Production and characterization of stabilized lignin

Thèse N° 9236

Présentée le 25 janvier 2019

à la Faculté des sciences de base
Laboratoire des procédés durables et catalytiques
Programme doctoral en chimie et génie chimique

pour l'obtention du grade de Docteur ès Sciences

par

MASOUD TALEBI AMIRI

Acceptée sur proposition du jury

Prof. P. J. Dyson, président du jury
Prof. J. Luterbacher, directeur de thèse
Prof. F. Vogel, rapporteur
Prof. K. Barta, rapporteuse
Dr J. Van Herle, rapporteur

2019

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Acknowledgements

First and foremost, I would like to thank my thesis advisor, Professor Jeremy Luterbacher, for giving me the chance to work in the field that I am very interested in, and also extending my doctoral studies for three months to finalize my work and obtain my Ph.D. degree.

On this journey, there were many people helped me to reach my goals and I am very grateful to each one of them. I start with Professor Fathollah Farhadi at the Sharif University of Technology, who taught me the basics of process engineering and raised my interest in this field by sharing his vast knowledge with his students, including me. His support had a significant effect on helping me to continue my academic path until now. I want to extend my thanks to Professor Lioubov Kiwi, from whom I learnt a lot, both in my academic and personal life. I was lucky to have her support until today which had an undeniable impact on successfully finishing my doctoral studies. In her courses and under her supervision, I learnt the fundamentals of catalysis in chemical engineering and moreover, how to do research properly. A crucial point that I did not realize its importance until I got more involved in academic research during my doctoral studies and observed that how easily the basics of doing proper research can be ignored by many researchers.

I am very grateful to Professor François Maréchal who was my supervisor for my master project and also my mentor during my doctoral studies. Apart from having the chance to learn from his great knowledge, I enjoyed every single conversation with him that also intellectually affected me. His intelligence, smartness and continuous support always kept me sure that I could overcome any challenge that I would face during my doctoral studies. Without his advice and support, I would never be able to finish this part of my academic path successfully.

I was lucky to have the chance to meet great professors at EPFL. Professor Vassily Hatzimanikatis who was always available for helping me with his valuable advice. Being the director of the doctoral programme and having a hectic schedule, he was always very welcoming whenever I was in need of his help. I will always be grateful to him for helping me to go through the toughest moments during my doctoral studies. I would like also to mention Professor Kevin Sivula, whom I

never had the chance to work with. However, I always enjoyed his company. His positive energy was a source of motivation for me and I always admire him.

Working in a field that I was very enthusiastic about was always satisfying. However, I went through very tough moments during the last two years of my doctoral studies that was the source of extremely high physical and mental pressure on me. Even though I am standing at a point that I made it to the end, this would not be possible without the support of all my colleagues, friends, and collaborators at EPFL. Therefore, I would like to thank all of them who helped or supported me in any way, and would like to mention some of my friends who always encouraged me on this journey: Ahmed H, Bartosz R, Benjamin LM, Jack D, Mary J, Marko S, Nadim A, Philip M, Rory K, Stefanos G, Ydna Q, and last but not least Alexandra who always stood by me and supported me unconditionally.

I finish the acknowledgments by expressing my sincerest thanks to my parents and family for helping to be who I am today. I am especially grateful to my brother, Omid, who always supported me and helped me to make important and critical decisions.

Abstract

Lignocellulosic biomass is a sustainable source of renewable carbon, and it is the most abundant form of terrestrial biomass. The three main constituents of lignocellulosic biomass are cellulose, hemicellulose, and lignin. The monomers from these biopolymers can be valuable feedstocks for a future sustainable chemical industry, and they include glucose from cellulose, predominantly xylose from hemicellulose, and aromatic molecules from lignin. Even though both glucose and xylose are already established feedstocks for further upgrading in biorefineries, the upgrading of lignin to higher value-added chemicals has not achieved the same success, despite the tremendous need for renewable aromatic molecules. This is mainly due to the challenges associated with the methods for extraction and depolymerization of this natural biopolymer into its constituent monomers.

Developing methods for the high-yield depolymerization of lignin can dramatically increase the efficiency and profitability of biorefineries. Moreover, such methods should be compatible with existing biorefineries so they can be implemented industrially. This translates to providing a fractionation process that can efficiently separate the main three biopolymers of biomass with the least possible alteration to the chemical structure of each fraction.

The main limiting factor in producing high-quality lignin is the inter-unit carbon-carbon linkages in the chemical structure of lignin, and also the formation of these linkages during the extraction process, known as repolymerization or condensation. In the first part of my thesis, I study the possibility of avoiding the condensation of lignin by preventing the formation of C-C linkages. This is achieved by adding formaldehyde, as a protection group, to the pretreatment of lignocellulosic biomass. This leads to the reaction of formaldehyde with an α,γ-diol structure on the side chain of extracted lignin to form a 1,3-dioxane structure, consequently stabilizing the lignin. The results show a near theoretical yield of monomers from the stabilized lignin after depolymerization by hydrogenolysis. This method was then further expanded by screening other protecting groups in order to develop a comprehensive picture of the chemistry involved in this process, and to optimize the process notably to be able to yield a narrower product distribution.

Even though this method provides high-quality lignin, there is a need for developing a fractionation process for large-scale production of isolated lignin. This can facilitate the implementation of the process in biorefineries and also further studies on developing new upgrading pathways. The second part of this thesis presents a protocol that fractionates lignocellulosic biomass into pure parts of three main constituents in large laboratory-scale with high extraction yields. The results show extraction yields higher than 95% for each biopolymer. The highly digestible cellulose fraction and high-quality lignin can be depolymerized into their constituent monomers with conventional methods. However, the hemicellulose fraction is obtained as functionalized sugars due to its reaction with aldehyde during pretreatment.

The third part of my work presents a non-destructive method for the prediction of the potential yield of monomers from the depolymerization of isolated lignin. Currently, the depolymerization of the aldehyde-stabilized lignin is done by hydrogenolysis over noble metal catalysts. Even though this destructive method provides valuable information on the amount of produced monomers, it does not provide any information on the chemical structure of the isolated lignin. Furthermore, the product distribution and overall monomer yields are highly dependent on process conditions such as reaction temperature, reaction time, and the catalyst. A novel 2D heteronuclear single quantum coherence nuclear magnetic resonance (2D-HSQC NMR) spectroscopy technique can be used for simultaneous identification and quantification of chemical groups in the structure of lignin. Standard 2D-HSQC NMR spectroscopy does not provide accurate quantification for polymers and oligomers such as lignin due to errors that are principally caused by differences in relaxation times for different parts of the oligomer chain. However, these errors can be avoided by extrapolation to time-zero ^{13}C HSQC ($HSQC_0$) for series of spectrums, acquired with different repetition times. The prediction results show agreement within a few percentage points with experimental results determined by gas chromatography (GC) after hydrogenolysis.

In summary, this thesis presents research on the valorization of lignin from lignocellulosic biomass by introducing a novel pretreatment method for preserving the lignin which can still be depolymerized at theoretical yields. The stabilization chemistry that is involved is also discussed in depth. This discussion is followed by the presentation of a detailed protocol for fractionation of lignocellulosic biomass that is compatible with biorefineries and provides high purity fractions of the three major polymers in biomass. Finally, a non-destructive quantification analysis has been developed for prediction of the lignin’s upgradeability which does not have the drawbacks of conventional depolymerization techniques.

Keywords

Lignocellulosic biomass, lignin, carbohydrates, depolymerization, fractionation, aldehyde, heteronuclear single quantum coherence nuclear magnetic resonance spectroscopy, biorefinery, biofuel, bioproduct.

Résumé

La lignocellulose est une source renouvelable de carbone et la source terrestre la plus abondante de biomasses. Les trois principaux constituants de la lignocellulose sont la cellulose, l'hémicellulose et la lignine. Les monomères de ces bio-polymères sont de précieuse matières premières pour l'industrie chimique. Les monomères pouvant être extrait sont le glucose à partir de la cellulose, la xylose à partir de l'hémicellulose et certaines molécules aromatiques à partir de la lignine. Le glucose et la xylose ont d'ores et déjà été établis comme matières premières dans les bioraffineries. Cependant, la lignine reste un élément difficile à obtenir, malgré le fait qu'elle soit une candidate importante pour la production de molécules aromatiques renouvelables. Cette limitation est principalement due à l'étape d'extraction et de dépolymérisation de la lignine.

Le développement de méthodes permettant de dépolymériser la lignine tout en conservant un rendement important pourrait grandement améliorer la production et la profitabilité des bioraffineries. Toutefois, ces procédés devraient pouvoir être facilement utilisable à l'échelle industrielle. En d'autres termes, le procédé de fractionnement idéal devrait pouvoir séparer les trois principaux biopolymers de la biomasse sans altérer leur intégrité chimique ou leur potentiel de valorization.

La principale limitation à la production de monomères issus de la lignine est la présence de liaisons carbone-carbone composant la structure chimique de cette dernière mais aussi la formation additionelle de ce type de liaison durant le procédé d'extraction. Ce phénomène étant connut sous le nom de repolymerization ou réaction de condensation. Dans la première partie de ma thèse, j'étudie ce problème de repolymerization. Le but a été de trouver un moyen d'empêcher la formation de ces

liaisons C—C. Cela a été possible en utilisant du formaldéhyde en tant que groupe protecteur lors du prétraitement de la lignocellulose. Le formaldéhyde réagit avec le groupe diol α,γ présent dans la structure des monomères de la lignine, pour former un groupe 1,3 dioxane qui permet ainsi de stabiliser la lignine. Les résultats ont montré qu'il était capable d'obtenir ces monomères avec un rendement proche du maximum théorique, après la déplomyerization et l'hydrogénation de la lignine isolée. Il a ensuite été possible de tester différents aldéhydes pour ainsi mieux comprendre les mécanismes chimiques impliqué dans ce procédé. Ces résultats nous ont permis d'optimiser ce procédée afin d'obtenir une distribution de produits la plus étroite possible.

Bien que cette méthode permette de produire de la lignine de haute qualité, il reste toutefois indispensable de développer un précéder de fractionnement pouvant être appliqué à une échelle de production bien plus importante et pouvant produire des fractions purifiées des trois composants de la biomasse. La deuxième partie de ma thèse présente un protocole qui permet de fractionner la lignocellulose en ces trois principaux constituants avec une grande pureté et à une échelle de laboratoire conséquente tout en conservant des rendements d'extraction élevés. Les résultats ont montré qu'il était possible d'obtenir ces biopolymer avec un rendement de 95%. La cellulose et la lignine peuvent être dépolymérisée, et l'hémicellulose est isolée sous forme de sucres fonctionalisés à cause de sa réaction avec l'aldéhyde présente durant le prétraitement.

La troisième partie de ma thèse présente une méthode non destructive pour la prédiction du rendement de dépolymérisation des monomères de la lignine. A ce jour, la dépolymérisation de la lignine stabilisée par l'aldéhyde et la mesure des rendements de monomères est effectuée grâce à une réaction d'hydrogénolyse catalysée par des métaux nobles. Bien que cette méthode destructive fournisse des informations importantes sur la quantité produite de monomères, elle ne permet pas d'obtenir des informations quant à la structure chimique de la lignine extraite. De plus, le rendement ainsi

que la distribution des différents types de monomères dépendent grandement des conditions de réactions tels que la température, le temps de réaction et du catalyseur. Il est possible d'identifier et de quantifié les groupements chimiques présent dans la structure de la lignine grâce à la technique de résonnance magnétique nucléaire (RMN) nommée 2D-HSQC-NMR. D'ordinaire, cette technique ne permet pas d'obtenir des informations quantitatives pour des systèmes polymeriques ou oligomériques tel que la lignine principalement à cause des différents temps de relaxation des groupements chimiques au centre et aux extrémités des oligomères. Cependant ces différences peuvent être corrigée par extrapolation au temps zéro d'une séries de spectre obtenue grâce à la RMN du ^{13}C à différents temps de répétition, rendant les mesures RMN quantitatives. Ces prédictions issues de la RMN se montrée en accord proche avec les données expérimentales obtenues grâce à la chromatographie en phase gazeuse qui suivait l'hydrogénolyse.

En conclusion, cette thèse présente les résultats sur l'étude de la valorisation de la lignine extraite de la lignocellulose. De nouvelles méthodes de prétraitements ont été développées permettant ainsi d'obtenir des produits de haute qualité. Une étude approfondie des mécanismes chimiques impliquée dans cette méthode a été menée. Il a aussi été possible d'établir un protocole pour le fractionnement de la lignocellulose, produisant des fractions de biomasse purifiées compatible avec les techniques utilisées en bioraffinerie. Finalement, une technique non destructive de quantification et de prédiction de la qualité de la lignine a été développée.

