in 2D-HSQC and ^{13}C spectra were identified using reputable publications.[24-29] The following equation was used for quantification of lignin substructures:

$$\frac{X}{100\ Ar} = \frac{2D_x}{2D_{IS}} * \frac{13C_{IS}}{13C_{163-103}} * 600$$

where X is the substructure being quantified, $2D_x$ is the integral of the 2D-HSQC signal used for structural quantification, IS is the integral cluster or region for 2D-HSQC and ^{13}C (respectively), $^{13}C_{163\text{-}103}$ is the integral of the aromatic region of ^{13}C spectra, and 600 is derived from 600 aromatic carbons per 100 aromatic rings. Computed values are viewed as the quantity/percentage of substructures per lignin unit. Unless otherwise noted in discussion, Cα substructure signals are integrated for quantitation. In addition, non-condensed S/G ratio was measured utilizing the following equation:

$$S/G = \frac{(\frac{2D_{S(2,6)} + 2D_{S'(2,6)}}{2})}{(2D_{G(6)} + 2D_{G'(6)})}$$

where $2D_{S\ or\ G}$ denotes the aromatic signals integrated for evaluation of the ratio between S and G aromatic ring structures. For non-woods, H/G was estimated using the same previously described equation except for implementation of H signals in the denominator. Finally, substructure quantities denoted as not detected (*nd*) failed to distinctively resonate in separated fashion at any spectral contour. For the substructures labelled as not quantified (*nq*), their computed values were less than 0.5/100 Ar. All integrals relating in substructure values of <0.5/100 Ar were rejected based upon instrument insensitivity.

3. Discussion

3.1. Analytical approach

In acknowledgement of the significant differences between non-wood lignin and hardwood lignin, this manuscript is divided into two sections which focus individually upon each trio of hardwoods and non-woods analysed. Discussion will begin with interpretation of inter-trio differences in protolignin extractability, functionality, and inter-lignin substructural profiles and abundance. Next, lignin redistribution due to biorefinery processing is addressed and compared. Lignin preparations from both lignin-containing process streams are then characterized along the same lines as were protolignin preparations. Finally, implementation of all lignin preparation's measured properties into individual biorefinery mass balances is discussed, in attempt to tease out the underlying protolignin properties most influential upon biorefinery lignin.

3.2. Hardwood protolignin

Due to hardwood's relatively lower complexity in comparison with non-woods, the first set of biomasses discussed is hardwoods. Starting on Table 1, it can be seen that the three hardwoods tested (maple, sweetgum, *Eucalyptus nitens* (nitens)) bear highly similar quantities of raw lignin (25.2, 25.2, 24.5 g total lignin/100g oven-dry "o.d." raw biomass). Because of this, it will be important to look beyond lignin content to understand how these lignin sources will respond to biorefinery processing.

The two protolignin preparations sequentially isolated from woodmeal combined to represent 41%, 39%, and 36% (wt% of total raw lignin) of maple, sweetgum, and nitens

protolignin (respectively, Table 2). The slightly lower protolignin extractability for nitens versus maple and sweetgum begins to elucidate nitens as an outlier compared to maple and sweetgum, however, such claims are not able to be made solely based upon ~3-4% lower protolignin extractability. It is important to note that CELp from nitens contains a much higher carbohydrate content (12.3g total carbohydrate/100g CELp) relative to other protolignin preparations (~6-9g total carbohydrate/100g protolignin preparation). For this reason, it is possible to assume that nitens CELp will likely contain greater quantities of LCC substructures, relative to other hardwood protolignin preparations. Unfortunately, the total representativity of nitens protolignin preparations (36%) does not enable this assessment to be made for the entire nitens protolignin relative to maple and sweetgum protolignin.

Quantitation of aliphatic and phenolic hydroxyl groups in hardwood protolignin preparations is shown in Table 3. Aliphatic hydroxyl groups were seen to be slightly higher in CELp preparations versus MWLp. As discussed in our previous work[19], this quantitation is infiltrated by the presence of carbohydrates in each preparation. For this reason, it is not useful to compare these numbers for describing lignin's aliphatic hydroxyls (attributable to the "side-chain" of C9 units). For total phenolic hydroxyl groups, the amounts were seen to be higher in all MWLp preparations. This is due to the greater extent of degradation induced by ball milling in this preparation as opposed to CELp. However, it is also important to note that the phenolic hydroxyl groups in CELp are also derived from ball milling, due to the preparations being isolated sequentially. Each preparation contained more C3 & C5 bound phenolic hydroxyl groups compared to C3 or C5 bound phenolic hydroxyls (with the exception of sweetgum CELp), however, the inability to specifically assign these values to syringyl "S" or

guaiacyl “G” structures prohibits further interpretation of this information to produce phenolic hydroxyl S/G ratios.[23]

Native substructure quantities and non-condensed S/G ratios of the hardwood protolignin preparations are shown in Figure 4. The non-condensed S/G ratios of all CELp preparations were found to be higher compared to MWLp (CELp: 2.1, 2.7, 3.7; MWLp: 1.4, 2.0, 3.0). This indicates that CEL preparations, rendered extractable through scission of carbohydrates, contain a greater abundance of non-condensed S structures compared to the lignin rendered extractable exclusively from ball milling. The same trend was observed for total β-O-4’ substructures, with CELp bearing ~7-11 more than MWLp. This is most likely due to degradation of β-O-4’ substructures during ball milling, with the insoluble lignin (but extractable following CEL protocol) being constituted by a greater amount of these substructures. As mentioned in the discussion around phenolic hydroxyls, ball milling still exerted effects (albeit lesser) upon the CELp preparations, therefore the quantities of β-O-4’ substructures provided in Table 4 are still likely underestimates of what is actually present in the whole protolignin structure. No significant differences in β-β’ or β-5’ substructure quantities could be deduced from all protolignin preparations, which can be explained by these substructures not existing in high quantities within each protolignin. If there were more β-β’ or β-5’ in each protolignin, it would be reasonable to expect some sort of substructural differentiation between the two protolignin isolation protocols (MWL and CEL). Finally, the amount of LCC substructures in each protolignin preparation is minor, with most possessing ~1% uronosil ester substructures. However, it is important to point out that maple MWLp contained the greatest abundance of uronosil esters (2/100 Ar) as well as a quantifiable amount

of benzyl ethers (with primary carbohydrate alcohol attachment, 1/100 Ar). This data, however minor, can be interpreted to suggest that maple protolignin bears slightly greater LCC abundance compared to sweetgum and nitens. Finally, the high carbohydrate content of nitens CELp did not translate to a significant difference in LCC substructure quantities. This can be taken to indicate that the minor amount of LCCs present (1/100 Ar) bear longer chains of carbohydrates attached via LCC. The purification protocol applied enables such a suggestion, as it is assumed to remove all un-bound carbohydrates from the crude protolignin preparation.

3.3 Hardwood biorefinery processing

3.3.1. Autohydrolysis of hardwoods

Chemical composition of each hardwood after autohydrolysis pretreatment is shown in Table 5. Each biomass exhibited a similar amount of total solids yielded and solid carbohydrate recovery. Lignin solubilization was nearly identical for maple and sweetgum (3.9 and 3.8 g soluble lignin/100 g o.d. raw biomass, ~15% of total raw lignin), however, the amount of lignin solubilized from nitens was slightly greater. 4.7g soluble lignin/100g o.d. raw biomass was measured for nitens, representing ~19% of total raw lignin. This can be interpreted as another indication that the protolignin structure of nitens is unlike sweetgum and maple's.

The percent of adsorbable autohydrolyzate-soluble lignin ("XADL") was very high for all species (93%, 98%, 99% recovery of autohydrolyzate-soluble lignin) (Table 6). However, two significant observations can be made. First, the amount of un-adsorbed lignin was greatest for maple. This suggests a difference in the chemical structure of maple autohydrolyzate-

soluble lignin compared to the others, when considering both 1) the amount recoverable from sweetgum (98% recovery, same quantity of soluble lignin) and 2) maple CELp was constituted with the greatest amount of LCC substructures (compared to other protolignin preparations). We interpret this finding to indicate that the remaining ~7% of maple autohydrolyzate-soluble lignin is sufficiently hydrophilic via LCC connectivity to avoid adsorption to the utilized hydrophobic solute-targeting resin.[17, 30-31]

Chemical composition of each XADL preparation was, expectedly, carbohydrate-enriched (48.2, 57.1, and 39.6 g total carbohydrate/100 g adsorbate). Our previous works[17, 19] encountered similar carbohydrate-enrichment, which was best explained by adsorbate's prominent degree of lignin-carbohydrate chemical connectivity. A difference between adsorbates was observable, specifically concerning the amount of total lignin in each adsorbate. Maple and sweetgum, which bear 34.4 and 29.0 wt% total lignin (respectively), possess lower lignin than was quantifiable in nitens adsorbate (47.9 wt% total lignin). In fact, nitens adsorbate was the only adsorbate which was comprised of more lignin than carbohydrate (by mass). This further indicates nitens as demonstrative of outlier behaviour, which was suggested in previous discussion.

3.3.2. Biorefinery processing of autohydrolyzed hardwood solids

Autohydrolyzed solids were subjected mechanical refining followed by high-dosage (20 FPU/o.d. g solid) cellulolytic hydrolysis to render an extractable portion of lignin that is akin to protolignin CEL. The procurable lignin is appropriately labelled "AHCEL". We have made the assumption that the lignin content of autohydrolyzed solids remained constant

throughout both mechanical refining and cellulolytic hydrolysis. Preparation yields from cellulolytically-hydrolyzed solids were calculated from autohydrolyzed lignin abundance, and are shown in Table 7. AHCELp preparation yield was 25%, 30%, and 42% for maple, sweetgum, and nitens (respectively). The difference in lignin extractability was greatest when comparing nitens to maple and sweetgum. Still, the range of lignin extractabilities does indicate differences in chemical and molecular properties of the whole autohydrolyzed lignin. Nitens' strong extractability can be interpreted to indicate that it has undergone the most native substructure degradation, resulting in more-extractable low molecular weight lignin molecules. Alternatively, this observation could suggest that nitens has undergone the least amount of extraction-hampering inter-lignin condensation, which results in lignin molecules with significantly greater molecular mass following said chemical reaction. By this logic, the range of extractabilities may also indicate that maple has undergone the most inter-lignin condensation, and nitens the least. Finally, chemical composition of each AHCELp preparation was highly similar, each bearing significant lignin contents and minor carbohydrate constituency. This is in line with the effects caused by biorefinery processing, which targets carbohydrate fractions.

3.3.3. Hardwood biorefinery lignin preparation characterization

Aliphatic hydroxyl quantities in hardwood biorefinery lignin preparations (Table 8) remain contaminated with carbohydrate-derived response. Non-lignin response is especially visible in XADL, which has exacerbated aliphatic hydroxyl quantities of ~14-29 mmol/g total lignin in XADL. The heightened carbohydrate contents of hardwood XADL are the cause for

these unrealistically high values. AHCELp, which contains much lower carbohydrate, still contains enough for the reported aliphatic hydroxyl amounts (~4 mmol/g total lignin in AHCELp) to be beyond what is theoretically possible. For this reason, aliphatic hydroxyl groups remain unconsidered in terms of significant conclusions, and will not be included in the mass balance fitting of quantified lignin characteristics.