Mots-clés

Lignocellulose, Lignine, glucides, dépolymérization, fractionnement, aldéhyde, HSQC, bioraffinerie, biocarburant, bioproduit.

Content

Acknowledgements i

Abstract iii

Keywords v

Résumé vi

Mots-clés viii

Content ix

List of Figures 12

List of Tables 14

List of Equations 15

List of Abbreviations 17

Chapter 1 Introduction 21

1.1 Lignocellulosic biomass as a renewable energy source 21

1.2 Composition of lignocellulosic biomass 26

1.2.1 Sugars 29

1.2.2 Lignin 32

1.3 Utilization of biomass for production of biofuels and bioproducts 36

1.4 Lignocellulosic biomass pretreatment 44

1.4.1 Fractionation based on carbohydrate conversion 45

1.4.1.1 Acid catalyzed carbohydrate conversion 48

1.4.1.2 Enzymatic hydrolysis 50

1.4.1.3 Thermal carbohydrate conversion: pyrolysis 51

1.4.2 Fractionation based on delignification 52

1.4.2.1 Acid delignification methods 54

1.4.2.2 Alkaline delignification methods 56

1.4.2.3 Other delignification techniques 60

1.4.2.4 Reductive Catalytic Fractionation (RCF) 62

1.5 Lignin chemistry 64

1.5.1 Acid-catalyzed lignin 65

1.5.2 Base-catalyzed lignin 66

1.5.3 Oxidative lignin depolymerization 68

1.5.4 Thermal lignin depolymerization 70

1.5.5 Reductive lignin depolymerization 71
1.6 Challenges in lignocellulosic biomass pretreatment for lignin valorization 72
1.6.1 Preventing structural lignin degradation during pretreatment 73
1.6.2 Fractionation efficiency 75
1.6.3 Identifiation and quantification of lignin’s structural features 75
1.7 Objectives 76
1.7.1 Preventing lignin condensation during the pretreatment process 77
1.7.2 Developing a high-efficiency fractionation process 78
1.7.3 Simultaneous identification and quantification of structural features of lignin by HSQC NMR... 79
Chapter 2 Protecting groups for stabilization of lignin during pretreatment 81
2.1 Utilization of formaldehyde for lignin stabilization 81
2.2 Protecting group optimization towards high-selectivity monomer production 90
Chapter 3 A protocol for fractionation of lignocellulosic biomass 101
3.1 Assessing lignin valorization strategies 102
3.2 Advantages, limitations, and alternative methods 103
3.3 Overview of the procedure 106
3.4 Anticipated results 109
3.4.1 Formaldehyde biomass fracrtionation protocol 109
3.4.2 Propionaldehyde biomass fracrtionation protocol 111
Chapter 4 Lignin monomer yield prediction by quantitative 2D-HSQC NMR 115
Chapter 5 Conclusion 127
5.1 Summary of the work 127
5.2 Outlook and future research 128
Appendix A Appendix for chapter 2 131
A.1 Materials and methods 131
A.1.1 Chemicals and materials 131
A.1.2 Experimental methods 132
A.1.3 Analytical methods 136
A.2 Experimental studies 139
A.2.1 Monomer identification and quantificaiton 139
A.2.2 Model Compound Studies with Vanillyl Alcohol 142
A.2.3 Recovering the Principal Biomass Polysaccharides after Lignin Extraction 143
A.2.4 Formaldehyde Mass Balance and Economic Analysis 144
A.2.5 Figures 150
A.2.6 Tables 171
Appendix B Appendix for chapter 3 179
B.1 Experimental design 179

B.1.1 Prior to the fractionation ... 179

B.1.2 Fractionation ... 180

B.1.3 Depolymerization of the extracted biopolymers ... 181

B.2 Procedure ... 182

B.2.1 Formaldehyde biomass fractionation protocol ... 182

B.2.1.1Pretreatment of the biomass ... 182

B.2.1.2Cellulose collection ... 183

B.2.1.3Formaldehyde-stabilized lignin collection ... 184

B.2.1.4Formylated C_5 sugar collection ... 185

B.2.2 Propionaldehyde biomass fractionation protocol ... 187

B.2.2.1Pretreatment of the biomass ... 187

B.2.2.2Cellulose collection ... 188

B.2.2.3Propionaldehyde-stabilized lignin collection ... 189

B.2.2.4Propylated C_5 sugar collection ... 191

B.2.3 Enzymatic cellulose hydrolysis ... 193

B.2.3.1Enzymatic hydrolysis of cellulose ... 193

B.2.3.2Determine the cellulose hydration ... 194

B.2.3.3Determine the glucose content of the cellulose ... 194

B.2.4 Lignin hydrogenolysis ... 196

B.3 Equipment setup ... 198

B.4 Boxes ... 202

B.5 Hemicellulose derivatives characterizations ... 205

B.6 Figures ... 207

Appendix C Appendix for chapter 4 ... **221**

S1 Chemicals and Materials ... 221

S2 Experimental Methods ... 221

S2.1 Lignin isolation ... 221

S2.2 Hydrogenolysis of lignin into lignin monomers ... 224

S3 Analytical Methods ... 224

References 237

Curriculum Vitae ... **256**

List of Figures

Figure 1 Remaining oil resources. 21

Figure 2 Global GHG emission by gas 22

Figure 3 World total primary energy supply by source 23

Figure 4 Structural binding of lignocellulosic biomass 26

Figure 5 Structure of lignocellulosic biomass and building blocks of its three main components 28

Figure 6 Cellulose structure 30

Figure 7 Hemicellulose structure model 31

Figure 8 Lignin structure model 33

Figure 9 Linkages ocurring in lignin structure 35

Figure 10 Seven proposed platform bioproducts to be produced from lignocellulose-derived sugars. 40

Figure 11 Carbohydrate-first lignocellulosic biomass pretreatments 47

Figure 12 Lignin-first lignocellulosic biomass pretreatments. 53

Figure 13 Acid-catalyzed lignin chemistry 66

Figure 14 Alkaline lignin chemistry. Phenolic rings are labeled with A–D. 67

Figure 15 Oxidative lignin chemistry 69

Figure 16 Reductive lignin chemistry 72

Figure 17 Lignin monomer production by extraction followed by hydrogenolysis. 82

Figure 18 Acid-catalyzed depolymerization of veratrylglycerol-β-guaiacyl ether (VG). 84

Figure 19 Production of lignin monomers from lignocellulosic biomass. 86

Figure 20 Mass balance of polysaccharide (glucan and xylan) and lignin fractions. 89

Figure 21 Hydrogenolysis of birch lignin extracted with the addition of different protection reagents.91

Figure 22 HSQC spectra of the product mixture 93

Figure 23 Reaction mechanism for the hydrogenolysis of aldehyde-stabilized lignin. 94

Figure 24 Monomer yields (on the basis of Klason lignin) and product distribution 96

Figure 25 Monomer yields (on the basis of Klason lignin) and product. 98

Figure 26 Mass balance of polysaccharides (glucan and xylan) and lignin fractions 99

Figure 27 The products of the propionaldehyde-based fractionation protocol. 102

Figure 28 Hydrogenolysis Data for the Formaldehyde and Propionaldehyde Stabilized Lignins 105

Figure 29 The Formaldehyde Biomass Fractionation Protocol. 108

Figure 30 . The Propionaldehyde Biomass Fractionation Protocol 109

Figure 31 Experimental sequence. 116

Figure 32 2D-HSQC NMR spectra of isolated lignins. 118

Figure 33 Extrapolation of 2D HSQCi (i=1, 2, 3) integrated peak volumes (Vi), to find V_0 values.. 120

Figure 34 Comparison of different monomer yield models based on extracted lignin. 123

Figure 35 Predicted yields versus experimental yields.. 125

Figure S36 The t_1 noise caused by internal standard (TMS) 231

Figure S37 The main sources of t_1 noise in lignin samples 232

List of Tables

Table 1 Availability of renewable energy sources. 24

Table 2 Composition of three main polymers in different types of lignocellulosic biomass 29

Table 3 Concept of different types of biorefineries 38

Table 4. Composition of Biomass Used for the Fractionation Protocols. 112

Table 5. Yield of Monomers from the Direct Hydrogenolysis of the Biomass on a Dry Basis. 113

Table 6. Yield of Lignin Monomers from the Hydrogenolysis of the Isolated Formaldehyde-Stabilized Lignin Powder 113

Table 7. Yield of Lignin Monomers from the Hydrogenolysis of the Isolated Propionaldehyde-Stabilized Lignin Powder 113

Table 8. Composition and Enzymatic Hydrolysis of the Extracted Cellulose from the Aldehyde Biomass Fraction-ation Protocols. 114

Table 9. Masses and Yields of Isolated Fractions from the Fractionation Proceduresa. 114

List of Equations

Equation 1 120
Equation 2 121
Equation 3 121
Equation 4 122
Equation 5 122

List of Abbreviations

AAP	anhydrous ammonia pretreatment
AFEX	ammonia fiber expansion
APL	alkaline pretreatment liquor
AQ	anthraquinone
ARP	ammonia recycled percolation
AIL	acid insoluble lignin
ASL	acid soluble lignin
CAH	concentrated acid hydrolysis
CEL	cellulolytic enzyme lignin
DAH	dilute acid hydrolysis
DAP	dilute acid pretreatment
DFRC	derivatization followed by reductive cleavage
DL	depolymerized lignin
DMR	deacetylation and mechanical refining
DOE	(United States) department of energy
EAP	extractive ammonia pretreatment
EMAL	enzymatic mild acidolysis lignin
EU	european union
FA	formaldehyde
FID	flame ionization detector
FT	flow through
FT-DAP	flow-through dilute acid process

FT-HWP	flow-through hot water pretreatment
GC	gas chromatography
GHG	greenhouse gases
GVL	γ-valerolactone
HCA	hydroxycinnamic acid
HSP	high syringyl poplar
HSQC	heteronuclear single quantum coherence
HWP	hot water pretreatment
IL	Ionic liquid
LCA	life-cycle analysis
LP	lignin precipitate
LR	lignin residue
MeOH	methanol
MeTHF	2-methyltetrahydrofuran
MIBK	methyl isobutyl ketone
MWL	milled wood lignin
NBO	nitrobenzene oxidation
NMR	nuclear magnetic resonance
PA	propionaldehyde
POHG	4-n-propanolguaiacol
POHS	4-n-propanolsyringol
PS	propylsyringol
Q-HSQC NMR	quantitative HSQC NMR
RCF	reductive catalytic fractionation
SEP	steam explosion pretreatment

THF	tetrahydrofuran
TMS	tetramethylsilane
UV	ultra violet
VG	veratrylglycerol-β-guaiacyl ether

Chapter 1 Introduction

1.1 Lignocellulosic biomass as a renewable energy source

The world currently relies on fossil fuels to fulfill its energy needs. Almost every sector uses some form of fossil fuel in running its daily operations. However, the over-reliance on petroleum presents a significant challenge for the future as fossil resources are bound to get depleted. With the increase in industrial activities and population growth, it is expected that the world's energy demands will keep increasing. In comparison to the trend in energy needs, the remaining reservoirs of fossil fuels, which are non-renewable sources of energy, depict a consistent decline that predicts their eventual exhaustion in the future (Figure1).[1]

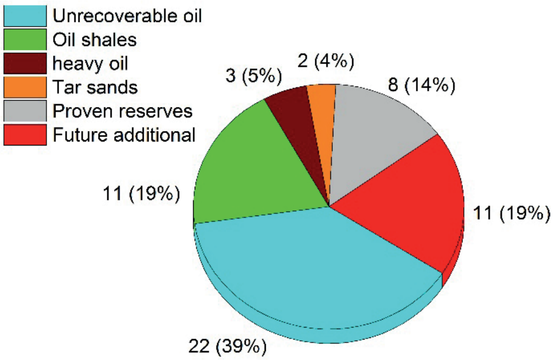


Figure 1 Remaining oil resources.[1]
Numbers are in ZJ with the total count of 57 ZJ of reserves and the annual consumption of 0.18 ZJ.