Total phenolic hydroxyl groups in both biorefinery lignin preparations were greater than protolignin preparations. In addition, greater variation between types of phenolic hydroxyls can be observed. First, total phenolic hydroxyls in XADL ranged were equal in maple and sweetgum (5.3 mmol/g total lignin in XADL), and nitens contained a much lower phenolic hydroxyl abundance (3.4 mmol/g total lignin in XADL). This can be taken to indicate that nitens XADL is of higher molecular mass than maple and sweetgum XADL, because genesis of phenolic hydroxyl groups is a function of native substructure degradation. By the same logic, maple and sweetgum XADL should be highly similar in molecular mass, based on phenolic hydroxyl abundance. The profile of phenolic hydroxyls in XADL demonstrated an abundance of C3 & C5 bound phenolic hydroxyls. This was also observable in AHCELp preparations. Unfortunately, lack of discrimination between S and G phenolic hydroxyls by this analysis disallows this to indicate which structure is more prevalent. For AHCELp, total phenolic hydroxyl amounts ranged from ~3-4 mmol/g total lignin in AHCELp. This amount was lower than XADL in all cases except for nitens. However, the limited isolation yield of AHCELp (25%, 30%, 42% representativity) does not rule out the possibility of additional phenolic hydroxyl functionality in the un-extractable lignin.

Estimated molecular mass of biorefinery lignin preparations is provided in Table 9. First, it can be seen that XADL from maple and sweetgum are indeed highly similar in molecular mass (400 Mw_n, 800 Mw_w), as previously speculated when discussing phenolic hydroxyls. XADL from nitens was also found to be significantly higher in approximate molecular mass (800 Mw_n, 1100 Mw_w), which also supports previous speculation. Importantly, each XADL fraction is orders lower in molecular mass compared to AHCELp. AHCELp estimated molecular masses ranged from ~7900-9200 g/mol (Mw_w). This proves that XADL is a much more degraded fraction of lignin, with the degradation induced by autohydrolysis pretreatment. It is important to note that these values demonstrate the distribution of lignin molecules present in each preparation, therefore polydispersity (Mw_w/Mw_n) can be taken as a metric of lignin heterogeneity. Based on the results, it is conceivable that the less polydisperse XADL fractions, which are easily isolated by the applied method, would be more valorizable based upon the limited differences between lignin molecules. In addition, XADL's water-solubility may be a significant advantage for development of biorefinery lignin co-products.

Native substructure abundancies of hardwood biorefinery lignin preparations is displayed in Table 10. First, it can be seen that the non-condensed S/G of all XADL preparations were quite high (4.9, 6.6, 6.8). This demonstrates that the C3 and C5 bound phenolic hydroxyl moieties measured in XADL are likely derived from S aromatic structures. Maple XADL, with the lowest non-condensed S/G (4.9), is comprised of much more G structures than sweetgum and nitens. For AHCELp, S/G was lower than XADL but still far from unity (2.6, 3.4, 4.6). This logically explains XADL's S enrichment, where the lignin

found in cellulolytic hydrolysis residue (represented by AHCELp) has undergone a loss of S structures during autohydrolysis. The total amount of β-O-4' substructures in XADL (34-46/100 Ar) was much greater than what was measured in AHCELp (18-24/100 Ar). Said observation suggests that the heavily degraded autohydrolyzate-soluble lignin (well represented by XADL) is dimeric-to-oligomeric in molecular structure, with native linkages remaining between lignin units. AHCELp's lower β-O-4' character is representative of solubilization of β-O-4' substructures. It is important to note that the un-extracted lignin from cellulolytic hydrolysis residue likely also contains β-O-4' character, however, some of these substructures are likely condensed with aromatic ring structures at the Cα position. It is likely due to this condensation that limited lignin extractability occurred.

As with protolignin preparations, no significant differences between biorefinery lignin preparations could be found in the amounts of β-β' and β-5' substructures. Significantly, however, was the quantities of LCC substructures observed exclusively in XADL. Total LCC quantities in XADL ranged from 7-11/100 Ar. This finding is in line with previous work's conclusions that such connectivity is assisting in the dimeric-to-oligomeric lignin's solubility. Finally, AHCELp from maple displayed a greater LCC character compared to the other hardwoods, however minor it may be (2/100 Ar). A similar observation was made in maple protolignin preparations, concluding that there is LCC "enrichment" in maple relative to sweetgum and nitens. This enrichment is derived from a greater presence of benzyl ether LCC substructures (with attachment to carbohydrate primary hydroxyls).

3.3.4. Biorefinery mass balance of hardwood lignin preparation properties

Tables 11, 12, and 13 display the incorporation of the measured properties in each hardwood lignin preparation into the mass balance of the investigated biorefinery process. Prior to discussion, it is imperative to note that two factors contribute most to the obtained quantities (normalized per 100g o.d. raw biomass). The first factor is the actual value measured. The second factor, perhaps most important, is the preparation yield (original lignin representativeness). It is not valid to claim that the characterization information provided is representative of the entirety of each respective original lignin, however, relative similarities in preparation yields across hardwoods enables discussion around the subset of lignin which was extractable.

Phenolic hydroxyls were generated in great abundance through biorefinery processing. With 11-15 mmol phenolic hydroxyl/100g o.d. raw biomass measured in protolignin preparations, this amount was quantified to be much greater in biorefinery lignin preparations (36-47 mmol/100g o.d. raw biomass). This demonstrates the chemical degradation effects of biorefinery processing upon hardwood lignin. Nitens AHCELp is observed to contribute the most to phenolic hydroxyl contents amongst all biorefinery lignin preparations, best explained by the preparation's high yield (42%), as the phenolic hydroxyl quantity of said preparation was not significantly higher than what was procured from maple and sweetgum (3.7 vs. 3.0 and 3.2 mmol total phenolic hydroxyl/g total lignin in AHCELp).

The abundances of non-condensed S and G units, derived from preparation S/G ratio, demonstrate scrutable behavior around biorefinery processing. As indicated by all hardwood lignin preparation's S/G, each was S enriched to different extents. The total amount of non-

condensed S units quantifiable across hardwood protolignin preparations was 33-34 mmol/100g o.d. raw biomass. After biorefinery processing, this amount increased to 34-54 mmol/100g o.d. raw biomass. This indicates that a greater amount of S units were liberated through biorefinery processing compared to protolignin isolation. However, as noted in discussion around preparation yield's effect upon the applied quantitation, this data could be more elucidative if higher protolignin preparation yields had been obtained. This could be achievable through increased ball milling times, solvent adjustment, abandonment of purification protocol[19], or utilization of a more degradative protolignin isolation protocol (enzyme mild hydrolysis lignin[32-33]). One conclusion can be drawn concerning non-condensed S structures: inter-lignin condensation of non-condensed S units appears to be unquantifiable from the data sets procured across all hardwood species studied. This conclusion is supported by the increased abundance of non-condensed S structures in biorefinery lignin preparations versus protolignin preparations.

Unlike non-condensed S structures, non-condensed G structural quantities were quite affected by autohydrolysis. Quantities of non-condensed G structures, measured as 11-19 mmol/100g o.d. raw biomass in protolignin preparations, fell to 9-10 mmol/100g o.d. raw biomass after biorefinery processing. Across hardwoods, the greatest loss to non-condensed G recovery was found in maple (52% recovery), followed by sweetgum (60% recovery), and finally, nitens (91% recovery). This is an absolute indicator of inter-lignin condensation occurring to native non-condensed G aromatic structures, with maple and sweetgum experiencing the greatest degree of inter-lignin condensation. A similar trend is also visible in AHCELp preparation yields, where maple (25% yield), sweetgum (30% yield), and nitens

(42% yield) demonstrate a strong linear correlation (R^2=0.99) between non-condensed G recovery and lignin extractability from cellulolytic hydrolysis residue (Figure 3). Higher degrees of inter-lignin condensation result in higher molecular weight lignin molecules, which tend to demonstrate insolubility in lignin-friendly solvents. Finally, nitens appears to lack proclivity for inter-lignin condensation amongst its non-condensed G structures (~10% loss), which is perhaps explainable by its protolignin's great degree of S enrichment.

β-O-4' substructure quantities, similar across hardwood protolignin preparations (20-24/100g o.d. raw biomass), were found to decrease in different extents following biorefinery processing. The combined amount of recovered β-O-4' substructures from biorefinery lignin preparations was 12-18/100g o.d. raw biomass, representing ~50-90% β-O-4' recovery. A large range in β-O-4' recovery suggests β-O-4' acidolysis is most prominent in maple (50% recovery), followed by sweetgum (76% recovery), and least in nitens (90% recovery). This indicates that maple's degree of inter-lignin condensation is not specifically related to degree of β-O-4' acidolysis. Instead, it seems that the exaggerated loss of both non-condensed G structures and β-O-4' substructures in maple appears to be a result of competing degradative reactions: 1) acidolysis around β-O-4' substructures containing at least one non-condensed G constituent, and 2) inter-lignin condensation involving non-condensed G structures. Part of this claim is supported by the conclusions drawn by Matsumoto et al[34], who demonstrated that non-condensed G structures are more prone to acidolysis than non-condensed S structures. Finally, no significant conclusions could be drawn from re-distribution of LCC substructures. However, it is clear that autohydrolyzate-soluble lignin (well-represented by XADL) is enriched with the LCCs that were minor substructures in each respective protolignin.

The cumulative sum of these findings mostly indicates that hardwood lignin's biorefinery degradation patterns are most influenced by protolignin S/G, which defines the availability of inter-lignin condensable or acidolysis-reactive G units. The differences in S/G are relatable to inter-lignin condensation around G structures (Figure 3) as well as β-O-4' recovery (Figure 4). These results indicate that protolignin with higher G character undergoes a greater amount of both acidolysis as well as inter-lignin condensation. In contrast, the highest S/G protolignin, nitens, demonstrated minor losses to non-condensed G units, β-O-4' substructures, and provided cellulolytic hydrolysis residue with the greatest amount of extractable lignin. In all, it can be concluded that biorefinery processing of high S/G hardwood species render heavier but less polydisperse autohydrolyzate-soluble lignins, as well as a cellulolytic hydrolysis residue with less condensed lignin structures, which contain in-tact native substructure character.

3.4. Non-wood protolignin

As mentioned in previous discussion, non-wood protolignin is acknowledged to be less comprehensible compared to hardwood protolignin. This is explainable by several traits exclusive to non-wood protolignin. Such distinct characteristics include the presence of *para*-coumaric acid and ferulic acid, which are lignin-like in chemical structure and may or may not be chemically attached to the protolignin structure. Next, hardwood lignins also contain 4-hydroxyphenyl aromatic rings, referred to as non-condensed or condensed "H" structures. Finally, dibenzodioxicin substructures ("DBDO") are present within various non-wood

protolignins to varying extents. This substructure's unique aromatic carbon-carbon connectivity and dual aryl ether bonds can further complicate lignin reactivity.

Revisiting Table 1, all three non-woods contain mostly similar amounts of lignin, ranging from 20.5-22.3 wt% total lignin. This range is slightly greater than what was measured in hardwood. Based on previous elucidation of the wide range of hardwood lignin responses in spite of similar lignin contents, a similar range of results is to be expected for non-woods. Beginning with non-wood protolignin preparation yields (Table 14), MWLp yields were found to be 17%, 20%, and 25% for wheat straw, switchgrass, and sugarcane bagasse (respectively). For CELP, preparation yields were 17%, 23%, and 30% for wheat straw, sugarcane bagasse, and switchgrass (respectively. The cumulative sum of protolignin preparation yields was 38% (wheat straw), 48% (sugarcane bagasse), and 50% (switchgrass). Identifying itself as an outlier, sugarcane bagasse required 10 hr of ball milling (6 hr all others) to obtain MWLp preparation yields greater than 6% (6 hrs) and 13% (8 hrs) (data not tabulated). In stark difference, 6 hrs of ball milling rendered MWLp preparation yields of 17% (wheat straw) and 30% (switchgrass). This observation already alludes to the vast differences in protolignin structure across the investigated non-wood biomasses.

Hydroxyl group quantitation in non-wood protolignin preparations is displayed in Table 15. As discussed around hardwood lignin aliphatic hydroxyls, the quantities obtained from non-wood protolignin preparations (MWLp: 4.7-6.7 mmol/g total lignin in preparation; CELp: 6.3-6.8 mmol/g total lignin in preparation) are disfigured by the presence of carbohydrates. For this reason, aliphatic hydroxyl quantities in non-wood protolignin preparations, as well as in biorefinery lignin preparations, will not be further discussed.

However, phenolic hydroxyl groups, which are exclusive to lignin, are demonstrative of phenolic hydroxyl abundance and distribution. Within the MWLp preparations, total phenolic hydroxyl was found to be highest in sugarcane bagasse (3.0 mmol/g total lignin in MWLp) compared to wheat straw and switchgrass (1.7-1.8 mmol/g total lignin in MWLp). This may indicate that sugarcane bagasse protolignin bears an enrichment of free phenolic hydroxyls, however, the effect of additional ball milling time most likely rendered the observed phenolic hydroxyl enrichment. Of significance, sugarcane bagasse bears significant enrichment of C3 & C5 unbound phenolic hydroxyls compared to the other non-wood protolignin preparations. This functional group is solely attributable to non-condensed H structures and *para*-coumaric acid, therefore it can be expected that further characterization will reveal sugarcane bagasse to bear an enrichment of H structures and/or *para*-coumaric acid. CELp preparations were found to contain 1.5-2.3 mmol phenolic hydroxyl/g total lignin in each respective preparation. This amount was slightly lower compared to MWLp preparations (with the exception of wheat straw). Sugarcane bagasse CELp also contained an enrichment of C3 and C5 unbound phenolic hydroxyls, confirming that sugarcane bagasse protolignin uniquely bears an abundance of H structures and/or *para*-coumaric acid that is dissimilar to wheat straw and switchgrass protoligin.

Non-condensed S/G, H/G, and substructural quantification from non-wood protolignin preparations (Table 16) rendered several biomass-specific characteristics. Non-condensed S/G was highest in sugarcane bagasse (1.1-1.2), versus wheat straw and switchgrass (0.7 and 0.4-0.5, respectively). Non-condensed H/G in sugarcane bagasse was also significantly higher compared to the others (1.5 & 1.8 versus 0.3-0.5). This confirms that sugarcane bagasse bears

non-condensed H enrichment derived through both non-condensed H aromatic structures and *para*-coumaric acid. Based on the elevated non-condensed S/G and H/G ratios in sugarcane bagasse protolignin preparations, it can be insinuated that the biomass bears very little non-condensed G constituency. For switchgrass and wheat straw, small differences in each respective non-condensed H/G and S/G suggests a lack of significant monolignol enrichment.

Quantification of *para*-coumaric acid and ferulic acid revealed that sugarcane bagasse is significantly more enriched in *para*-coumaric acid (20 & 29/100 Ar) compared to the other non-wood biomasses (6-14/100 Ar). For this reason, the non-condensed H/G enrichment in sugarcane bagasse is due in part to the strong presence of *para*-coumaric acid moieties. Ferulic acid displayed a minor presence across all non-wood protolignin preparations, ranging from 2-5/100 Ar. Because *para*-coumaric acid was more prevalent than ferulic acid in all protolignin preparations, it can be suggested that *para*-coumaric acid is the predominant hydroxycinnamic acid produced during non-wood protolignin biosynthesis.

Total β-O-4' varied relatively insignificantly across all non-wood protolignin preparations, with wheat straw demonstrating the greatest β-O-4' character (30, 31/100 Ar) and sugarcane bagasse demonstrating the least (23, 25/100 Ar). Similar to hardwood, β-β' and β-5' substructures are minor constituents of non-wood protolignin. However, the amount of β-5' measured is slightly higher in non-woods (3-5/100 Ar) than what was measured in hardwoods (2-4/100 Ar). DBDO presents itself in varying amounts across the non-wood protolignin preparations. However, the structure consistently appears nearly-exclusively in CELp preparations (2-7/100 Ar) compared to MWLp preparations (0-1/100 Ar). These results reveal DBDO to be best represented by CEL preparations, suggesting that the methodology for

CEL procurement specifically renders DBDO enrichment. It can be logically speculated that DBDO near-exclusive presence in CELp will also correlate with an increase LCC substructure abundance. However, results revealed this speculation untrue, with the amount of total LCC substructures in MWL (5-11/100 Ar) being greater than CELp (4-9/100 Ar) in all non-wood protolignin preparations. Therefore it can be concluded that CELp's DBDO enrichment is due to an unknown factor, having ruled out enhanced LCC substructure quantities being responsible.

3.5 Non-wood biorefinery processing

3.5.1. Autohydrolysis of non-woods

Chemical compositions of each non-wood biomass following autohydrolysis is displayed in Table 17. First, it can be seen that a range of solid mass yields occurs (53g, 55g, 59g solid recovery for switchgrass, wheat straw, and sugarcane bagasse (respectively)). This may be in part due to the greater amount of autohydrolyzate-soluble lignin released during autohydrolysis of switchgrass (7.6 g), followed by wheat straw (6.7 g), and finally sugarcane bagasse (6.5 g autohydrolyzate-soluble lignin/100g o.d. raw biomass). Another possibility for sugarcane bagasse's heightened solid recovery may be due to structurally-incorporated inorganic ash exerting a buffering effect upon autohydrolyzate pH.[35] As shown, sugarcane bagasse retains over twofold the amount of inorganic ash (1.7 g) than wheat straw (0.7 g) and switchgrass (0.4 g/100g o.d. raw biomass).

Table 6 displays the extent of adsorbable non-wood autohydrolyzate-soluble lignin. The strongest degree of adsorption occurred in wheat straw autohydrolyzate, producing an

XADL preparation with 98% representativeness. Next was switchgrass (94% adsorption) followed by sugarcane bagasse (92% adsorption). Chemical composition of the non-wood XADL preparations is tabulated in Table 18. As expected, these too also bear significant carbohydrate enrichment, ranging from 36-48 g total carbohydrate/100g o.d. XADL. With knowledge that these are due to the presence of LCC substructures, it is also feasible to hypothesize that sugarcane bagasse (47.9 g) will bear the greatest LCC character, followed by switchgrass (46.9 g), and finally, wheat straw (36.4 g total carbohydrate/100g o.d. XADL). The total lignin contents of XADL preparations did vary across the non-woods. Specifically, the XADL with the most total lignin was wheat straw (35.4 g), followed by switchgrass (29.7 g), and finally, sugarcane bagasse (23.8 g/100g o.d. XADL). Lignin content of each adsorbate was found to not correlate with the amounts of autohydrolyzate-soluble lignin produced per gram o.d. raw biomass. Finally, the amount of uncharacterized components was much higher for non-wood XADL than hardwood. We hypothesize this to be due to non-wood-exclusive hydroxycinnamic acids avoiding explicit quantitation during conventional compositional analysis.

3.5.2. Biorefinery processing of autohydrolyzed non-wood solids

Chemical composition and preparation yields of AHCELp preparations derived from non-woods is displayed in Table 18. The total lignin abundance in each was very high in similarity to hardwood AHCELp. Specifically, AHCELp total lignin contents ranged from ~86-90 g total lignin/100g o.d. AHCELp. Diminished carbohydrate abundances were also observed in non-wood AHCELp, with wt% total carbohydrate ranging from 5-6%. The

dominant remaining carbohydrate was xylan, with switchgrass AHCELp containing the greatest xylan quantity (3.8 g) compared to wheat straw (2.5 g) and sugarcane bagasse (2.3 g xylan/100g o.d. AHCELp). This suggests that this xylan was both resistant to autohydrolysis as well as cellulolytic enzyme hydrolysis (including hemicellulose). For that reason, it can be suggested that the LCC substructures quantified in AHCELp either directly or indirectly include xylan.

3.5.3. Non-wood biorefinery preparation characterization

A range of total phenolic hydroxyl quantities was measurable across non-wood XADL preparations (4.3-7.3 mmol total phenolic hydroxyl/g total lignin in XADL) (Table 19). Sugarcane bagasse XADL again demonstrated significant enrichment of C3 and C5 unbound phenolic hydroxyls compared to other preparations. This suggests that the abundance of these functional groups measured in the protolignin preparations was not exclusively related to extended ball milling times. Instead, it appears that the protolignin itself bears enrichment of non-condensed H phenolic hydroxyls and/or *para*-coumaric acid. Phenolic hydroxyl quantities were mostly similar in non-wood AHCELp, with recorded values of 2.9 mmol (wheat straw and switchgrass) and 3.6 mmol (sugarcane bagasse) total phenolic hydroxyl/g total lignin in AHCELp. As with protolignin and XADL, sugarcane bagasse AHCELp does contain a greater amount of C3 and C5 unbound phenolic hydroxyls compared to other non-wood AHCELp preparations.

Molecular mass estimates (Table 20) revealed some distinguishing factors for certain XADL and AHCELp preparations. Concerning XADL, weight-averaged molecular mass

ranged from 1000-2000 g/mol, with wheat straw being greatest mass and sugarcane bagasse the lowest. This range perfectly correlates with the total phenolic hydroxyl abundancy in each XADL preparation (Figure 5), demonstrating a strong relationship between non-wood autohydrolyzate lignin's phenolic functionality and weight-averaged molecular mass. A similar relationship was visible for hardwood XADL, however, the absence of a range of phenolic hydroxyls and XADL Mw_w prohibited useful correlation of this information. As discussed in previous sections, the presence of phenolic hydroxyls in biorefinery lignin preparations is due to degradation of ether-containing native substructures. Therefore it can be assumed that sugarcane bagasse XADL will contribute the least to native substructure recovery amongst the non-wood XADLs.

Table 21 contains non-condensed S/G and H/G of all non-wood biorefinery lignin preparations, as well as quantities of native substructures. First, sugarcane bagasse continues to demonstrate its non-condensed H and *para*-coumaric acid enrichment (H/G = 7.0, 18-20 *para*-coumaric acid/100 Ar). Non-condensed S/G was found not to vary across XADL preparations, however, it does vary across AHCELp. The non-condensed S/G ratios across AHCELp preparations are measured to be 0.7 (switchgrass), 1.3 (wheat straw), and 2.3 (sugarcane bagasse). As noted during protolignin discussion, sugarcane bagasse demonstrated a native non-condensed S enrichment with respect to the other non-woods. This observation holds true when evaluating the sugarcane bagasse lignin isolated from cellulolytic hydrolysis residue.

As touched on in the previous paragraph, the quantities of *para*-coumaric acid and ferulic acid wildly varied between preparations as well as between biomasses. For XADL,

para-coumaric acid was measured to be only 4 for wheat straw, yet as high as 20 (sugarcane bagasse) and 22/100 Ar (switchgrass). Ferulic acid was also dominant in XADL, ranging from 11-18/100 Ar across XADL preparations. The total amount of hydroxycinnamic acids in XADL preparations demonstrates that a significant amount of these structures are liberated into autohydrolyzate during autohydrolysis, likely contributing to overall autohydrolysis-caused delignification. Interestingly, wheat straw XADL exclusively favors ferulic acid, in contrast to *para*-coumaric acid prevalence throughout all other XADL and AHCELp preparations. Hydroxycinnamic acids were less present in AHCELp preparations, however, *para*-coumaric acid was still found to enrich sugarcane bagasse AHCELp (18/100 Ar). This is further demonstrative of sugarcane protolignin biosynthesis promoting production of a significant quantity of these structures.