Furthermore, the continued use of fossil fuels presents significant threats to the environment. A major part of the effects involves air pollution through the emission of greenhouse gases (GHG) produced by the combustion of fossil feedstocks. Evidence shows that these greenhouse gases contribute significantly to causing climate change.[2] Fossil fuels are responsible for the emission of vast amounts of pollutants leading to climate change, with carbon dioxide being the primary pollutant emitted as high as 35 Gtons per year contributing to 65% of the total emitted GHG (Figure 2). [3]

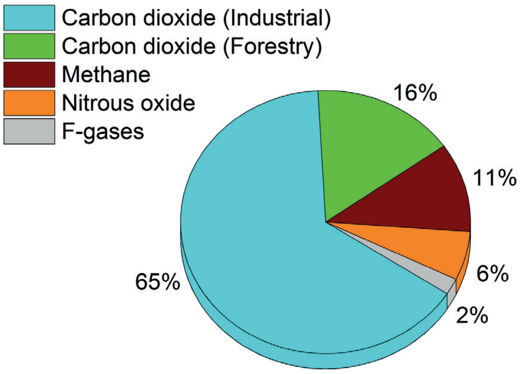


Figure 2 Global GHG emission by gas [3]

Considering the massive role played by the combustion of fossil fuels in skyrocketing GHG emissions, the need for alternative sources of energy cannot be overemphasized. Unlike fossil fuels, the use of renewable energy results in minimal impact on the global climate. Even when accounting for the full life-cycle of renewable energy production, the result points to minimal global warming emissions resulting from such sources of energy.[4] Efforts in taping renewable resources have resulted in a gradual increase in the overall supply of renewable energy reaching up to 13% of global supply (Figure 3). Apart from nuclear energy, as a debatable renewable energy source, and hydroelectric power, other emerging sources such as solar, wind, and biomass are also increasingly contributing to the renewable energy grid.[5]

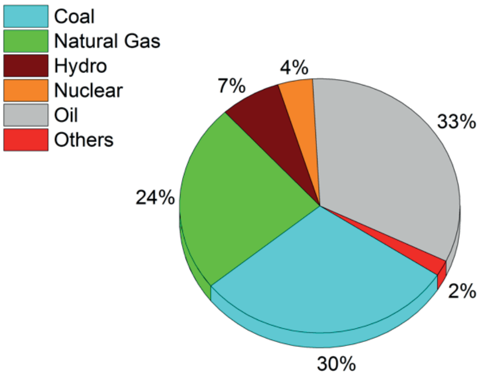


Figure 3 World total primary energy supply by source [5]

Keeping up with the demand for renewable energy calls for concerted efforts to increase the energy generation by more than 75% from 18 TW-year to 32 TW-year by 2050.[5] The energy sustainability goals require a comparison of the estimated global energy demands to the potential generation of energy from renewable sources (Table 1). Due to the shortcomings in the generation of renewable energy, the promise of utilizing such sources show solar energy as the sole source that is potentially capable of supplying the required energy[6]. Unfortunately, even the latest technologies for utilizing solar energy face considerable challenges. In practice, these technologies are suffering from drawbacks such as low conversion yields[7], stability issues[8,9], high costs[10], and storage.[11] Similarly, the uneven distribution of solar energy worldwide limits its use in all countries.[11] Given the massive need for renewable energy in the face of significant challenges in tapping solar energy, the need for advanced technologies for using other renewable sources of energy such as wind, geothermal or biomass remains of utmost importance.

Table 1 Availability of renewable energy sources

	Available (TW-yr)	Extractable (TW-yr)
Demand [5]	32	
Solar Energy [12,13]	1.2×10^5	600
Biomass [14,15]	25	5-7
Wind [16]	72	4-7
Hydro [17]	4-5	1.7
Geothermal [18]	12	<1*

* The extractable amount for electricity production

One of the potential sources of renewable energy is biomass. The relevance of biomass arises principally from the vast amounts of biomass produced on a yearly basis, which makes it a highly potent renewable energy feedstock. Out of the 25 TW of biomass grown annually, 5 to 7 TW could be used for energy production.[15] Even though this amount is not sufficient to cover the world's primary energy needs, it plays an important role for transitioning from fossil fuels to renewable energy as complementary to solar energy.[19] It is also crucial to mention that solar energy with its current conversion and storage technologies, cannot replace all forms of fuels such as aviation fuels.[20] Aviation fuels need high energy density to meet the requirements for reliable long distance transportation.[21] The main difference between biomass and solar energy is that biomass can function as a direct source of renewable carbon.[19] Similar to fossil fuels, biomass is capable of producing carbon-based fuels that offer high energy density fuels and can be also used in other conversion technologies such as fuel cells for heat and power generation.

Moreover, the use of biomass can also gradually replace petroleum for the production of non-energy products with high market demand such as polymers, additives, and resins. This is an important aspect of biomass utilization as there are insufficient alternatives apart from petroleum for the production of carbon-based chemicals. One of the only other sustainable alternatives is to capture atmospheric CO_2 and convert it to higher value-added chemicals.[22] However, conversion of CO_2 to many products (including aromatic molecules) has not been demonstrated at reasonable yields.[23] Furthermore, the scalability along with the cost and energy consumption for such processes remains much too high for most products and cannot currently meet the market demands.[24]

Biomass is defined as all organic matter derived from animals or plants. The classification of biomass as a source of energy can result in the definition of two feedstock categories: waste and energy crops.[25] Wastes comprise of all residues obtained from livestock, forestry or other biological solid wastes. On the other hand, energy crops consist of all plants grown purposely as sources of biomass for energy production that can be harvested several times during the year to increase the yield of biomass production. The biomass resource obtained from terrestrial plants referred to as lignocellulosic biomass. Carbohydrate polymers (cellulose and hemicellulose) form the major part of the structure of lignocellulosic biomass while lignin binds these compounds to maintain their overall structure (Figure 4). The primary use of lignocellulosic biomass is in pulp and paper industry which focuses on separation of cellulosic fraction from lignin.[26] It can be also used for production of various chemicals that bear the name bioproducts.[27]

In this section, we gave an overview of the importance and role of lignocellulosic biomass in replacing petroleum for providing energy and high market demand chemicals. For better understanding of chemical processes for production of biofuels and bioproducts, first we need to be familiar with the composition of lignocellulosic biomass and the sources of carbon it can offer. Therefore, the next chapter presents the constituents of lignocellulosic biomass and their chemistry, which is necessary to understand before we can continue with valorization possibilities of these components.

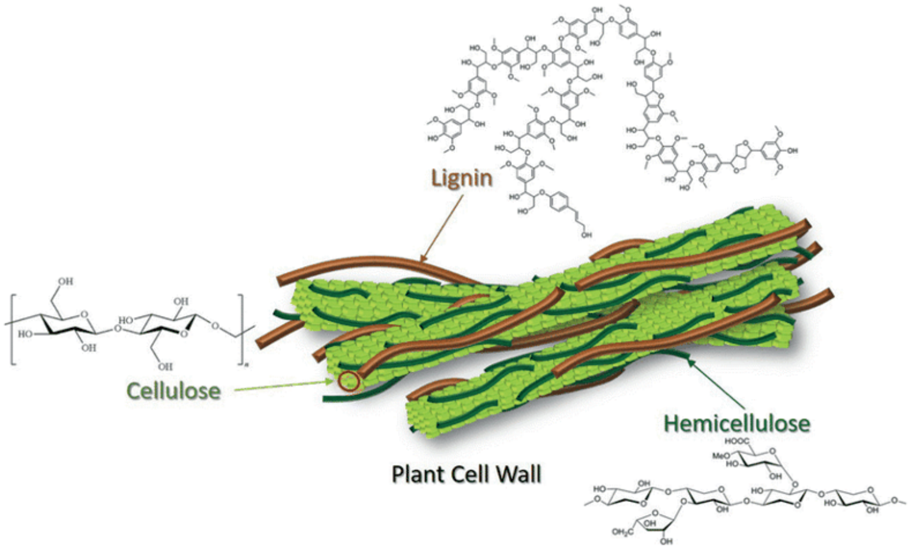


Figure 4 Structural binding of lignocellulosic biomass

Reprinted by permission from **Springer Nature Customer Service Centre GmbH**: **Springer Nature**, **Biomass Conversion and Biorefinery**, Fundamentals of Hydrofaction™: Renewable crude oil from woody biomass, Claus Uhrenholt Jensen, Julie Katerine Rodriguez Guerrero, Sergios Karatzos et al**, COPYRIGHT** (2017)

1.2 Composition of lignocellulosic biomass

The major constituents of plant biomass include three polymers known as cellulose, hemicellulose, and lignin along with lesser amounts of extractives, protein, pectin, ash, and moisture.[28] With evolution, the structure of lignocellulose maximized its capability to resist degradation by benefiting from the crystallinity of cellulose, hydrophobicity of lignin, and encapsulation of cellulose by the lignin-hemicellulose matrix.[29] The binding of the three main components of lignocellulosic biomass and their building blocks is shown in Figure 5. Within the plant structure, the type of lignocellulosic biomass influences the complex non-uniform three-dimensional organization of these polymers and also their relative content in the plant. For instance, the content of cellulose is higher in hardwoods while leaves and grasses contain higher contents of hemicellulose.[30] The variations in the polymer ratios also exist in a plant depending on the number of years, growth stage and source among other factors.

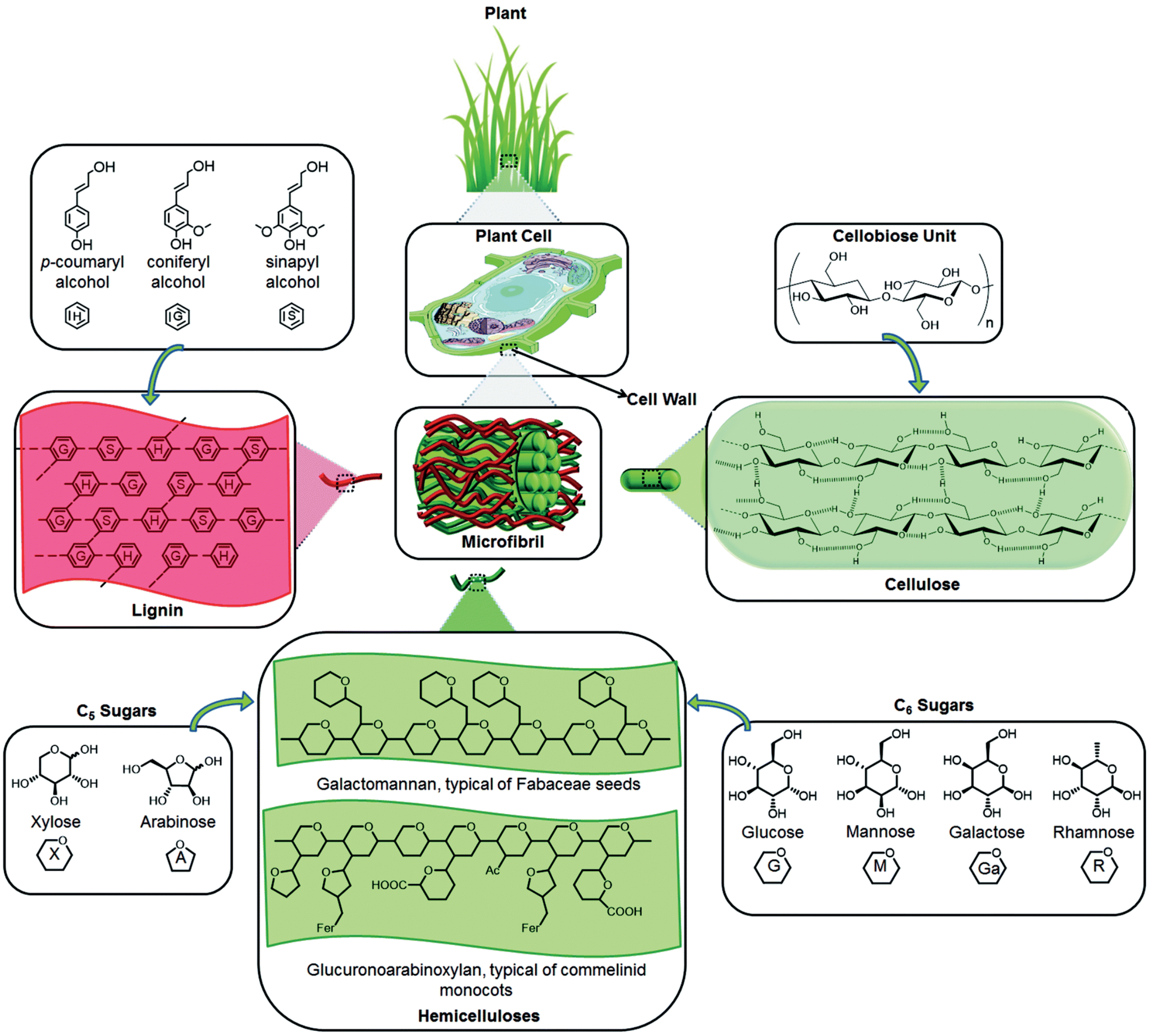

Plant
Plant Cell
Cell Wall
Cellobiose Unit
p-coumaryl alcohol
coniferyl alcohol
sinapyl alcohol
H
G
S
Lignin
Microfibril
Cellulose
C5 Sugars
Xylose
Arabinose
X
A
Galactomannan, typical of Fabaceae seeds
Glucuronoarabinoxylan, typical of commelinid monocots
Hemicelluloses
C6 Sugars
Glucose
Mannose
Galactose
Rhamnose
G
M
Ga
R
HOOC
Ac
Fer
COOH
OH
HO
n

Figure 5 Structure of lignocellulosic biomass and building blocks of its three main components

Reprinted from "Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers", Furkan H. Isikgor, C. Remzi Becer, Polym. Chem., 2015, 6, 4497 (**DOI:** 10.1039/C5PY00263J) - Published by The Royal Society of Chemistry.