Total β-O-4' character did not vary significantly across both XADL and AHCELp preparations, 10-16/100 Ar. Implementation of these values into a mass balance will produce more informative evaluation of the migration of non-wood β-O-4' substructures around biorefinery processing. DBDO again revealed itself to be prevalent in AHCELp (1-4/100 Ar), and especially within XADL (2-8/100 Ar). The abundance of DBDO in autohydrolyzate-soluble lignin is especially interesting considering its naturally condensed 5-5' bond, which is typically viewed as an insoluble structure. As discussed around DBDO's exclusive enrichment in CELp, DBDO quantities in XADL do not directly correlate with any specific LCC substructural abundance nor the total sum of quantified LCC substructures. DBDO's strong inclusion in CELp and XADL still remains unexplained. LCC substructures were abundant in all non-wood biorefinery lignin preparations (8-14) with the exception of wheat straw

AHCELp (3/100 Ar). This indicates that non-wood LCC substructures are resilient to degradation by autohydrolysis. However, it cannot be definitely said they total recovery of LCC substructures is observed. This information will be obtained after implementation of non-wood protolignin and biorefinery lignin preparations into the process' mass balance.

3.5.4. Biorefinery mass balance of non-wood lignin preparation properties

Fitting of all the previously discussed non-wood protolignin and biorefinery lignin properties into each respective biomass' biorefinery mass balance is found in Table 22, Table 23, and Table 24. As observed during similar analysis of hardwood lignin, the phenolic hydroxyl quantities are significantly higher in biorefinery lignin preparations relative to each respective protolignin preparation. This is best explained as chemical degradation of aryl-ether substructures to liberate new phenolic hydroxyls, something that is better induced by autohydrolysis than planetary ball milling.

Quantities of non-condensed S structures in biorefinery lignin preparations (15-22 total) were greater than what was quantifiable in the protolignin preparations (13-16 total mmol non-condensed S/100g o.d. raw biomass). Liberation of non-condensed S structures through biorefinery processing was also observed in the hardwood biomasses, leading to a greater quantity of non-condensed S structures being quantifiable from biorefinery lignin preparations versus protolignin preparations. A similar phenomenon is observable for the hardwoods, albeit to a lesser extent. In similarity, non-condensed H structures were greater than what was quantified from protolignin preparations, however to a very significant extent. Specifically, the units quantifiable in protolignin preparations summated to 8-23 mmol/100g o.d. raw

biomass. In stark contrast, the amount quantifiable in biorefinery lignin preparations ranged from 24-39 mmol/100g o.d. raw biomass. This suggests a phenomenon similar to non-condensed S units also occurs for non-condensed H units, i.e. liberation of aromatic structures not isolatable using the applied protolignin isolation protocol. Finally, a range of non-condensed G structures were lost from biorefinery lignin preparations that were quantifiable in protolignin preparations. Specifically, recovery of non-condensed structures was 63% for switchgrass, 67% for sugarcane bagasse, and 83% for wheat straw. This is an indicator of inter-lignin condensation also occurring in non-wood non-condensed G units, akin to what was observable in the hardwoods.

No significant difference in *para*-coumaric acid distributions could be identified across the non-wood lignin preparations. Recovery of *para*-coumaric acid was 100% in both sugarcane bagasse and wheat straw, and the amount quantified in switchgrass biorefinery lignins was actually higher than what could be found in its protolignin preparations. This indicates that *para*-coumaric acid is not involved in a quantifiable amount of inter-lignin condensation, suggesting that it is mostly inert with respect to said degradative chemical reaction. Concerning ferulic acid, a large amount was liberated and found soluble in each XADL preparation. Based on both hydroxycinnamic acid's displayed behaviour, it appears that autohydrolysis results in either liberation or cleavage of lignin-hydroxycinnamic acid bonds, resulting in the production of soluble (but inert) *para*-coumaric and ferulic acid. A similar hypothesis can be drawn for the behavior of DBDO substructures, which are found in highest abundance in XADL preparations. For this reason, it is possible to assume that DBDO is chemically attached to low molecular weight lignin fragments, or is associated with a

network of autohydrolysis-labile bonds that render lignin molecules containing DBDO substructures soluble. LCC substructures also dominate XADL, representing the majority of LCC substructures quantifiable across both protolignin and biorefinery lignin preparations. The hydrophilic nature of chemically attached carbohydrates seems to promote lignin solubility. This solubility promotion is likely greater when the LCC includes a particularly long carbohydrate chain, or when the lignin molecule attached is minor in degree of polymerization (dimer, trimer, etc.).

β-O-4' demonstrated greater instability in non-woods than was observable in hardwoods. Across all non-wood biomasses, recovery of β-O-4' substructures ranged from 56-62%. The presence of non-condensed S structures did not protect β-O-4' from acidolysis as was observable in nitens (hardwood). Instead, β-O-4' acidolysis must be occurring mostly around H-containing β-O-4' structures, and possibly even S-containing β-O-4' substructures. Acidolysis of β-O-4' could be rendering greater amounts of non-condensed H and S structures quantifiable after biorefinery processing due to the decrease in molecular mass, which enables characterization of a greater amount of biorefinery lignin compared to protolignin.

Totalling the findings of Tables 22-24, it is surmisable that the behavior of non-wood lignin around biorefinery processing is not as clearly understood as hardwood. While there is an observable interplay between inter-lignin condensation (measured only for G structures) and β-O-4' acidolysis, specific patterns with respect to both cannot be drawn from the compiled data. This is because no trends were observable between residual lignin extractability nor recovery of non-condensed aromatic structures. What is certain, however, is that autohydrolysis results in major liberation of non-condensed H structures and minor liberation

of non-condensed S structures. In addition, significant amounts of *para*-coumaric acid and ferulic acid, as well as lignin molecules bearing LCC and DBDO substructure enrichment, are rendered soluble within autohydrolyzate. Based on these findings, it appears that the only labile substructure identified (β-O-4'), results in the release of lignin fragments with said substructures in-tact.

4. Conclusions

A wealth of lignin characterization data acquired from protolignin and biorefinery lignin preparations was compiled into the biorefinery's mass balance with respect to each individually tested biomass. The findings from such analysis almost definitively elucidated how hardwood protolignin influences biorefinery lignin properties. Protolignin non-condensed S/G was identified as the key driving force towards lignin reactivity during autohydrolysis, with lower S/G hardwoods demonstrating evidence of increased inter-lignin condensation and acidolysis compared to higher S/G hardwoods. This suggest that the more degraded (both acidolyzed and condensed) lignin will be present for a biorefinery process using hardwoods with low S/G. This degradation will influence lignin's applicability in different valorization schemes. For non-woods, no clear conclusion could be drawn in spite of observation of inter-lignin condensation around G structures and acidolysis of β-O-4' substructures. It was found that a significant amount of hydroxycinnamic acids, LCCs, and DBDO substructures are liberated from lignin during autohydrolysis, suggesting their attachment to the observed acidolyzed β-O-4' substructures.

REFERENCES

1. Phillips, R. B.; Jameel, H.; Chang, H.-m. Integration of pulp and paper technology with bioethanol production. *Biotechnology for Biofuels* **2013,** *6* (13).

2. Treasure, T.; Gonzalez, R.; Jameel, H.; Phillips, R. B.; Park, S.; Kelley, S. Integrated conversion, financial, and risk modeling of cellulosic ethanol from woody and non-woody biomass via dilute acid pre-treatment. *Biofuels, Bioproducts, & Biorefining* **2014,** *8*, 755-769.

3. Giuliano, A.; Cerulli, R.; Poletto, M.; Raiconi, G.; Barletta, D. Process pathways optimization for a lignocellulosic biorefinery producing levulinic acid, succinic acid, and ethanol. *Industrial & Engineering Chemistry Research* **2016,** *55*, 10699-10717.

4. Wang, C.; Li, Q.; Tang, H.; Zhou, W.; Yan, D.; Xing, J.; Wan, Y. Clarification of succinic acid fermentation broth by ultrafiltration in succinic acid bio-refinery. *Chemical Technology and Biotechnology* **2013,** *88*, 444-448.

5. Pileidis, F. D.; Titirici, M.-M. Levulinic acid biorefineries: new challenges for efficient utilization of biomass. *ChemSusChem* **2016,** *9*, 562-582.

6. Mazzoli, R.; Bosco, F.; Mizrahi, I.; Bayer, E. A.; Pessione, E. Towards lac acid bacteria-based biorefineries. *Biotechnology Advances* **2014,** *32*, 1216-1236.

7. Kamat, S.; Khot, M.; Zinjarde, S.; Ravikumar, A.; Gade, W. N. Coupled production of single cell oil as biodiesel feedstock, xyliotl and xylanase from sugarcane bagasse in a biorefinery concept using fungi from the tropical mangrove wetlands. *Bioresource Technology* **2013,** *135*, 246-253.

8. Beckham, G. T.; Johnson, C. W.; Karp, E. M.; Salvachua, D.; Vardon, D. R. Opportunities and challenges in biological lignin valorization. *Current Opinion in Biotechnology* **2016,** *42*, 40-53.

9. Cline, S. P.; Smith, P. M. Opportunities for lignin valorization: an exploratory proces. *Energy, Sustainability and Society* **2017,** *7* (26), 1-12.

10. Karkas, M. D., Matsuura, B. S., Monos, T. M., Magallanes, G., Stephenson, C. R. J. Transition-metal catalyzed valorization of lignin: the key to a sustainable carbon-neutral future. *Organic & Biomolecular Chemistry* **2016,** *14* (6), 1853-1914.

11. Raveendra, G., Srinivas, M., Pasha, N., Prasada Rao, A. V., Sai Prasad, P. S., Lingaiah, N. Heteropoly tungstate supported on tantalum oxide: a highly active acid catalyst for the selective conversion of fructose to 5-hydroxy methyl furfural. *Reaction Kinetics, Mechanisms, and Catalysis* **2015,** *115*, 663-678.

12. Linger, J. G., Vardon, D. R., Guarnieri, M. T., Karp, E. M., Hunsinger, G. B., Franden, M. A., Johnson, C. W., Chupka, G., Strathmann, T. J., Pienkos, P. T., Beckham, G. T. Lignin valorization through integrated biological funneling and chemical catalysis. *Proceedings of the National Academy of Sciences* **2014,** *111* (33), 12013-12018.

13. Ma, X.; Cui, K.; Hao, W.; Ma, R.; Tian, Y.; Li, Y. Alumnia supported molybdenum catalyst for lignin valorization: Effect of reduction temperature. *Bioresource Technology* **2015,** *192*, 17-22.

14. Narron, R. H.; Kim, H.; Chang, H.-m.; Jameel, H.; Park, S. Biomass pretreatments capable of enabling lignin valorization in a biorefinery process. *Current Opinion in Biotechnology* **2016,** *38*, 39-46.

15. Ragauskas, A. J.; Beckham, G. T.; Biddy, M. J.; Chandra, R.; Chen, F.; Davis, M. F.; Davison, B. H.; Dixon, R. A.; Gilna, P.; Keller, M.; Langan, P.; Naskar, A. K.; Saddler, J. N.; Tschaplinski, T. J.; Tuskan, G. A.; Wyman, C. E. Lignin valorization: improving lignin processing in the biorefinery. *Science* **2014,** *344* (6185), 709-720.

16. Boes, K. S.; Narron, R. H.; Chen, Y.; Park, S.; Vinueza, N. R. Characterization of biofuel refinery byproduct via selective electrospray ionization tandem mass spectrometry. *Fuel* **2017,** *188*, 190-196.

17. Narron, R. H.; Chang, H.-m.; Jameel, H.; Park, S. Soluble lignin recovered from biorefinery pretreatment hydrolyzate characterized by lignin-carbohydrate complexes. 2017.

18. Narron, R. H.; Han, Q.; Park, S.; Chang, H.-m.; Jameel, H. Lignocentric analysis of a carbohydrate-producing lignocellulosic biorefinery process. *Bioresource Technology* **2017,** *241*, 857-867.

19. Narron, R. H.; Park, S.; Chang, H.-m.; Jameel, H. Following a hardwood protolignin through biorefinery processing. 2017.