As pertains to the biofuel production process, the content of the three major polymers can influence the optimal conversion processes needed for maximum energy or carbon recovery. Even though the distribution of each polymer is uneven throughout the plant, the composition of lignocellulosic biomass is generally between 30-60% for cellulose, 20-45% for hemicellulose and 10-35% for lignin (Table 2).[31] The summary of the composition and chemistry of these three main fractions of lignocellulosic biomass is discussed in this section.

Table 2 Composition of three main polymers in different types of lignocellulosic biomass[32]

Biomass	Cellulose	Hemicellulose	Lignin
Hardwoods	40-55	24-40	18-25
Softwoods	45-50	25-35	25-35
Wheat straw	35	50	15
Corn cobs	45	35	15
Grasses	30-40	35-50	10-20
Switchgrass	45	35-40	10-15

1.2.1 Sugars

The major sources of sugars within the lignocellulosic biomass are hemicellulose and cellulose. Cellulose is the most abundant constituent of lignocellulosic biomass which contributes to the structure and rigidity of the plant cell wall and is made up of repeating disaccharide cellobiose units (Figure 6). In its more detailed configuration, cellulose is a homopolymer of glucose in the pyranose form without branching bonds. Glucopyranoses bind together through glycosidic bonds that result in the formation of a polymer.[33] Despite the biofuel processing challenges presented by the natural structure of cellulose, its occurrence as a uniform and highly abundant source of carbon offers unique opportunities in biofuel and bioproduct production. The cellulose chain can have a degree of polymerization ranging from 10’000 to 15’000 glucopyranose units for wood and cotton respectively. [34] Individual chains of cellulose link together in groups of 20-300 to form structures referred to as microfibrils. A similar grouping of the microfibrils results in the formation of cellulose fibers. Hydrogen bonds occur predominantly within the cellulose chains with van der Waals forces causing the eventual grouping of the chains into microfibrils. These microfibrils are covered with hemicellulose and lignin cross-linking. Since the glucose molecules forming cellulose bind through β-(1,4)-glycosidic linkages, it becomes possible to break the bonds through acidic or enzymatic reaction to release D-glucose that can be fermented.[35]

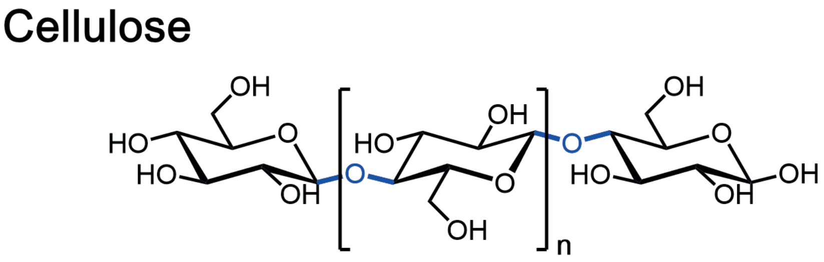


Figure 6 Cellulose structure

As the main structure forming the plant cell wall, cellulose occurs in biomass in two different forms. The one form is the crystalline cellulose that is the predominant form of cellulose. The other form is amorphous cellulose that occurs in small quantities within cellulose chains. This form is also the most susceptible to enzymatic breakdown.[36] Where the hydrogen bonds occur in an orderly sequence, the resulting structure takes a crystalline appearance. Conversely, when such bonds are disordered, amorphous cellulose is formed.[37]

Hemicellulose is generally the second most abundant polymer in biomass. Hemicellulose, though also a polysaccharide, lacks homogeneity within its chemical structure in contrast to cellulose.

Different monosaccharide sugars join to form linear and branched structures (Figure 7). These monosaccharides are pentoses, hexoses and uronic acids. Three common pentoses include xylose, rhamnose, and arabinose; three common hexoses include glucose, mannose and galactose; and uronic acids include 4- O-methylglucuronic, D-glucuronic and D-galactouronic acids. The formation of the major structure in hemicellulose requires glycosidic bonds of the short chain compounds to form a polymer. These glycosidic bonds are mainly β-(1,4)-glycosidic linkages and occasionally β-(1,3)-glycosidic linkages. Other than predominantly comprising aldose sugars, hemicelluloses contain some component of acetyl sugars that form heteropolymers such as heteroxylan.[38] Some of the main units of hemicellulose are shown in Figure 5. Due to its amorphous structure, hemicellulose has less resistance to degradation compared to cellulose[39], but can provide structural strength with its complex bonding network that links cellulose fibers into microfibrils and cross-links with lignin. The composition of hemicellulose can differ depending on the species. Hardwood, such as grass or straw, contains xylan-rich hemicellulose and softwood contains mainly glucomannans.

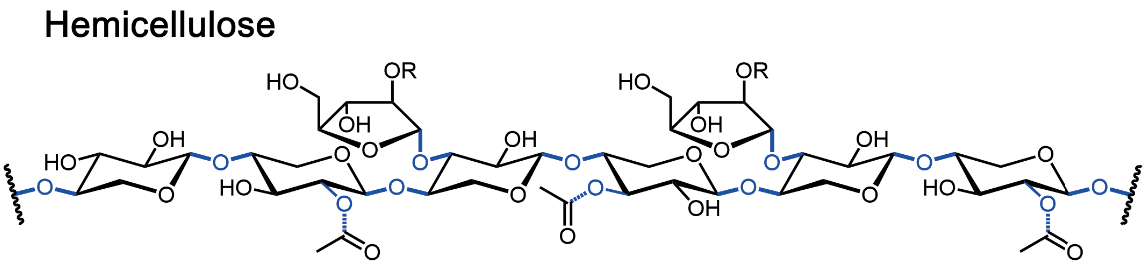


Figure 7 Hemicellulose structure model

Xylans are heteropolysaccharides in many plants that have 1,4- linked β-D-xylopyranose bonds. As a unit comprised of numerous polysaccharides, xylan is extracted easily when the biomass is put in an alkaline or acidic medium. On the other hand, the homogeneity of the polysaccharides formed by glucomannan necessitates the use of a concentrated alkaline medium for its extraction.[32] Considering the fact that hemicellulose covers the cellulose-fibrils in the plant's cell wall, it can inhibit the digestive processes targeting cellulose. As such, efficient biological deconstruction of cellulose requires the removal of at least 50% of hemicellulose to enhance cellulose digestibility.[40] All the same, the process requires sufficient severity and optimized conditions to ensure that removing hemicellulose does not result in production of degradation compounds such as furans that are known to prevent the fermentation of sugars.[41] More often than not, the initial process is optimized to ensure maximum retrieval of cellulose sugars while at least limiting hemicellulose degradation and ensuring its removal.[42]

1.2.2 Lignin

In the plant biomass, lignin is generally the third most abundant fraction. The structure of lignin consists of phenyl propanoid sub-unit linked irregularly together by several different motifs to form a largely linear heteropolymer (Figure 8).[43] The structure of lignin allows it to keep all the other components of lignocellulosic biomass bound together, hence forming a complex hydrophobic structure. By closely adhering to cellulose microfibrils, lignin acts as an inhibitor against hydrolysis of these microfibrils. Its complex structure also leads to its protective role in the cell wall by providing resistance to pathogens and various oxidation processes.[38]

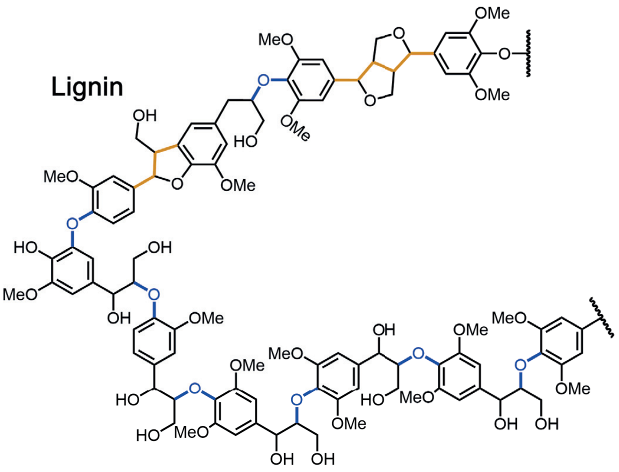


Figure 8 Lignin structure model

The phenyl propionic alcohol building blocks of lignin are called monolignols and include:

1. Coniferyl alcohol (guaiacyl propanol) abbreviated as G units.
2. Coumaryl alcohol (p-hydroxyphenyl propanol) abbreviated as H units.
3. Sinapyl alcohol (syringyl alcohol) abbreviated as S units.[44]

These units differ in the number of methoxy groups on the phenolic center (Figure 5). Plants with high quantities of wood contain high quantities of lignin while herbaceous plants like grasses contain smaller lignin quantitities.[40] Moreover, the distribution of monolignols also differs in different plant species. Softwood plants such as spruce or pine are exclusively formed by coniferyl alcohol (G units), while the lignin in hardwood plants such as poplar or birch is formed of both coniferyl alcohol and sinapyl alcohol (G and S units). On the other hand, herbaceous plants contain all three monolignols (G, S and H units).[45] Apart from the three mentioned major building blocks, lignin can also contain other compounds (mainly phenolic compounds) such as hydroxycinnamates[46], tricin[47], p-hydroxybenzoate[48] and acetate[49].

The bonds linking the monolignols within the lignin structure comprise of different ether linkages (alkyl-aryl, alkyl-alkyl and aryl-aryl) that form a complex structure.[32] Among all different type of linkages, the ether linkages of β-O-4 and α-O-4 are more interesting due to their higher reactivity for cleaving reactions.[50] Carbon-carbon inter-unit linkages are the other important linkages that occur in lignin. These linkages can be in the form of 5-5, β-5, β-1 and β-β (Figure 9).[51] The distribution of monolignols in lignin dictates the amount of inter-unit linkages. Softwood lignin, which is mainly formed of G units, contains higher amount of carbon-carbon linkages compared to hardwood.[31] Due to the difficulties in isolating lignin that results in structural changes during the extraction process, it is difficult to measure the molecular weight of native lignin. However, some studies were able to report the number average molecular weight for close-to-native lignin that ranges between 2500 to 10’000 g mol^{-1}.[52]

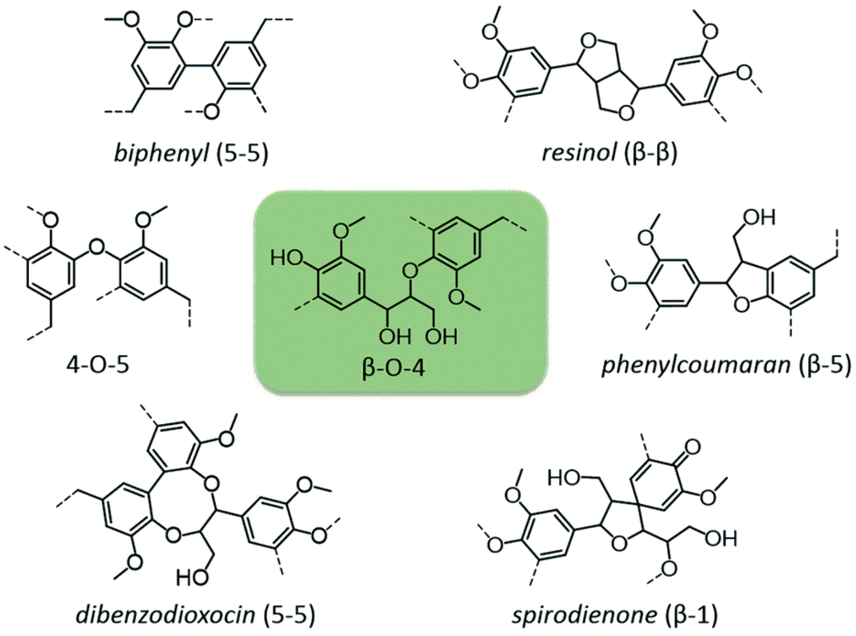


Figure 9 Linkages ocurring in lignin structure

Adopted by permission from **ROYAL SOCIETY OF CHEMISTRY**: **Chemical Society reveiws**, Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading, W. Schutyser, T. Renders, S. Van den Bosch, S.-F. Koelewijn, G. T. Beckham and B. F. Sels**, COPYRIGHT** (2018)

Cellulose, hemicellulose and lignin are the predominant components of lignocellulosic biomass that can be used as a source of carbon for further upgrading. Therefore the main focus of biofuel and bioproduct production is on valorizing these polymers. Studying other components of biomass such as protein or pectin is also very important to understand their effect on biomass valorization processes or their potential of being upgraded to higher value-added products.[53,54] To avoid complexity, the chemistry and effects of these components are not presented in this thesis. The next section provides an overview of lignocellulosic biomass valorization processes and some of the existing challenges with the focus on the role of sugars and lignin.

1.3 Utilization of biomass for production of biofuels and bioproducts

Prior to the industrial revolution, human beings depended mostly on wood for their energy supply. Shifts over time have resulted in a substantial decline in the use of biomass to the point that it barely reaches 10% of the overall energy supply today, of which two-third is used in cooking and heating.[55] Recently, the interest in biofuels has been revived as result of the need for generating petroleum substitutes to reduce the impact of burning fossil fuels.

With the environmental threat posed by continued use of fossil fuels, policymakers have also sought to enact pertinent rules towards accelerating the research on biofuel production. Economic communities such as the European Union are pushing the use of renewable energy in transportation. By 2020, the EU aims to have 10% of the transport fuel of every EU country come from renewable sources such as biofuels. Moreover, they impose requirements on fuel suppliers that mandate them to decrease the potency of emissions in the fuels supplied. Fuel suppliers are required to reduce the GHG intensity of the EU fuel mix by 6% by 2020 in comparison to 2010.[56]

In the endeavor to reduce the emissions of greenhouse gases, a major step involves gaining a comprehensive understanding of CO_2 emissions as they represent the major fraction of greenhouse gases. With the promising prospect of using biofuels in place of fossil fuels, it is important to explore the carbon footprints of biofuel generation.[57] Another important issue is the issue of prioritizing biofuels at the expense of food production which may result in unintended consequences in respect to the food prices and availability.[57] On the one hand, natural forest reserves and grasslands could be used as grounds for growing energy crops. However, further analysis reveals that such approaches might result in even higher emissions of carbon as compared to fossil fuels. For example, converting rainforests or grasslands for production of biofuels from food crops releases 17 to 420 times more CO_2 than the annual GHG reductions that these biofuels would provide by replacing fossil fuels.[58] As such, the solution lies in utilizing waste biomass to produce biofuels. To augment the biomass supply, growing energy crops on derelict land also seems feasible and sustainable as it could contribute to the ultimate objective of reduced greenhouse emissions.[57]

Depending on the biofuel that is targeted, the resulting carbon footprint varies due to the process and the feedstock used. For example, different processes for generating biodiesel lead to sharp differences in CO_2 emission; from 13 gr emitted per MJ of bio-diesel produced from cooking oil to 73 gr/MJ when soy beans are used as the feedstock which results in almost equal emissions to conventional petroleum-based diesel.[59] Similarly, the production of bioproducts has to be carefully designed to ensure that they actually lead to a substantial decreases in carbon emission. Therefore, there is a need to combine the different biomass conversion routes in a network of technologies that can receive different feedstock as input to produce a range of biofuels and bioproducts. This can help us to perform technological optimization and life-cycle analysis (LCA) towards studying the feasibility of the implementation of such technologies along with increasing the profitability and decreasing carbon emission. Such network of technologies for valorizing biomass is called biorefinery.[60]

Essentially, the biorefinery concept borrows from that of typical oil refineries used in petrochemistry: a production facility should have the capacity to convert all types of biomass feedstocks and fractions to produce different classes of bioproducts and biofuels by employing various conversion technologies.[61,62] The biorefinery needs to ensure an expeditious process that not only generates the needed energy to satisfy worldwide demands but also effectively reduces GHG emissions.[63] Table 3 summarizes the concepts of different types of biorefineries.[60]

Table 3 Concept of different types of biorefineries[60]

Concept	Type of feedstock	Predominant technology	Phase of development
Green Biorefineries (GBR)	wet biomass: green grasses and green corps	pretreatment, pressing, fractionation, separation, digestion	Pilot plant (and R&D)
Whole Crop Biorefineries (WCBR)	whole crop (including straw) cereals such as rye, wheat and maize	dry or wet milling, bioproduct conversion	Pilot plant (and Demo)
Lignocellulosic Feedstock Biorefineries (LCBFR)	lignocellulosic-rich biomass such as straw, chaff, reed, wood	pretreatment, chemical and enzymatic hydrolysis, fermentation, separation	R&D/Pilot plant (EC), Demo (US)
Two Platform Concept Biorefineries (TPCBR)	all types of biomass	Combination of sugar platform (bioproduct conversion) and syngas platform (thermochemical conversion)	Pilot plant
Thermochemical Biorefineries (TCBR)	all types of biomass	thermochemical conversion: torrefaction, pyrolysis, gasification, HTU, separation, catalytic reaction	Pilot plant (R&D and Demo)
Marine Biorefineries (MBR)	aquatic biomass: microalgae and macroalgae (seaweed)	cell disruption, extraction and separation	R&D (and Pilot plant)

So far, only lignocellulosic feedstock biorefineries are industrially developed at the demo phase. Notably, when compared with petroleum products, lignocellulosic biomass contains a greater content of oxygen and consequently lower relative amounts of carbon and hydrogen. This results in a decreased heat content of the produced biofuels and limits their blending with conventional fuels.[64]

For compatible use of biofuels in the transport sector, biorefineries seek to ensure rapid depolymerization and deoxygenation of lignocellulosic biomass to ensure efficient combustion. On the other hand, the structure of lignocellulosic biomass enables these technologies to potentially generate several products that are not directly accessible from petroleum and could lead to the production of higher value-added bioproducts.[65]

During the processing of biomass within the biorefinery, in past years, the priority has been given to the extraction of sugar compounds as the release of these compounds paves the way for the production of various products that result from further processing of these sugars.[66] When biomass is broken down to release cellulose, the resulting degradation of cellulose results in the production of glucose. Similarly, the breakdown of polymers within hemicellulose gives rise to various C_5 and C_6 sugars.

In their research, scientists in the US Department of Energy (DOE) propose to use sugars to synthesize high market value chemicals. Specifically, using C_5 and C_6 sugars, twelve platform bioproducts are considered to be produced through the proposed synthetic process.[67] In addition to this list, many other different chemicals can be achieved through biomass processing. Figure 10 summarizes the routes to some of these platform chemicals and their derivatives.[68]

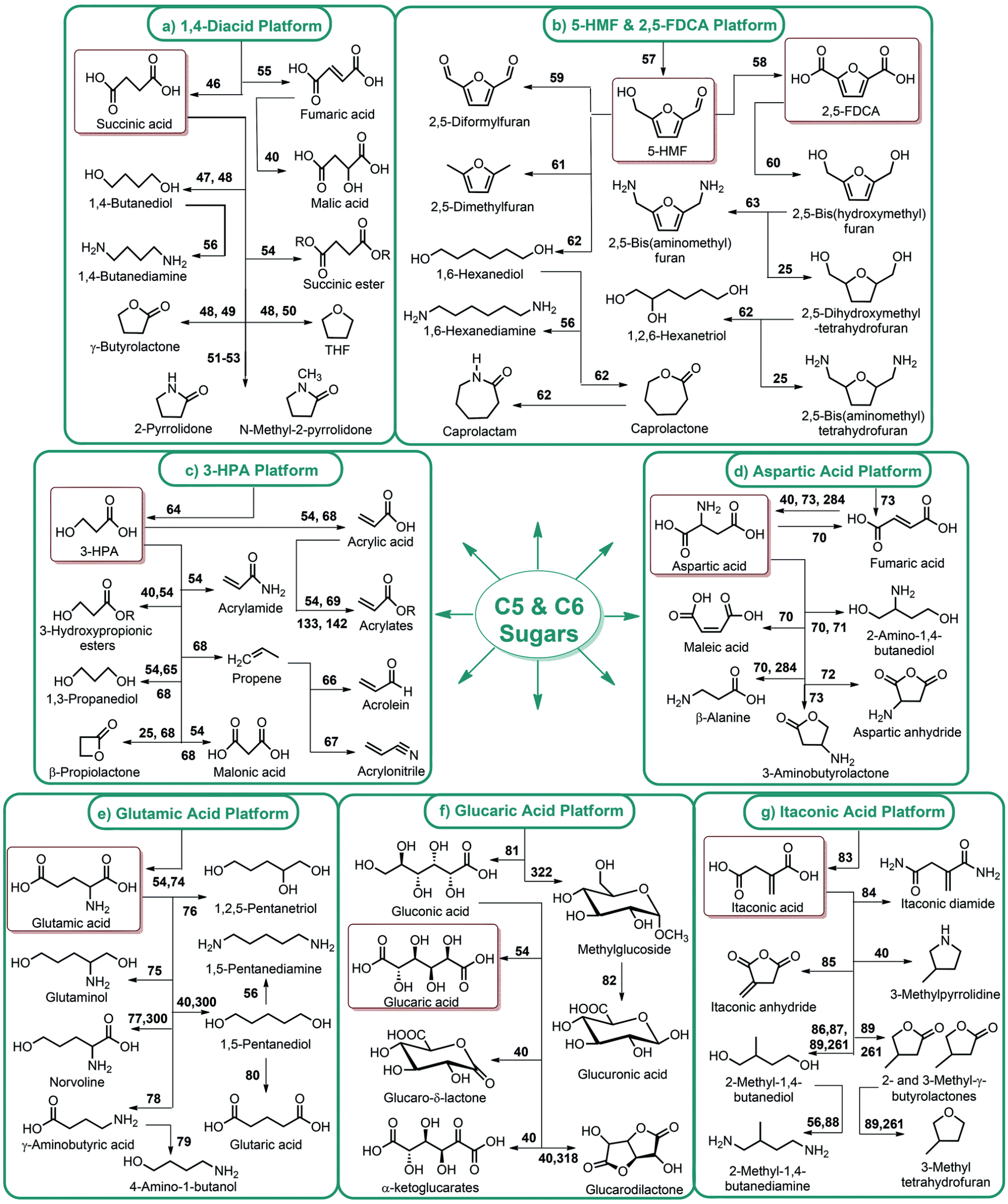

a) 1,4-Diacid Platform
Succinic acid
Fumaric acid
Malic acid
1,4-Butanediol
1,4-Butanediamine
Succinic ester
γ-Butyrolactone
THF
2-Pyrrolidone
N-Methyl-2-pyrrolidone
b) 5-HMF & 2,5-FDCA Platform
2,5-Diformylfuran
5-HMF
2,5-FDCA
2,5-Dimethylfuran
2,5-Bis(hydroxymethyl) furan
2,5-Bis(aminomethyl) furan
1,6-Hexanediol
1,6-Hexanediamine
1,2,6-Hexanetriol
2,5-Dihydroxymethyl -tetrahydrofuran
2,5-Bis(aminomethyl) tetrahydrofuran
Caprolactam
Caprolactone
c) 3-HPA Platform
3-HPA
Acrylic acid
3-Hydroxypropionic esters
Acrylamide
Acrylates
Propene
1,3-Propanediol
Acrolein
β-Propiolactone
Malonic acid
Acrylonitrile
C5 & C6 Sugars
d) Aspartic Acid Platform
Aspartic acid
Fumaric acid
Maleic acid
2-Amino-1,4-butanediol
β-Alanine
Aspartic anhydride
3-Aminobutyrolactone
e) Glutamic Acid Platform
Glutamic acid
1,2,5-Pentanetriol
Glutaminol
1,5-Pentanediamine
Norvoline
1,5-Pentanediol
γ-Aminobutyric acid
Glutaric acid
4-Amino-1-butanol
f) Glucaric Acid Platform
Gluconic acid
Methylglucoside
Glucaric acid
Glucuronic acid
Glucaro-δ-lactone
α-ketoglucarates
Glucarodilactone
g) Itaconic Acid Platform
Itaconic acid
Itaconic diamide
Itaconic anhydride
3-Methylpyrrolidine
2-Methyl-1,4-butanediol
2- and 3-Methyl-γ-butyrolactones
2-Methyl-1,4-butanediamine
3-Methyl tetrahydrofuran

Figure 10 Seven proposed platform bioproducts to be produced from lignocellulose-derived sugars

Reprinted from "Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers", Furkan H. Isikgor, C. Remzi Becer, Polym. Chem., 2015, 6, 4497 (**DOI:** 10.1039/C5PY00263J) - Published by The Royal Society of Chemistry.