20. Sluiter, J. B.; Ruiz, R. O.; Scarlata, C. J.; Sluiter, A. D.; Templeton, D. W. Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. *Journal of Agricultural and Food Chemistry* **2010,** *58* (16), 9043-9045.

21. Li, S.; Ogunkoya, D.; Fang, T.; Willoughby, J.; Rojas, O. J. Carboxymethylated lignins with low surface tension toward low viscosity and highly stable emulsions of crude bitumen and refined oils. *Journal of Colloid and Interface Science* **2016,** *482*, 27-38.

22. Cui, C.; Sun, R.; Argyropoulos, D. S. Fractional precipitation of softwood Kraft lignin: isolation of narrow fractions common to a variety of lignins. *ACS Sustainable Chemistry & Engineerng* **2014,** *2* (4), 959-968.

23. Balakshin, M.; Capanema, E. On the quantification of lignin hydroxyl groups with ^{31}P and ^{13}C NMR spectroscopy. *Journal of Wood Chemistry and Technology* **2015,** *35* (3), 220-237.

24. Capanema, E.; Balakshin, M.; Kadla, J. A comprehensive approach for quantitative lignin characterization by NMR spectroscopy. *Journal of Agricultural and Food Chemistry* **2004,** *52* (7), 1850-1860.

25. Balakshin, M.; Capanema, E.; Gracz, H.; Chang, H.-m.; Jameel, H. Quantification of lignin-carbohydrate linkages with high-resolution NMR spectroscopy. *Planta* **2011,** *233* (6), 1097-1110.

26. Ralph, J.; Akiyama, T.; Kim, H.; Lu, F.; Schatz, P.; Marita, J.; Ralph, S.; Reddy, M.; Chen, F.; Dixon, R. Effects of coumarate 3-hydroxylase down-regulation on lignin structure. *Journal of Biological Chemistry* **2006,** *281* (13), 8843-8853.

27. Yelle, D. J.; Ralph, J.; Frihart, C. R. Characterization of nonderivatized plant cell walls using high-resolution solution-state NMR spectroscopy. *Magnetic Resonance in Chemistry* **2008,** *46* (6), 508-517.

28. Yelle, D. J.; Kaparaju, P.; Hunt, C. G.; Hirth, K.; Kim, H.; Ralph, J.; Felby, C. Two-dimensional NMR evidence for cleavage of lignin and xylan substituents in wheat straw through hydrothermal pretreatment and enzymatic hydrolysis. *BioEnergy Research* **2013,** *6* (1), 211-221.

29. Kim, H.; Ralph, J.; Akiyama, T. Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d_6. *BioEnergy Research* **2008,** *1*, 56-66.

30. Schwartz, T. J.; Lawoko, M. Removal of acid-soluble lignin from biomass extracts using Amberlite XAD-4 resin. *BioResources* **2010,** *5* (4), 2337-2347.

31. Huang, C.; Jeuck, B.; Du, J.; Yong, Q.; Chang, H.-m.; Jameel, H.; Phillips, R. B. Novel process for the coproduction of xylo-olgiosaccharides, fermentable sugars, and lignosulfonates from hardwood. *Bioresource Technology* **2016,** *219*, 600-607.

32. Guerra, A.; Filpponen, I.; Lucia, L. A.; Argyropoulos, D. S. Comparative evaluation of three lignin isolation protocols for various wood species. *Journal of Agricultural and Food Chemistry* **2006,** *54*, 9696-9705.

33. Zoia, L.; Orlandi, M.; Argyropoulos, D. S. Microwave-assisted lignin isolation using the enzymatic mild acidolysis (EMAL) protocol. *Journal of Agricultural and Food Chemistry* **2008,** *56*, 10115-10122.

34. Matsumoto, Y.; Yokoyama, T.; Akiyama, T.; Shioya, T. Formation rate of benzyl cation intermediate from p-Hydroxyphenyl, guaiacyl, or syringyl nucleus in acidolysis of lignin. *Journal of Wood Chemistry and Technology* **2017,** *37* (2), 75-86.

35. Huang, C.; Wu, X.; Huang, Y.; Lai, C.; Li, X.; Yong, Q. Prewashing enhances the liquid hot water pretreatment efficiency of waste wheat straw with high free ash content. *Bioresource Technology* **2016,** *219*, 583-588.

TABLES

Table 1. Chemical composition of six lignocellulosic biomasses

Biomass / Chemical component	Sugarcane bagasse	Wheat straw	Switchgrass	Maple	Sweetgum	Nitens
Glucan	42.3	39.0	34.3	43.2	40.2	44.4
Xylan	20.9	18.8	24.1	13.0	15.7	14.3
Galactan	0.9	1.2	0.9	0.7	0.8	1.1
Arabinan + Mannan[a]	2.6	3.1	3.1	3.4	2.2	1.4
Total carbohydrate	66.7	62.1	62.4	60.3	58.9	61.2
Acid-insoluble lignin	18.2	19.4	18.4	22.1	21.5	19.9
Acid-soluble lignin	2.3	2.9	2.9	3.1	3.7	4.6
Total lignin	20.5	22.3	21.3	25.2	25.2	24.5
Extractives[b]	3.9	1.9	2.1	2.0	1.2	1.1
Ash	2.7	5.6	3.0	0.3	0.6	0.2
Uncharacterized	6.2	8.1	11.2	12.2	14.1	13.0

[a] Co-eluting monosaccharides

[b] Measured gravimetrically, removed prior to compositional analysis

Table 2. Chemical composition and preparation yields of hardwood protolignin preparations

Biomass / Chemical component	Maple		Sweetgum		Nitens	
	MWLp	CELp	MWLp	CELp	MWLp	CELp
Glucan	2.0	2.8	2.0	2.5	1.5	3.2
Xylan	3.3	1.2	1.2	1.3	2.0	2.3
Galactan	0.7	1.0	0.8	1.1	1.6	2.6
Arabinan + Mannan[a]	2.2	2.1	1.9	1.6	3.6	4.2
Total carbohydrate	8.2	7.1	5.9	6.5	8.7	12.3
Acid-insoluble lignin	82.5	86.0	86.6	90.1	83.3	76.6
Acid-soluble lignin	0.3	0.2	0.3	0.3	0.4	0.3
Total lignin	82.8	86.2	86.9	90.4	83.7	76.9
Uncharacterized	9.0	6.7	7.2	3.1	7.6	10.8
Gravimetric solid yield[b]	24.6	23.8	27.0	18.3	26.1	18.8
Protolignin preparation percent yield[c]	20%	21%	23%	16%	22%	14%

[a] Co-eluting monosaccharides
[b] Preparation mass relative to lignin in meal
[c] Preparation yield (corrected for carbohydrates, wt% of total raw lignin)

Table 3. Hydroxyl group functional abundance of hardwood protolignin preparations

Biomass / Hydroxyl group[a]	Maple		Sweetgum		Nitens	
	MWLp	CELp	MWLp	CELp	MWLp	CELp
Aliphatic OH	6.8	7.3	7.2	7.5	7.6	8.5
C3 & C5 bound phenolic OH	0.8	0.7	0.9	0.5	0.7	0.6
C3 or C5 unbound phenolic OH	0.7	0.5	0.7	0.5	0.6	0.4
C3 & C5 unbound phenolic OH	0.1	0.1	0.2	0.1	0.1	0.0
Total phenolic OH	1.6	1.3	1.8	1.1	1.4	1.0

[a] mmol/g total lignin in solid

Table 4. Native substructure quantities from hardwood protolignin preparations (#/100 Ar)

Substructure		Maple		Sweetgum		Nitens	
		MWLp	CELp	MWLp	CELp	MWLp	CELp
Non-condensed S/G		1.4	2.1	2.0	2.7	3.0	3.7
β-O-4’		38	50	39	47	40	50
β-O-4’ with Cα carbonyl		2	1	2	1	2	1
Total β-O-4’		40	51	41	48	42	51
β-β’		4	3	5	4	4	3
β-5’		4	3	3	3	2	2
Spirodienone		1	2	2	2	2	2
Cγ acyl		2	1	1	1	1	1
LCC	Uronosil ester	2	1	1	1	2	1
	Benzyl ether, 1° OH	1	*nd*	*nd*	*nd*	*nd*	*nd*
	Benzyl ether, 2° OH[a]	*nd*[b]	*nq*	*nd*	*nd*	*nd*	*nd*
	Phenyl glycoside	*nq*[c]	*nq*	*nq*	*nq*	*nq*	*nq*
	Total LCC	3	1	1	1	2	1

[a] Semi-quantitative
[b] Not detected
[c] Not quantified (<0.5/100 Ar)

Table 5. Autohydrolyzed hardwood chemical composition, autohydrolyzate carbohydrate and lignin concentrations (g/100 g o.d. raw biomass)

Chemical component	Maple		Sweetgum		Nitens	
	Solid	Liquid	Solid	Liquid	Solid	Liquid
Solid yield[a]	69.2	-	66.5	-	68.8	-
Glucan	40.1	1.0	37.5	2.3	43.3	0.6
Xylan	3.0	9.2	3.7	10.4	2.5	10.7
Galactan	*nd*[d]	0.6	*nd*	0.5	*nd*	1.1
Arabinan + Mannan[b]	1.5	1.5	1.3	0.5	0.6	0.7
Total carbohydrate	44.6	12.3	42.5	13.7	46.4	13.1
Acid-insoluble lignin	20.4	-	20.3	-	18.7	-
Acid-soluble lignin	0.9	-	1.1	-	1.1	-
Total lignin	21.3	3.9[e]	21.4	3.8[e]	19.8	4.7[e]
Ash	*nd*	-	0.1	-	*nd*	-
Uncharacterized[c]	4.8%	-	3.8%	-	3.8%	-

[a] Mass of extensively-washed autohydrolyzed solids
[b] Co-eluting monosaccharides
[c] wt% of extensively-washed autohydrolyzed solids
[d] Not detected
[e] Obtained by back-calculation

Table 6. Autohydrolyzate-soluble lignin (XADL) adsorption yields

Biomass	Sugarcane bagasse	Wheat straw	Switchgrass	Maple	Sweetgum	Nitens
Soluble lignin abundance[a]	6.5	6.7	7.6	3.9	3.8	4.7
Reduced autohydrolyzate UV response	60	98	104	42	43	50
Reduced post-adsorption autohydrolyzate UV response	5	2	6	3	1	1
Soluble lignin adsorption yield[b]	92%	98%	94%	93%	98%	99%

λ = 280 nm
[a] From Table 7
[b] % of absorbance lost

Table 7. Chemical composition of hardwood biorefinery lignin preparations, AHCELp preparation yields

Biomass / Chemical component	Maple		Sweetgum		Nitens	
	XADL	AHCELp	XADL	AHCELp	XADL	AHCELp
Glucan	2.4	2.4	3.5	2.5	1.6	1.9
Xylan	39.1	1.2	47.2	1.3	33.5	1.2
Galactan	0.5	1.3	0.7	1.1	0.4	1.7
Arabinan + Mannan[a]	6.3	2.4	5.7	1.6	4.0	2.5
Total carbohydrate	48.2	7.3	57.1	6.5	39.6	7.3
Acid-insoluble lignin	7.8	91.5	7.1	90.1	7.8	92.4
Acid-soluble lignin	26.6	0.2	21.8	0.3	40.1	0.2
Total lignin	34.4	91.7	29.0	90.4	47.9	92.6
Ash	*nd*[d]	*nd*	*nd*	*nd*	*nd*	*nd*
Uncharacterized	17.4	1.0	13.9	3.1	12.5	0.1
Gravimetric solid yield[b]	-	27.3	-	32.6	-	45.3
Preparation percent yield[c]	-	25%	-	30%	-	42%

[a] Co-eluting monosaccharides
[b] Preparation mass (wt% of autohydrolyzed lignin)
[c] Preparation yield (corrected for carbohydrates, wt% of autohydrolyzed lignin)
[d] Not detected