Although the prior DOE report was skeptical of the role of ethanol due to its huge demand in the context of transportation fuels, which would require massive production, a review by Bozell and Petersen resulted in the inclusion of ethanol.[69] Their reason was that in its definition, platform chemicals such as ethanol could be dehydrated to produce other more important bioproducts like ethylene. Apart from ethanol, the same review identified acetone as a potential platform chemical due to the high-value chemicals generated from its conversion. Further studies on the feasibility of using ethanol, acetone, and butanol as platform chemicals revealed the huge potential of these chemicals in synthesizing other products including alkenes and chlorides that greatly influence the industrial production of important polymers. [69] Another important chemical omitted was lactic acid as its potency seemed doubtful. However, further analysis revealed that lactic acid has significant bio-functionality which means that it could be converted to other intermediate chemicals for further upgrading.[69] As such, the breakdown of sugar polymers remains the primary upgrading focus for production of biomass-derived platform chemicals in most biorefinery concepts.

The need for an efficient process for producing sugars has resulted in the development of various hydrolysis methods, including an early developed pretreatment involving concentrated hydrochloric acid which remains an important option for the extraction of low-cost digestible sugars from lignocellulosic biomass in industry.[70] The major shortcoming with this process remains the challenge of sufficient acid recovery. As such, this challenge demands the need for further research to facilitate the separation of chemicals including acids from the products of hydrolysis.[71] To attain full exploitation of biomass given its compositional variety, advanced treatment methods are required[72] for biomass fractionation which will be discussed in the next section.

Another important point is that separating lignin from polysaccharides significantly facilitates the subsequent upgrading to and from sugars, which is vital for the production of biofuel.[73] Apart from physically impeding the process of breaking down plant biomass, lignin leads to undesirable effects in processing biomass, such as:

1. Reducing the efficiency of hydrolytic enzymes through competitive adsorption on lignin.
2. Similar binding of cellulytic enzymes to the lignin-carbohydrate complexes, rendering them ineffective.
3. Lignin by-products' toxicity for some of the microorganisms.[32]

Research into the properties of lignin could facilitate its removal prior to biomass digestion. Through the feedstock's pretreatment, it becomes possible to efficiently remove lignin even in very different plant species. While not all of the pretreatment processes can completely remove lignin from biomass, the chemical structure of lignin can be altered even without extraction. In both cases, as long as the internal surface area of the biomass increases, the result is a significant increase in biomass digestibility which further increases the efficiency of the entire process.[74]

The removal process can be performed by subjecting the biomass to a process of sufficient severity that leads to deconstruction or solubilization of lignin and sometimes condensation upon cooling.[75] The result of the process causes an alteration in the chemical structure of lignin, allowing for precipitation and removal.[76] The chemical pretreatment of lignocellulosic biomass to extract lignin is known as delignification and can have following effects:

1. Swelling of the biomass.

2. Structural disruption of lignin.
3. Higher internal surface area.
4. Improved enzymatic activities through higher accessibility of cellulose fibers for cellulolytic enzymes.[32]

Lignin can account for up to 30% of the weight or 40% of the energy value of the lignocellulosic biomass.[77] However, despite making up a significant portion of the overall biomass, measures to fully exploit the extracted lignin remain under-developed. As a result, lignin's main use in the biorefinery is combustion to release energy.[72] That being said, lignin's phenolic heteropolymeric structure could, in theory, make it an attractive source for production of different valuable products such as phenolic resins, adhesives, and aromatics.[78] Valorization of extracted lignin along with enzyme improvements and logistic challenges are regarded as the most important challenges for increasing profitability of biorefineries.[60]

Valorization of lignin waste comes to a greater importance knowing that lignin is also a waste in other industries apart from biorefineries. The amount of extracted lignin resulting only from pulp and paper production reaches 50 million tons per year where the majority of this lignin is also used as a low value fuel to produce heat and energy and only 2% of the lignin is upgraded to higher value chemicals.[72] Finding a solution for upgrading lignin can enable biorefineries to valorize the wastes from such industries. Overcoming this challenge requires a comprehensive understanding of different types of lignocellulosic biomass pretreatment as the first process in a biorefinery.

1.4 Lignocellulosic biomass pretreatment

As discussed in the previous section, the first objective in refining lignocellulosic biomass is to separate and purify the three main components of lignocellulosic biomass.[71] Due to the differences in their structural properties, it remains impractical to utilize only one step process to attain this complete separation.[79] In fact, such a method would face inconveniences arising from the numerous compounds generated from the breakdown process. As such, inefficiency in this initial biomass processing calls for improved methods while minimizing extra costs to the overall process. From the various processes employed to this effect, the common aspect is the requirement to use high temperatures.[80] Subjecting the biomass to different pretreatment procedures results in significant alterations in the properties of the components, especially by breaking the lignin-carbohydrate bonds. Breaking such bonds paves the way for further separation of the three components. Different categories of pretreatment approaches apply in the biorefinery and include biological, mechanical, and chemical methods among others.[81] In some cases, it is necessary to combine two or more approaches to attain maximum biofuel output. Among the methods employed towards separating the components of biomass, there are: use of radiation, grinding, use of steam, use of ammonia and other chemicals in breakdown the bonds. Although using different materials, all the pretreatment methods aim at increasing biomass surface area for enzymatic activity.[71]

As a result of pretreatment, the array of downstream valorization possibilities can either be widened or become limited. Limitations usually result from lignin alterations due to sulfur incorporation and structural alteration, among others.[82] Several processes of fractionation ranging from traditional pulp and paper industry to more environmentally friendly and sophisticated innovations have been developed. Different lignin products are produced by each of these methods. Such products are often isolated in the form of a lignin precipitate, a solid residue or directly as a depolymerized product mixture.[31] The diverse and extensive possibilities of biomass fractionation will be discussed in this section. The discussion will be organized around the characteristics of the resulting lignin product. These several fractionation technologies can be divided into two classes:

1) The first class involves those methods that focus on the conversion and solubilization of the portions of carbohydrate. Within these processes, lignin is usually isolated as a lignin precipitate (LP) or insoluble lignin residue (LR).
2) The second class is delignification which covers methods targeting separation of lignin from the biomass while the (hemi)cellulose carbohydrates are kept in delignified pulp form. The lignin is usually isolated either as depolymerized lignin oil (DL) or as a lignin precipitate (LP). This depends on the method used.

Lignin precipitate refers to the fractionated lignin that is isolated by methods that first solubilize the lignin in solution followed by a precipitation step, such as the Kraft process[26]; and lignin residue is obtained from the techniques that solubilize the carbohydrates, while the lignin fraction remains as a residual solid, such as in the Klason lignin method.[83]

1.4.1 Fractionation based on carbohydrate conversion

As mentioned above, the recovered products of lignin in these processes, which is mainly a side product, leads to a lignin precipitate (LP) or insoluble lignin residue (LR). Figure 11 shows a summary of the carbohydrate-first fractionation processes, which is further discussed below based on their carbohydrate deconstruction mechanism.

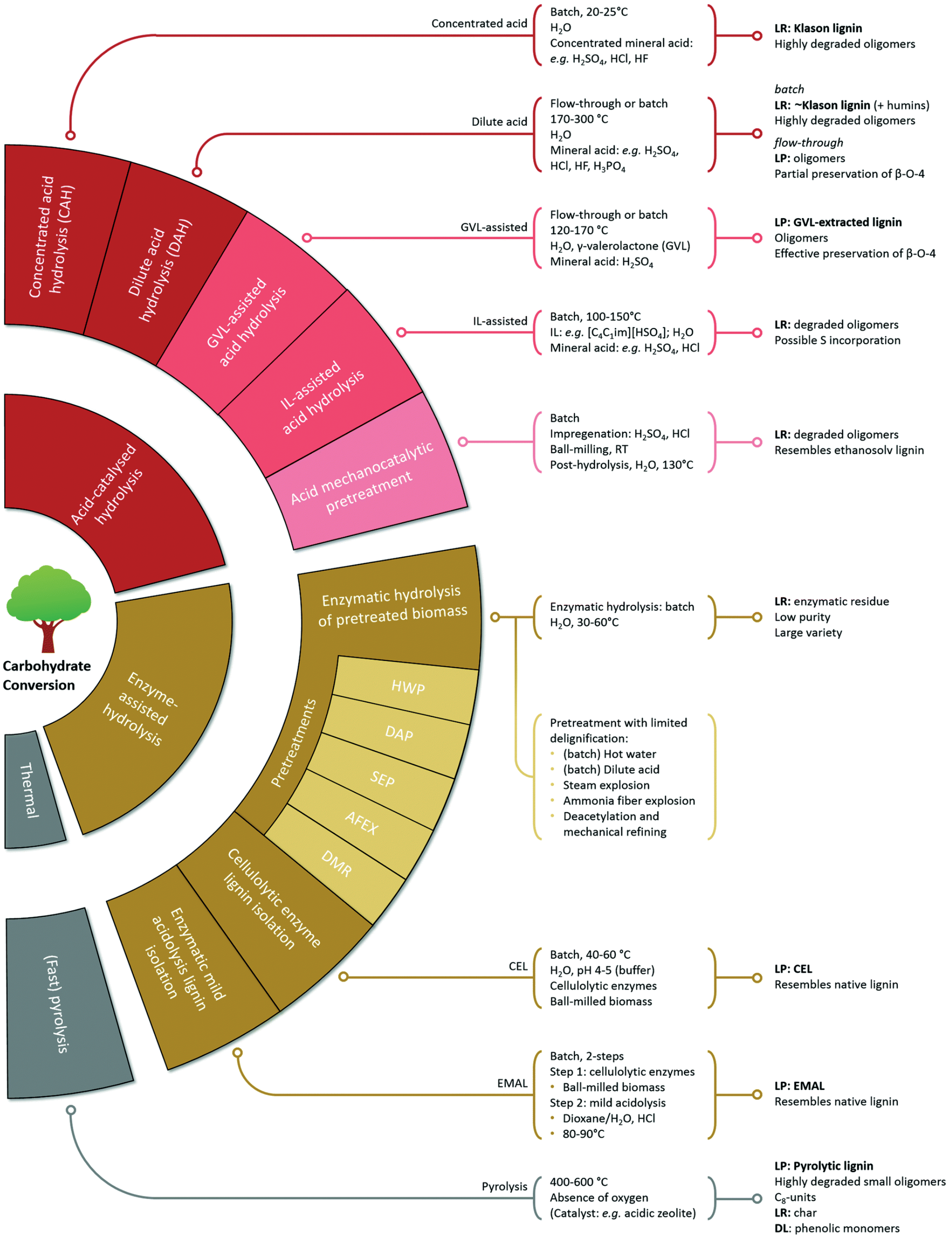

Carbohydrate Conversion
Acid-catalysed hydrolysis
Enzyme-assisted hydrolysis
Thermal
Concentrated acid hydrolysis (CAH)
Dilute acid hydrolysis (DAH)
GVL-assisted acid hydrolysis
IL-assisted acid hydrolysis
Acid mechanocatalytic pretreatment
Enzymatic hydrolysis of pretreated biomass
Pretreatments
HWP
DAP
SEP
AFEX
DMR
Cellulolytic enzyme lignin isolation
Enzymatic mild acidolysis lignin isolation
(Fast) pyrolysis
Concentrated acid
Batch, 20-25°C
H_2O
Concentrated mineral acid: *e.g.* H_2SO_4, HCl, HF
LR: Klason lignin
Highly degraded oligomers
Dilute acid
Flow-through or batch
170-300 °C
H_2O
Mineral acid: *e.g.* H_2SO_4, HCl, HF, H_3PO_4
batch
LR: ~Klason lignin (+ humins)
Highly degraded oligomers
flow-through
LP: oligomers
Partial preservation of β-O-4
GVL-assisted
Flow-through or batch
120-170 °C
H_2O, γ-valerolactone (GVL)
Mineral acid: H_2SO_4
LP: GVL-extracted lignin
Oligomers
Effective preservation of β-O-4
IL-assisted
Batch, 100-150°C
IL: *e.g.* $[C_4C_1im][HSO_4]$; H_2O
Mineral acid: *e.g.* H_2SO_4, HCl
LR: degraded oligomers
Possible S incorporation
Batch
Impregenation: H_2SO_4, HCl
Ball-milling, RT
Post-hydrolysis, H_2O, 130°C
LR: degraded oligomers
Resembles ethanosolv lignin
Enzymatic hydrolysis: batch
H_2O, 30-60°C
LR: enzymatic residue
Low purity
Large variety
Pretreatment with limited delignification:
(batch) Hot water
(batch) Dilute acid
Steam explosion
Ammonia fiber explosion
Deacetylation and mechanical refining
CEL
Batch, 40-60 °C
H_2O, pH 4-5 (buffer)
Cellulolytic enzymes
Ball-milled biomass
LP: CEL
Resembles native lignin
EMAL
Batch, 2-steps
Step 1: cellulolytic enzymes
Ball-milled biomass
Step 2: mild acidolysis
Dioxane/H_2O, HCl
80-90°C
LP: EMAL
Resembles native lignin
Pyrolysis
400-600 °C
Absence of oxygen
(Catalyst: *e.g.* acidic zeolite)
LP: Pyrolytic lignin
Highly degraded small oligomers
C_8-units
LR: char
DL: phenolic monomers

Figure 11 Carbohydrate-first lignocellulosic biomass pretreatments

Reprinted by permission from **ROYAL SOCIETY OF CHEMISTRY**: **Chemical Society reveiws**, Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading, W. Schutyser, T. Renders, S. Van den Bosch, S.-F. Koelewijn, G. T. Beckham and B. F. Sels**, COPYRIGHT** (2018)

1.4.1.1 Acid catalyzed carbohydrate conversion

Concentrated acid hydrolysis (CAH) is a traditional technique used to produce monosaccharides from structural lignocellulosic carbohydrates. The method involves the partial depolymerization of polysaccharides to carbohydrate oligomers in aqueous solution after subjecting untreated biomass to a highly concentrated solution (e.g. 72 wt% H_2SO_4, HCl or H_3PO_4) of mineral acids at room temperature. To obtain high yields (80-100%) of sugar monomers, a post hydrolysis step is carried out at high temperature (>100 °C) using dilute acids (0.5–5 wt%).[84] Significant degradation of the lignin fraction occurs during this process through a combination of repolymerization reactions and acid-catalyzed cleavage of ether linkages. Consequently, a Klason lignin residue is obtained which is comprised of condensed insoluble lignin fragments. Interestingly, production of Klason lignin is a well-known gravimetric technique for the determination of so-called Klason lignin content.[85] In addition to this Klason lignin, acid soluble lignin (ASL) which is a minor fraction of lignin can be obtained spectrophotometrically.[83] The main components of ASL include oxygenated oligomers and monomers. Some of the limitations of this method of simple sugar production using concentrated acid are linked to the cost and difficulty of regenerating and recovering mineral acid, as well as the resulting corrosion of the equipment.

Dilute acid hydrolysis (DAH) is another technique related to CAH which can be used for carbohydrate conversion.[86] In this process, an elevated reaction temperature of over 170 ^{o}C is used to compensate for the relatively low concentration of acid which is less than 5 wt%. The hydrolysis of both hemicellulose and cellulose occurs in such an environment. Once again, significant degradation of lignin fraction results from several acid catalyzed reactions, and an insoluble residue with a highly altered structure is recovered. The resulting residue may also contain pseudo-lignin components (i.e.

components that may appear lignin-like), called humins which actually result from carbohydrate degradation.[87] Flow-through dilute acid process (FT-DAP) can be used to reduce the residence time of soluble intermediates by ensuring a shorter time of residence for species which are highly reactive. In such cases, a large percentage of the obtained lignin fraction consists of partially condensed oligomers. The remaining smaller portion comprises of coniferyl alcohol, syringaldehyde, and vanillin.[88]

Recently, combination of certain polar aprotic solvents with water and acid have been shown to promote biomass deconstruction. Notably, γ-valerolactone (GVL)/water mixture and sulfuric acid are utilized to promote acid-catalyzed reactions during biomass deconstruction.[89] GVL also promotes solubilization of lignin and deconstruction of hemicellulose, even at mild conditions (<120°C) thereby ensuring the biomass solubilizes completely. About 70-90% of the carbohydrate yield is mono- and oligosaccharides and also contains several by-products including furfural and levulinic acid.[90] Water addition to the mixture leads to precipitation of the solubilized lignin. Instead of using additional of water, CO_2 extraction of the water/GVL mixture can also be used to precipitate the lignin by selectively removing the organic solvent.[91,92] As a result, CO_2 separation can be used to produce aqueous solutions with high monosaccharide concentrations of up to 127 gL^{-1}. The mild conditions enabled by GVL-based fractionation also helps preserve β-O-4 ether linkages.[92]

Acidified ionic liquids can also be used as media for carrying out lignocellulose acid hydrolysis.[93] The benefit of ionic liquid is that lignocellulosic biopolymers can be solubilized due to the highly ionic environment breaking hydrogen bonds between cellulose strands. Such a step ensures an easier hydrolysis process due to easier access of glyosidic bonds. As a result, ILs prove to be more efficient in promoting carbohydrate hydrolysis compared to aqueous media.[94] The residue obtained after degradation of lignin acid-catalyzed process contain low content of β-O-4 ether linkages.

Mechanocatalytic depolymerization is another acid-catalyzed technique used in carbohydrate conversion.[95] The method consist of transforming biomass into products including lignin fragments and oligosaccharides which are soluble in water by milling of acid-impregnated biomass. High portions of monosaccharide yields are obtained after post-milling hydrolysis. The acid-catalyzed process of saccharification also leads to formation of a lignin precipitate that is similar to ethanosolv in its content of β-O-4.[96] Investigations have shown that the mechanocatalytic depolymerization is responsible for the massive structural alteration rather than the post-milling hydrolysis procedure. Still, post-hydrolysis can be done in a biphasic system to avoid precipitation. Biphasic systems that include MeTHF/water produce a lignin polymer which a considerably lower molecular weight compared with that of monophasic systems. It was projected that some recondensation is avoided when MeTHF phase is used during the lignin fragments extraction.[97]

1.4.1.2 Enzymatic hydrolysis

Enzymatic hydrolysis is a technique used for the extraction of monosaccharides by the breaking down of carbohydrate polymers within the lignocellulosic biomass. This process requires the isolation of a highly accessible solid residue from biomass which maximizes the presence of residual carbohydrates and can contain a lignin-rich fraction which is normally insoluble in water.[98] Usually, the raw untreated biomass is characterized by certain physico-chemical factors which blocks effective polysaccharide deconstruction. Therefore, the biomass is pretreated to mitigate recalcitrance and hence allow for direct biological deconstruction processes.[99] In some cases, pretreatment fails to significantly delignify the substrate. If that happens, the pretreated biomass is enzymatically hydrolyzed leaving behind a lignin-rich residue. Commonly used pretreatment methods include: ammonia fiber expansion (AFEX), dilute acid pretreatment (DAP), hot water pretreatment (HWP), steam explosion pretreatment (SEP), and DMR (deacetylation and mechanical refining).[100] Meanwhile, methods such as ultrasound and plasma pretreatments are less conventional.[101,102] The process of enzymatic hydrolysis produces residues that contain proteins, ashes and residual carbohydrates which lead to low levels of lignin purity.

To obtain a pure lignin with high β-O-4 and little structural alteration, two enzymatic laboratory scale strategies are utilized. The effectiveness of these methods depend on the extent to which extensively ball milled wood is hydrolyzed by cellulolytic enzymes.[52] Typically, the hydrolysis processes takes about 48 hours. Upon hydrolysis, lignin is extracted from the residual solids using water or dioxane. Subsequently, the lignin precipitated to produce cellulolytic enzyme lignin (CEL). Normally, the resultant CEL is less than 35% of the total raw biomass but the linkages of β-O-4 are preserved. The second extraction procedure involves further hydrolysis using HCl in water/dioxane. The process is referred to as mild acid hydrolysis and is allows the cleaving of soluble lignin and lignin-carbohydrate complexes. The product of the acidic hydrolysis is known as enzymatic mild acidolysis lignin (EMAL) which is about 25% to 65% lignin with incrasing degrees of structural modification as the extraction is increased. In spite of the high level of purity demonstrated by both EMAL and CEL, they are not considered realistic procedures for industrial application. Instead, they are important for research purposes including for the study of native lignin depolymerization.[103]

1.4.1.3 Thermal carbohydrate conversion: pyrolysis

Pyrolysis is a processes where lignocellulosic biomass is thermally decomposed in the absence of oxygen at 400°C to 600°C. The decomposition produces a gaseous substance along with char which is mostly derived from lignin.[104] When the gaseous product is condensed, the resultant liquid is known as bio-oil or pyrolysis oil. The quantity of oil produced can be maximized by employing high heating rates of about 300-1000°C min^{-1} for short durations (residence times of 1-2 seconds). Normally, the resultant oil product is characterized by a low energy content, instability, high water content, and immiscibility with petroleum-based fuels. Nevertheless, the quantity of oil yields can be as high as 75 wt%. Due to the fuel's immiscibility, in-situ or ex-situ catalytic upgrading is normally necessary for producing fuel-compatible products.[105]

The components of pyrolysis bio-oil are both lignin and carbohydrate-derived whereby the former consists of monomeric and oligomeric phenols. The monomers account for up to 20 wt% of the lignin and are made up of a wide array of compounds such as catechols, phenols, and guaiacols.[106] The carbohydrate derived compounds include short aldehydes, furfural, acids, and anhydrosugars. Generally, the lignin derived products are easier to precipitate from bio-oil compared to carbohydrates since they are more hydrophobic. However, the phenolic monomers mostly remain in the liquid phase even after the precipitation process. The resultant precipitate is referred to as pyrolytic lignin and has generally undergone significant degradation and condensation. The pyrolytic lignin consists of DP 4-9 short oligomers which are built from C_8 units instead of C_9 as the C_3 side chain is prone to degradation.[107]

1.4.2 Fractionation based on delignification

The second class of lignocellulosic biomass fractionation is based on extraction and purification of lignin or lignin-first methods. Figure 12 presents a summary of the methods that will be discussed.