Table 8. Hydroxyl group functional abundance of hardwood biorefinery lignin preparations

Biomass / Hydroxyl group[a]	Maple		Sweetgum		Nitens	
	XADL	AHCELp	XADL	AHCELp	XADL	AHCELp
Aliphatic OH	27.9	4.0	28.5	4.1	13.8	3.5
C3 & C5 bound phenolic OH	3.7	2.2	3.7	2.2	2.7	2.9
C3 or C5 unbound phenolic OH	1.4	0.9	1.2	0.7	0.6	0.7
C3 & C5 unbound phenolic OH	0.2	0.1	0.4	0.1	0.1	0.1
Total phenolic OH	5.3	3.2	5.3	3.0	3.4	3.7

[a] mmol/g total lignin in solid

Table 9. Estimated molecular mass of hardwood biorefinery lignin preparations

Biomass / Molecular mass[a]	Maple		Sweetgum		Nitens	
	XADL	AHCELp	XADL	AHCELp	XADL	AHCELp
Mw_n[b]	400	1700	400	1600	800	1500
Mw_w[c]	800	7900	800	8700	1100	9200
PDI[d]	2.0	4.6	2.0	5.4	1.3	6.1

[a] g/mol
[b] Number-averaged molecular mass
[c] Weight-averaged molecular mass
[d] Polydispersity index (Mw_w/Mw_n)

Table 10. Native substructure quantities for hardwood biorefinery lignin preparations (#/100 Ar)

Substructure		Maple		Sweetgum		Nitens	
		XADL	AHCELp	XADL	AHCELp	XADL	AHCELp
Non-condensed S/G		4.9	2.6	6.8	3.4	6.6	4.6
β-O-4		33	21	42	23	43	17
β-O-4’ with Cα carbonyl		1	1	1	1	3	1
Total β-O-4’		34	22	43	24	46	18
β-β’		3	3	4	3	2	2
β-5’		2	4	2	3	2	2
Cγ acyl		1	*nd*[c]	1	*nd*	*nq*	*nd*
LCC	Uronosil ester	8	1	8	*nq*	6	*nq*
	Benzyl ether, 1° OH	2	1	2	1	1	1
	Benzyl ether, 2° OH[a]	*nq*[b]	*nd*	*nq*	*nq*	*nq*	*nd*
	Phenyl glycoside	*nq*	*nd*	1	*nd*	*nq*	*nd*
	Total LCC	10	2	11	1	7	1

[a] Semi-quantitative
[b] Not quantified (<0.5/100 Ar)
[c] Not detected

Table 11. Biorefinery mass balance of maple lignin properties (X / 100g o.d. raw biomass)

Lignin property	Raw			Biorefinery			
	Wood	MWLp	CELp	Autohydrolyzate	XADL	Cellulolytic hydrolysis residue	AHCELp
Total raw lignin representativeness	100%	-	-	15%	-	85%	-
Lignin preparation representativeness	-	20%	21%	-	93%	-	25%
Total lignin, g	25.2	5.0	5.3	3.9	3.6	21.3	5.3
Phenolic hydroxyl, mmol	-	8	7	-	19	-	17
Non-condensed S units, mmol	-	15	18	-	15	-	19
Non-condensed G units, mmol	-	10	9	-	3	-	7
Total β-O-4', #	-	10	14	-	6	-	6
Total LCC, #	-	1	<1	-	2	-	1

Table 12. Biorefinery mass balance of sweetgum lignin properties (X / 100g o.d. raw biomass)

Lignin property	Raw			Biorefinery			
	Wood	MWLp	CELp	Autohydrolyzate	XADL	Cellulolytic hydrolysis residue	AHCELp
Total raw lignin representativeness	100%	-	-	15%	-	85%	-
Lignin preparation representativeness	-	23%	16%	-	98%	-	30%
Total lignin, g	25.2	5.8	4.0	3.8	3.7	21.4	6.4
Phenolic hydroxyl, mmol	-	10	4	-	20	-	19
Non-condensed S units, mmol	-	19	15	-	16	-	25
Non-condensed G units, mmol	-	10	5	-	2	-	7
Total β-O-4', #	-	12	9	-	8	-	8
Total LCC, #	-	<1	<1	-	2	-	<1

Table 13. Biorefinery mass balance of nitens lignin properties (X / 100g o.d. raw biomass)

Lignin property	Raw			Biorefinery			
	Wood	MWLp	CELp	Autohydrolyzate	XADL	Cellulolytic hydrolysis residue	AHCELp
Total raw lignin representativeness	100%	-	-	19%	-	81%	-
Lignin preparation representativeness	-	22%	14%	-	99%	-	42%
Total lignin, g	24.5	5.4	3.4	4.7	4.7	19.8	8.3
Phenolic hydroxyl, mmol	-	8	3	-	16	-	31
Non-condensed S units, mmol	-	20	13	-	20	-	34
Non-condensed G units, mmol	-	7	4	-	3	-	7
Total β-O-4', #	-	11	9	-	11	-	7
Total LCC, #	-	1	<1	-	2	-	<1

Table 14. Chemical composition and preparation yields of non-wood protolignin preparations

Biomass / Chemical component	Sugarcane bagasse		Wheat straw		Switchgrass	
	MWLp	CELp	MWLp	CELp	MWLp	CELp
Glucan	1.0	1.7	0.6	1.8	2.3	1.0
Xylan	11.6	8.0	0.9	7.3	2.8	2.4
Galactan	0.4	0.8	1.1	0.6	0.8	1.3
Arabinan + Mannan[a]	1.9	2.9	0.8	2.2	1.5	1.5
Total carbohydrate	14.9	13.4	3.4	11.9	7.4	6.2
Acid-insoluble lignin	68.9	81.2	94.9	84.2	81.5	89.2
Acid-soluble lignin	0.4	0.3	0.1	0.0	0.0	0.1
Total lignin	69.3	81.5	95.0	84.2	81.5	89.3
Uncharacterized	15.8	5.1	1.6	3.9	11.1	4.5
Gravimetric solid yield[b]	36.1	28.3	17.3	25.0	24.8	33.3
Protolignin preparation percent yield[c]	25%	23%	17%	21%	20%	30%

[a] Co-eluting monosaccharides
[b] Preparation mass relative to lignin in meal (g preparation/100 g total raw lignin)
[c] Preparation yield (corrected for carbohydrates, wt% of total raw lignin)

Table 15. Hydroxyl group functional abundance of non-wood protolignin preparations

Biomass / Hydroxyl group[a]	Sugarcane bagasse		Wheat straw		Switchgrass	
	MWLp	CELp	MWLp	CELp	MWLp	CELp
Aliphatic OH	6.7	6.4	4.7	6.7	6.3	6.8
C3 & C5 bound phenolic OH	0.8	0.4	0.6	0.5	0.5	0.3
C3 or C5 unbound phenolic OH	0.7	0.5	0.7	0.7	0.7	0.6
C3 & C5 unbound phenolic OH	1.5	1.4	0.5	0.6	0.5	0.6
Total phenolic OH	3.0	2.3	1.8	1.8	1.7	1.5

[a] mmol/g total lignin in solid

Table 16. Native substructure quantities from non-wood protolignin preparations (#/100 Ar)

Substructure		Sugarcane bagasse		Wheat straw		Switchgrass	
		MWLp	CELp	MWLp	CELp	MWLp	CELp
Non-condensed S/G		1.2	1.1	0.7	0.7	0.5	0.4
Non-condensed H/G		1.5	1.8	0.3	0.4	0.4	0.5
para-Coumaric acid		20	29	6	9	11	14
Ferulic acid		2	3	2	4	3	5
β-O-4’		25	23	30	29	26	27
β-O-4’ with Cα carbonyl		1	*nq*	1	1	1	1
Total β-O-4’		26	23	31	30	27	28
β-β’		*nq*[b]	*nq*	1	1	1	1
β-5’		3	3	5	4	5	4
Spirodienone		*nq*	1	1	1	1	1
Dibenzodioxicin		1	6	*nd*	2	1	7
Cγ acyl		1	1	2	2	2	1
LCC	Uronosil ester	8	8	3	3	6	5
	Benzyl ether, 1° OH	1	*nd*	1	*nd*	1	*nd*
	Benzyl ether, 2° OH[a]	*nd*[c]	*nd*	*nq*	*nq*	*nq*	*nd*
	Phenyl glycoside	2	1	1	1	3	1
	Total LCC	11	9	5	4	10	6

[a] Semi-quantitative
[b] Not quantified (<0.5/100 Ar)
[c] Not detected

Table 17. Autohydrolyzed non-wood chemical composition, autohydrolyzate carbohydrate and lignin concentrations (g/100 g o.d. raw biomass)

Chemical component	Sugarcane bagasse		Wheat straw		Switchgrass	
	Solid	Liquid	Solid	Liquid	Solid	Liquid
Solid yield[a]	59.0	-	54.9	-	53.1	-
Glucan	37.3	1.9	30.9	3.3	30.0	2.1
Xylan	3.5	10.4	4.3	9.3	5.1	14.0
Galactan	*nd*[d]	0.7	*nd*	1.0	*nd*	0.8
Arabinan + Mannan[b]	0.4	1.2	0.3	1.6	0.3	2.0
Total carbohydrate	41.2	14.2	35.5	15.2	35.4	18.9
Acid-insoluble lignin	13.2	-	14.9	-	13.1	-
Acid-soluble lignin	0.8	-	0.7	-	0.6	-
Total lignin	14.0	6.5[e]	15.6	6.7[e]	13.7	7.6[e]
Ash	1.7	-	0.7	-	0.4	-
Uncharacterized[c]	3.6%	-	5.6%	-	6.7%	-

[a] Mass of extensively-washed autohydrolyzed solids
[b] Co-eluting monosaccharides
[c] wt% of extensively-washed autohydrolyzed solids
[d] Not detected
[e] Obtained by back-calculation

Table 18. Chemical composition of non-wood biorefinery lignin preparations, AHCELp preparation yields

Biomass / Chemical component	Sugarcane bagasse		Wheat straw		Switchgrass	
	XADL	AHCELp	XADL	AHCELp	XADL	AHCELp
Glucan	5.8	0.7	4.5	0.7	4.2	0.9
Xylan	38.6	2.3	27.8	2.5	37.9	3.8
Galactan	0.2	0.9	0.7	0.9	0.3	1.0
Arabinan + Mannan[a]	3.4	0.5	3.4	0.5	4.5	0.6
Total carbohydrate	47.9	4.5	36.4	4.6	46.9	6.3
Acid-insoluble lignin	8.7	88.9	10.7	88.7	12.8	85.8
Acid-soluble lignin	15.2	1.1	24.7	0.6	16.9	0.4
Total lignin	23.8	90.0	35.4	89.4	29.7	86.2
Ash	*nd*[d]	0.1	*nd*	0.2	*nd*	*nd*
Uncharacterized	28.2	5.4	28.1	5.8	23.4	7.5
Gravimetric solid yield[b]	-	45.6	-	33.9	-	47.3
Preparation percent yield[c]	-	41%	-	30%	-	41%

[a] Co-eluting monosaccharides
[b] Preparation mass (wt% of autohydrolyzed lignin)
[c] Preparation yield (corrected for carbohydrates, wt% of autohydrolyzed lignin)
[d] Not detected

Table 19. Hydroxyl group functional abundance of non-wood biorefinery lignin preparations

Biomass / Hydroxyl group[a]	Sugarcane bagasse		Wheat straw		Switchgrass	
	XADL	AHCELp	XADL	AHCELp	XADL	AHCELp
Aliphatic OH	41.3	2.8	20.9	3.9	29.4	4.7
C3 & C5 bound phenolic OH	2.6	1.5	1.6	1.3	1.8	1.2
C3 or C5 unbound phenolic OH	2.6	0.8	1.9	1.0	2.3	1.1
C3 & C5 unbound phenolic OH	2.1	1.3	0.8	0.6	1.6	0.6
Total phenolic OH	7.3	3.6	4.3	2.9	5.7	2.9

[a] mmol/g total lignin in solid

Table 20. Estimated molecular mass of non-wood biorefinery lignin preparations

Biomass / Molecular mass[a]	Sugarcane bagasse		Wheat straw		Switchgrass	
	XADL	AHCELp	XADL	AHCELp	XADL	AHCELp
Mw_n[b]	600	1900	900	1500	700	1900
Mw_w[c]	1000	11000	2000	9000	1500	10900
PDI[d]	1.7	5.8	2.2	6.0	2.1	5.7

[a] g/mol
[b] Number-averaged molecular mass
[c] Weight-averaged molecular mass
[d] Polydispersity index (Mw_w/Mw_n)

Table 21. Native substructure quantities for non-wood biorefinery lignin preparations (#/100 Ar)

Substructure		Sugarcane bagasse		Wheat straw		Switchgrass	
		XADL	AHCELp	XADL	AHCELp	XADL	AHCELp
Non-condensed S/G		1.0	2.3	1.0	1.3	0.9	0.7
Non-condensed H/G		7.0	2.4	2.3	0.4	4.5	0.5
para-Coumaric acid		20	18	4	6	22	10
Ferulic acid		18	1	11	4	16	3
β-O-4		13	13	13	16	10	14
β-O-4’ with Cα carbonyl		*nq*[b]	1	*nq*	1	*nq*	1
Total β-O-4’		13	14	13	17	10	15
β-β’		*nq*	*nq*	1	1	*nq*	1
β-5’		1	4	2	6	2	7
Dibenzodioxicin		8	4	2	1	7	4
Cγ acyl		*nq*	*nd*[c]	*nq*	*nq*	*nq*	*nd*
	Uronosil ester	5	6	2	1	6	3
	Benzyl ether, 1° OH	*nq*	*nd*	1	1	*nq*	1
	Benzyl ether, 2° OH[a]	*nq*	*nd*	*nq*	*nd*	*nq*	*nd*
	Phenyl glycoside	4	2	5	1	8	4
	Total LCC	9	8	8	3	14	8

[a] Semi-quantitative
[b] Not quantified (<0.5/100 Ar)
[c] Not detected

Table 22. Biorefinery mass balance of sugarcane bagasse lignin properties (X / 100g o.d. raw biomass)

Lignin property	Raw			Biorefinery			
	Non-wood	MWLp	CELp	Autohydrolyzate	XADL	Cellulolytic hydrolysis residue	AHCELp
Total raw lignin representativeness	100%	-	-	32%	-	68%	-
Lignin preparation representativeness	-	25%	23%	-	92%	-	41%
Total lignin, g	20.5	5.1	4.7	6.5	6.0	13.9	5.7
Phenolic hydroxyl, mmol	-	15	11	-	44	-	21
Non-condensed S units, mmol	-	9	7	-	4	-	13
Non-condensed G units, mmol	-	8	7	-	4	-	6
Non-condensed H units, mmol	-	11	12	-	26	-	13
para-Coumaric acid, #	-	6	8	-	7	-	6
Ferulic acid, #	-	1	1	-	6	-	<1
Total β-O-4', #	-	7	6	-	4	-	4
Dibenzodioxicin, #	-	<1	2	-	3	-	1
Total LCC, #	-	3	2	-	3	-	3

Table 23. Biorefinery mass balance of wheat straw lignin properties (X / 100g o.d. raw biomass)

Lignin property	Raw			Biorefinery			
	Non-wood	MWLp	CELp	Autohydrolyzate	XADL	Cellulolytic hydrolysis residue	AHCELp
Total raw lignin representativeness	100%	-	-	30%	-	70%	-
Lignin preparation representativeness	-	17%	21%	-	98%	-	30%
Total lignin, g	22.3	3.8	4.7	6.7	6.6	15.6	4.7
Phenolic hydroxyl, mmol	-	7	8	-	28	-	14
Non-condensed S units, mmol	-	7	9	-	9	-	13
Non-condensed G units, mmol	-	11	12	-	9	-	10
Non-condensed H units, mmol	-	3	5	-	20	-	4
para-Coumaric acid, #	-	1	2	-	1	-	2
Ferulic acid, #	-	<1	1	-	4	-	1
Total β-O-4', #	-	7	8	-	5	-	4
Dibenzodioxicin, #	-	0	1	-	1	-	<1
Total LCC, #	-	1	1		3	-	1

Table 24. Biorefinery mass balance of switchgrass lignin properties (X / 100g o.d. raw biomass)

Lignin property	Raw			Biorefinery			
	Non-wood	MWLp	CELp	Autohydrolyzate	XADL	Cellulolytic hydrolysis residue	AHCELp
Total raw lignin representativeness	100%	-	-	34%	-	66%	-
Lignin preparation representativeness	-	20%	30%	-	94%	-	41%
Total lignin, g	21.3	4.3	6.4	7.3	6.9	13.7	5.6
Phenolic hydroxyl, mmol	-	7	10	-	39	-	16
Non-condensed S units, mmol	-	6	7	-	5	-	10
Non-condensed G units, mmol	-	13	19	-	6	-	14
Non-condensed H units, mmol	-	5	9	-	27	-	7
para-Coumaric acid, #	-	3	5	-	8	-	3
Ferulic acid, #	-	1	2	-	6	-	1
Total β-O-4', #	-	6	10	-	4	-	5
Dibenzodioxicin, #	-	<1	2	-	3	-	1
Total LCC, #	-	2	2	-	5	-	2

FIGURES

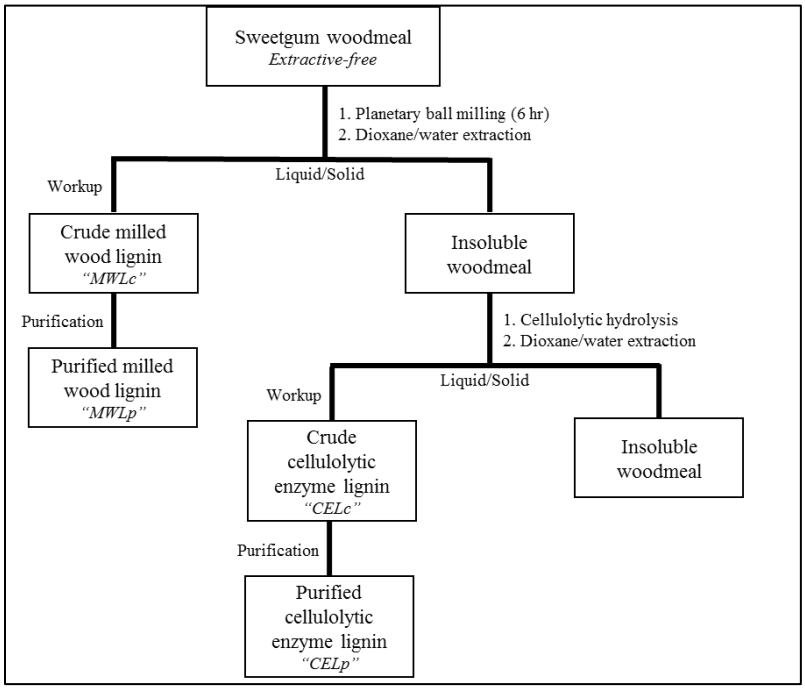

Figure 1. Protolignin preparation procurement protocol

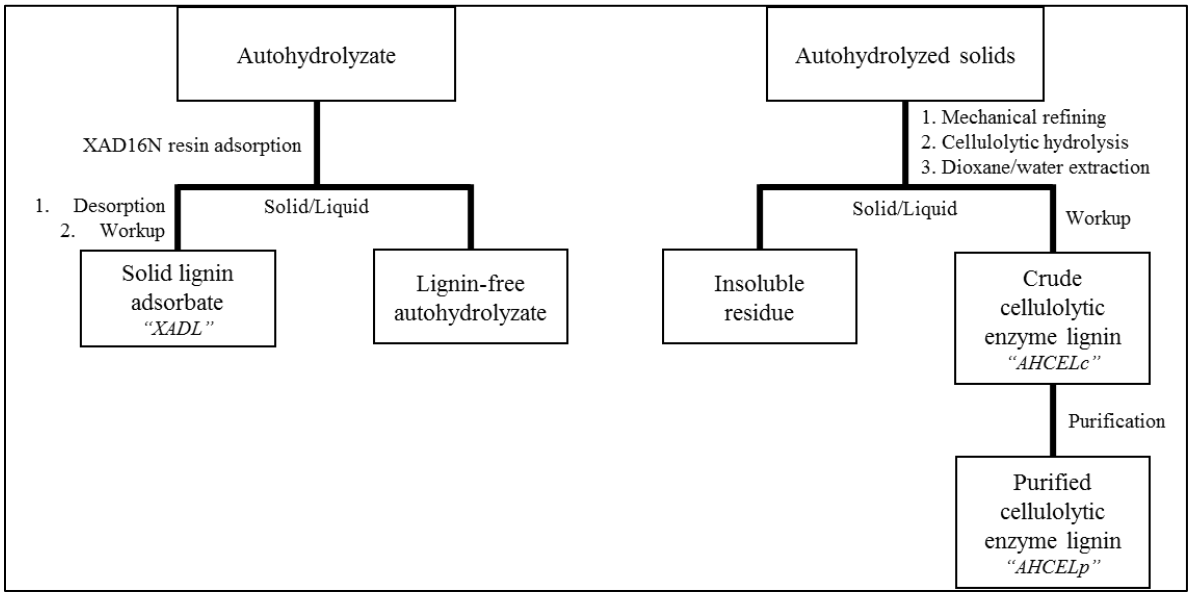

Figure 2. Isolation protocols for autohydrolyzate-soluble and cellulolytic hydrolysis-insoluble lignin

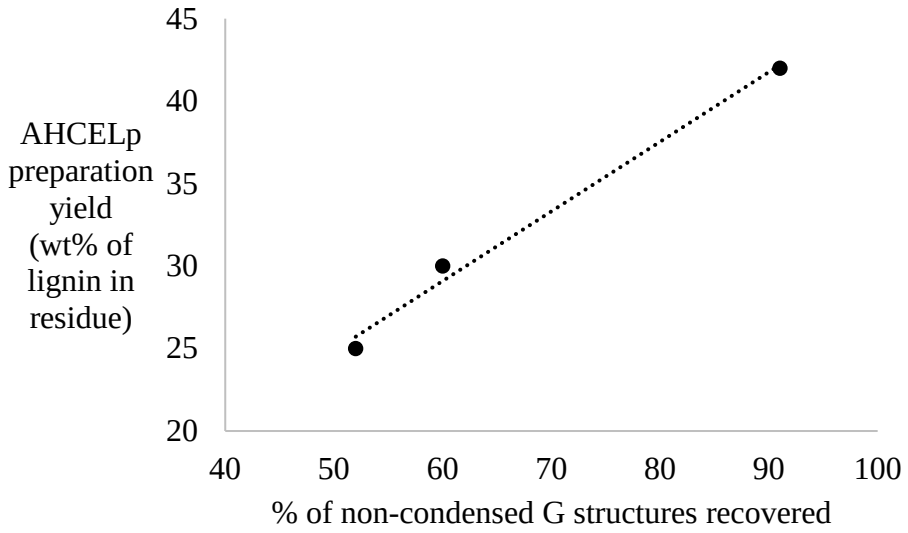

Figure 3. Correlation between hardwood AHCELp preparation yield and recovery of non-condensed G structures across hardwood lignin preparations

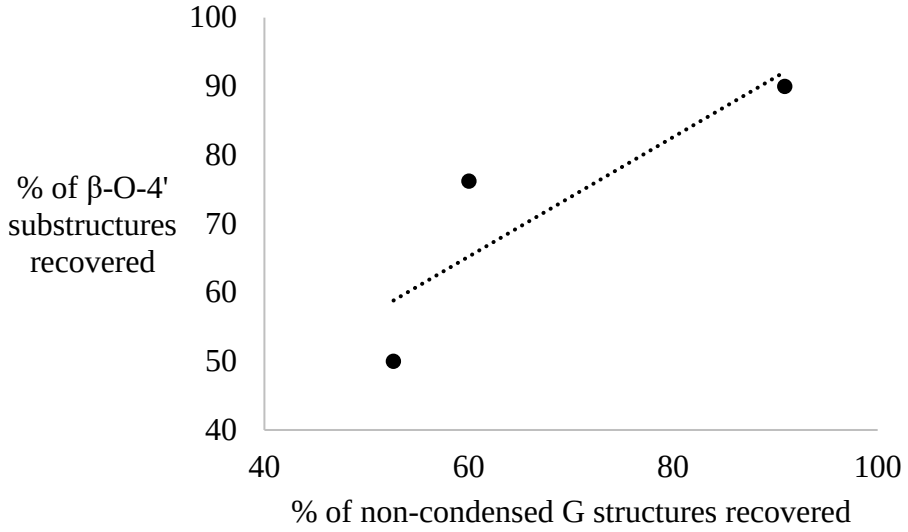

Figure 4. Correlation between recovery of non-condensed G structures and recovery of β-O-4' substructures across hardwood lignin preparations

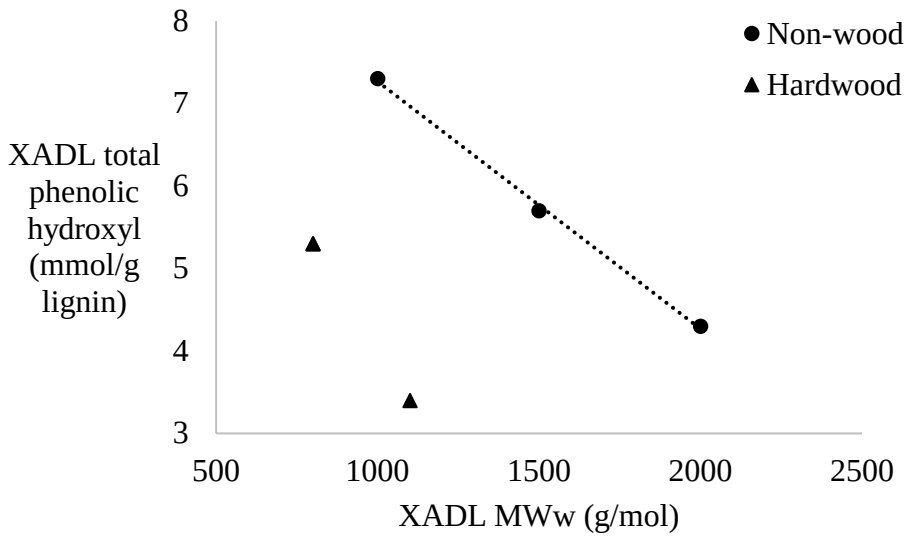

Figure 5. Correlation between non-wood XADL's weight-averaged molecular mass and total phenolic hydroxyl functionality

Chapter 6. Conclusions

An impactful literature review was performed in order to evaluate the state of knowledge around biorefinery pretreatment's effects upon biorefinery lignin properties. It was confirmed from reviewed sources that lignins subjected to more severe pretreatment (acidic or alkaline) are generally less useful for valorization. This was based upon the increased prominence of inter-lignin condensation. Pretreatment dampening to increase lignin's valorizability was prescribed.

Six different biomasses were funneled in identical manner through a biorefinery process at the laboratory scale. The process includes pretreatment using sequential mild autohydrolysis and mechanical refining, followed by low-dosage cellulolytic hydrolysis. The process was found to render strong degrees of saccharification yields, yet different yields were obtained from the botanically diverse feedstocks. The mass balance of all biomass components throughout said process was assembled to identify the exact quantities of carbohydrates and lignin produced by the investigated process upon a single biomass. In effort to relate the range of carbohydrate output to differences in protolignin, a lignocentric approach to carbohydrate generation was adopted. In this, several lignin properties (delignification, degree of condensation, non-wood hydroxycinnamic acid abundance, and more) were compared with saccharification results in attempt to identify the protolignin properties most influential upon carbohydrate output. Finally, it was concluded that hardwood species with higher S/V ratios resulted in greater carbohydrate output, attributable to the minor degree of inter-lignin condensation experienced by such hardwoods.

Comprehensive analysis of biorefinery lignin fractions was previously hindered by a lack of elucidative characterization of the lignin rendered soluble by autohydrolysis pretreatment. This fraction, however minor (~15 wt% of raw hardwood lignin, ~30 wt% of raw non-wood lignin), contains significant information regarding lignin degradation induced by autohydrolysis. An aromatic-selective resin was applied to selectively withdraw the soluble lignin from the autohydrolyzate solution matrix. Laboratory processing was employed to render the originally soluble lignin into a solid lignin preparation, which could then be subjected to a suite of analyses previously un-performable on bulk autohydrolyzate. Characterization of this lignin revealed its molecular weight to be a reflection of its strong degradation. However, the degree of degradation halted prior to generation of monomeric lignin fragments. Instead, the soluble lignin was revealed to be dimeric-oligomeric with solubility assistance from lignin-carbohydrate complex substructures. These findings shed light on a biorefinery lignin fraction that went previously uncharacterized to this extent.

Lignin preparations were procured from a single hardwood's biorefinery lignin fractions as well as from its protolignin. Comprehensive analysis was performed upon all lignin preparations, resulting in specific description of each lignin preparation's hydroxyl functionality, molecular weight, and native substructural abundance. The characterization data was then framed into the biorefinery mass balance, and the recovery and/or re-distribution of different lignin structures and functionalities were tracked. Biorefinery processing was found to induce a strong degree of chemical degradation upon lignin, resulting in a distribution of fragmented lignin molecules locatable both in autohydrolyzate

as well as in the process' resultant residues. Degradation of β-O-4' substructures by acidolysis-type reactions were noted, as was indirect evidence of inter-lignin condensation. The results provide an enhanced description of the lignin produced by the investigated biorefinery process.

To address the effects of varying protolignin sources upon biorefinery lignin properties, six dissimilar biomasses (three hardwoods and three non-woods) were subjected to an identical biorefinery process. Lignin preparations were procured from each respective protolignin as well as both biorefinery lignin fractions. These preparations were subjected to comprehensive characterization along the lines of hydroxyl functionality, molecular weight, and native substructural abundance. Once characterized, all the data was incorporated into the mass balance of lignin for each individual biomass. Comparison between lignin responses revealed that hardwood lignin's with low S/G ratios are more prone to both inter-lignin condensation and acidolysis, resulting in a more degraded soluble lignin fraction and a less valorizable insoluble lignin fraction. No significant identifiers were uncovered that could definitively describe non-wood lignin's response to the biorefinery process. However, noted occurrence of acidolysis in conjunction with soluble lignin containing unique substructure profile suggests that acidolysis leads to solubility of substructure-specific soluble lignin fragments.

Chapter 7. Future work

7.1. Recovery of protolignin lost during crude protolignin preparation purification

The applied purification protocol resulted in loss of lignin preparation yield. This effect was most pronounced in crude milled wood lignins ("MWLc") which saw its preparation yields drop from ~35-45% down to ~15-20% following purification. Other works have concluded that this is yield loss is attributable to unintended removal of lignin molecules bearing LCC substructures. However, these works did not feature selectivity when recovered the lost lignin, therefore the samples analyzed were an amalgam of lignin with chemically attached carbohydrates as well as un-bound carbohydrates.

To improve characterization of the lost LCC-enriched lignins, a lignin-selective adsorptive resin was applied to the LCC-containing solution. This ensured procurement of a lignin preparation free of un-bound carbohydrates, improving certainty surrounding discussion of the lignin's LCC substructures.

Preliminary results revealed 82-95% recovered of this portion of milled wood lignin, translating to recovery of 11-20% of the LCC-rich lignin lost during purification. Characterization of the lignin resulted in identification of LCC abundance, in agreement with previous literature. A range of total LCC substructures was found to be present in the three tested hardwoods (4-6 total LCC/100 Ar). Significant LCC enrichment was identified in the investigated non-wood lignin, which possessed 18 total LCC/100 Ar. Future work should involve procurement of enough of this material to enable further characterization, specifically: chemical composition, approximate molecular mass, and hydroxyl functionality.

7.2. Improved compositional analysis protocol for lignocellulosic solids high in acid-soluble lignin

Subpar mass balances were consistently delivered for all autohydrolyzate-soluble lignin adsorbates, regardless of biomass origin. This is believed to be in part due to the assumptions made during spectrophotometric quantification of acid-soluble lignin, found in the acid digestate produced during compositional analysis. For most biomasses, the erroneous assumptions made during acid-soluble lignin quantification do not overtly affect the analysis' mass balance. This is due to acid-soluble lignin generally representing a minor portion of the total biomass. However, such assumptions can be detrimental to characterization of solids containing higher orders of acid soluble lignin (e.g. non-wood adsorbate: 15 wt%, hardwood adsorbate: 27 wt% acid-soluble lignin).

To overcome the error introduced through the conventional assumptions made during compositional analysis, a modified protocol for characterization of acid-soluble lignin in acid digestate has begun development. This method involves acquisition of a UV response which is indiscriminate between lignin and furan aldehydes, measured in preliminary work at 280 nm. The non-specific UV response was then whittled down utilizing independent HPLC quantitation of non-lignin UV absorbing solutes. Specifically, furfural and 5-hydroxymethylfurfural were measured chromatographically. Using acid digestate concentration and molar absorptivity at the chosen wavelength, furan aldehyde contribute to UV response was subtracted and the remainder attributed to lignin. This approach was nearly successful, but resulted in slightly erroneous hardwood compositional analysis (105% mass balance) and highly erroneous non-wood compositional analysis (117% mass balance).

Future work should seek to quantify additional potential UV active solutes. For hardwood, the slight overestimation (~5%) is likely due to the UV response of hemicellulosic uronic acid moieties. If properly accounted for, it is feasible to assume that mass balance could be lowered below 100% into the realm of real results. For non-woods, in addition to uronic acids, independent quantification of hydroxycinnamic acids will also lower the absorbance attributable to acid-soluble lignin, and potentially produce non-erroneous mass balances using this modified method.

7.3. XAD16N adsorption kinetics around autohydrolyzates of different botanical and process origins

A more advanced study of how Amberlite® XAD16N interacts with autohydrolyzate solutes should be conducted. This prescription comes from observation of a range of adsorptions from different hardwood autohydrolyzates at identical solution-to-resin ratios. Specifically, only 93% adsorption was quantified for maple autohydrolyzate, compared to 98% for sweetgum and 99% for nitens. To determine if this difference is due to chemical differences between the unadsorbed maple solutes compared to the others, it will be important to create extensive adsorption curves to understand the adsorption limit with respect to individual biomass. If compounds are confirmed as un-adsorbable, then it is likely due to overwhelming lignin hydrophilicity granted by prolonged carbohydrate chains attached to lignin through LCC substructures.

A similar study could be performed upon other lignin-containing solutions, e.g. pulping liquor or bleaching effluent. Specific understanding of what lignin solutes engage

best with XAD will further application of this resin into biorefinery or even pulping processes. It is feasible to imagine a fractionation process tuned to different soluble lignin-containing solutions utilizing different XAD products in an array of unit operations which takes highly selective cuts of soluble lignin away from bulk solution.