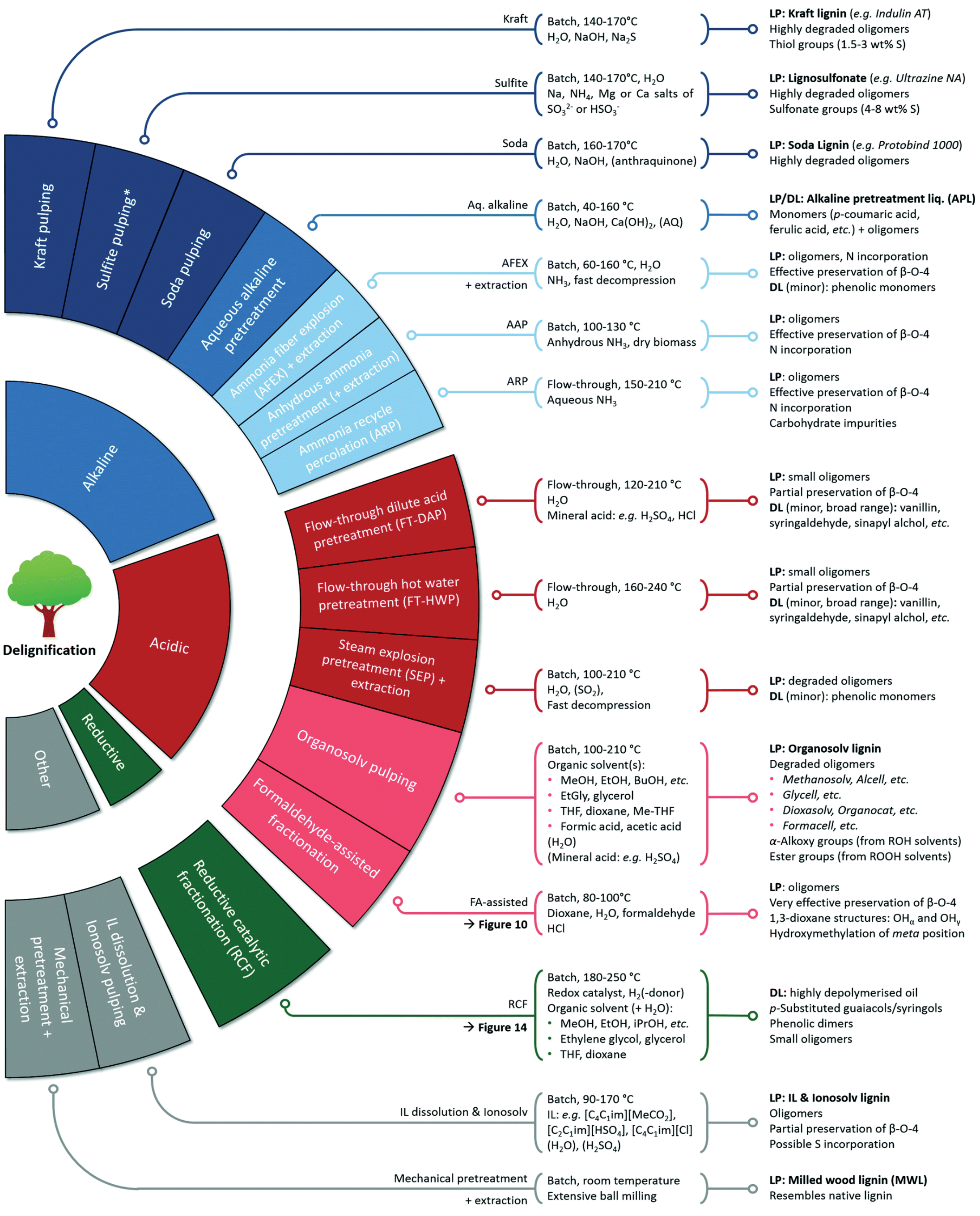

Delignification
Alkaline
Acidic
Reductive
Other
Kraft pulping
Sulfite pulping*
Soda pulping
Aqueous alkaline pretreatment
Ammonia fiber explosion (AFEX) + extraction
Anhydrous ammonia pretreatment (+ extraction)
Ammonia recycle percolation (ARP)
Flow-through dilute acid pretreatment (FT-DAP)
Flow-through hot water pretreatment (FT-HWP)
Steam explosion pretreatment (SEP) + extraction
Organosolv pulping
Formaldehyde-assisted fractionation
Reductive catalytic fractionation (RCF)
IL dissolution & Ionosolv pulping
Mechanical pretreatment + extraction
Kraft
Batch, 140-170°C
H_2O, NaOH, Na_2S
LP: Kraft lignin (*e.g. Indulin AT*)
Highly degraded oligomers
Thiol groups (1.5-3 wt% S)
Sulfite
Batch, 140-170°C, H_2O
Na, NH_4, Mg or Ca salts of SO_3^{2-} or HSO_3^-
LP: Lignosulfonate (*e.g. Ultrazine NA*)
Highly degraded oligomers
Sulfonate groups (4-8 wt% S)
Soda
Batch, 160-170°C
H_2O, NaOH, (anthraquinone)
LP: Soda Lignin (*e.g. Protobind 1000*)
Highly degraded oligomers
Aq. alkaline
Batch, 40-160 °C
H_2O, NaOH, $Ca(OH)_2$, (AQ)
LP/DL: Alkaline pretreatment liq. (APL)
Monomers (*p*-coumaric acid, ferulic acid, *etc.*) + oligomers
AFEX
+ extraction
Batch, 60-160 °C, H_2O
NH_3, fast decompression
LP: oligomers, N incorporation
Effective preservation of β-O-4
DL (minor): phenolic monomers
AAP
Batch, 100-130 °C
Anhydrous NH_3, dry biomass
LP: oligomers
Effective preservation of β-O-4
N incorporation
ARP
Flow-through, 150-210 °C
Aqueous NH_3
LP: oligomers
Effective preservation of β-O-4
N incorporation
Carbohydrate impurities
Flow-through, 120-210 °C
H_2O
Mineral acid: *e.g.* H_2SO_4, HCl
LP: small oligomers
Partial preservation of β-O-4
DL (minor, broad range): vanillin, syringaldehyde, sinapyl alchol, *etc.*
Flow-through, 160-240 °C
H_2O
LP: small oligomers
Partial preservation of β-O-4
DL (minor, broad range): vanillin, syringaldehyde, sinapyl alchol, *etc.*
Batch, 100-210 °C
H_2O, (SO_2),
Fast decompression
LP: degraded oligomers
DL (minor): phenolic monomers
Batch, 100-210 °C
Organic solvent(s):
• MeOH, EtOH, BuOH, *etc.*
• EtGly, glycerol
• THF, dioxane, Me-THF
• Formic acid, acetic acid
(H_2O)
(Mineral acid: *e.g.* H_2SO_4)
LP: Organosolv lignin
Degraded oligomers
• *Methanosolv, Alcell, etc.*
• *Glycell, etc.*
• *Dioxasolv, Organocat, etc.*
• *Formacell, etc.*
α-Alkoxy groups (from ROH solvents)
Ester groups (from ROOH solvents)
FA-assisted
→ **Figure 10**
Batch, 80-100°C
Dioxane, H_2O, formaldehyde
HCl
LP: oligomers
Very effective preservation of β-O-4
1,3-dioxane structures: OH_α and OH_γ
Hydroxymethylation of *meta* position
RCF
→ **Figure 14**
Batch, 180-250 °C
Redox catalyst, H_2(-donor)
Organic solvent (+ H_2O):
• MeOH, EtOH, iPrOH, *etc.*
• Ethylene glycol, glycerol
• THF, dioxane
DL: highly depolymerised oil
p-Substituted guaiacols/syringols
Phenolic dimers
Small oligomers
IL dissolution & Ionosolv
Batch, 90-170 °C
IL: *e.g.* $[C_4C_1im][MeCO_2]$, $[C_2C_1im][HSO_4]$, $[C_4C_1im][Cl]$
(H_2O), (H_2SO_4)
LP: IL & Ionosolv lignin
Oligomers
Partial preservation of β-O-4
Possible S incorporation
Mechanical pretreatment
+ extraction
Batch, room temperature
Extensive ball milling
LP: Milled wood lignin (MWL)
Resembles native lignin

Figure 12 Lignin-first lignocellulosic biomass pretreatments

Reprinted by permission from **ROYAL SOCIETY OF CHEMISTRY**: **Chemical Society reveiws**, Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading, W. Schutyser, T. Renders, S. Van den Bosch, S.-F. Koelewijn, G. T. Beckham and B. F. Sels**, COPYRIGHT** (2018)

1.4.2.1 Acid delignification methods

The first technique considered most popular in reducing the amount of hemicellulose in biomass is dilute acid pretreatment (DAP).[108] Upon release of the lignin components in DAP, they are partly dissolved in the hot acidic water. However, they rapidly condense and deposit on the surface of biomass when handled in batch. Although the lignin content is hardly affected by batch-DAP, the structure of the lignin is greatly altered. In contrast, the lignin structure alterations together with the redeposition process are reduced when operating in flow through mode (FT) due to the removal of soluble lignin components from the heated zone. As such, FT-DAP is a more effective technique for extracting hemicellulose and lignin compared to the batch mode.[109] The extracted hydrolysate consists of small quantities of lignin monomers, lignin oligomers, and hemicellulose monomers and oligomers. The oligomeric fraction's β-O-4 linkages can also be partially preserved. The precipitation process in FT-DAP is ineffective due to existence of oxygenated compounds and low molecular weight products. As a result, the total separation of hydrolysate and lignin is more difficult.[31]

Another common approach of biomass delignification is autohydrolysis which is also referred to as FT-hot water pretreatment (FT-HWP). The method is similar to FT-DAP but without any added acid. The sources of acidity during the process include: extracted organic acid such as acetate and increased dissociated water at higher temperatures.[110] Similar to dilute acid pretreatment (DAP), FT-HWP features acid-catalyzed condensation and acidolysis during lignin extraction although to a lesser degree. Eventually, the resulting lignin contains oligomeric components with some of the β-O-4 linkages preserved. Almost 30 different small quantities of monomeric phenols are also extracted.[111] The main compounds in the phenols include; syringaldehyde, vanillin, sinapyl alcohol, and p-hydroxybenzoic acid.[112] It is also crucial to point out that oxidative degradation is responsible for the production of acids related to syringaldehyde and vanillin. Overall, employing FT-HWP is just as difficult as using the FT-DAP in the process of lignin isolation from biomass.

Steam explosion treatment (SEP) is the third technique that can be employed for acidic delignification of biomass. The method uses a combination of DAP and ammonia fiber explosion/expansion (AFEX). It features an autohydrolysis process in which pressurized steam is used to treat the biomass. The mixture is then explosively depressurized, which ensures that the lignocellulose matrix is opened.[113] The pressure also alters the lignin fibrous structure and disrupts the physical ordering of its components. A combination of hemicellulose-derived compounds and lignin are extracted. Although the severity varies, in most cases, the lignin undergoes 50-100% β-O-4 linkages loss which translates to moderate to severe degradation.

In all the three delignification techniques analyzed, pure water is used as the main media. However, the techniques can be made more efficient using organic solvent media. This idea is illustrated in organosolv pulping in which water, mineral salts and organic solvents are all combined and used in the treatment of biomass.[114] Usually, the organic medium increases solubility of lignin thereby ensuring better extraction compared to HWP/DAP. After pulping, a precipitation can be performed to separate lignin from hemicellulose within pulping liquor which then leads to formation of organosolv lignin. As a result, organosolv pulping is an efficient procedure which facilitates the fractionation of lignocellulose into three main components including a lignin precipitate, solid cellulose pulp, and an aqueous solution of hemicellulose oligomers.[115] The solvents which can used in the fractionation include ketones (MIBK, acetone), organic acids (acetic acid, formic acid), cyclic ethers (dioxane, THF), polyols (glycerol, glycol, ethylene), and alcohols (butanol, ethanol, methanol). Alcohols with low boiling points are the most popular due to low cost and easy recovery. Besides, organosolv pulping can work with or without acid catalysts. Still, regardless of the technique used, lignin undergoes some condensation and depolymerization that is acid-catalyzed which ends up producing oligomeric fragments. Still, the severity of the process affects the degree of lignin structural alteration.[116] Only small quantities of residual β-O-4 motifs are obtained from industrial organosolv process including the Alcell (alcohol pulping and recovery) method (which involves pulping using 195 °C heated aqueous ethanol).[117]

1.4.2.2 *Alkaline delignification methods*

Lignin solubilization and biomass delignification can also be promoted by alkaline media. In fact, such mixtures are often used in the pulping industry. The primary process of pulping is Kraft pulping which produces more than 90% of all chemical pulps. The global dominance of Kraft pulping is due to (i) the low energy demand since the process is self-sufficient in terms of energy, (ii) the ability to recover the pulping chemicals and (iii) the excellent quality of the pulp produced.[118] In Kraft pulping, the middle lamella which is the most abundant source of lignin in the plant is usually removed in high alkaline conditions. Though the hemicellulose wood components are highly modified during pulping, the wood cellulose remains unchanged due to its high resistivity to the harsh alkaline conditions. This process consists of heating wood in an aqueous solution containing NaOH and Na_2S (White liquor). This liquor has a unique feature, which is that the presence of HS^--ions improves the lignin depolymerization and delignification. This occurs without simultaneously speeding up carbohydrate solubilization. Nevertheless, severe lignin repolymerization and degradation reactions are induced due to the harsh alkaline conditions.[119] Black liquor, which is mostly incinerated to recover the pulping chemicals and energy, is the spent processing liquor, containing mostly solubilized condensed lignin. Its combustion helps meet the energy demand in the pulp mill. Acid-induced precipitation can also be used to further isolate the solubilized lignin from the black liquor.[120] The precipitate that is highly condensed and contains low amounts of β-O-4 residual linkages is referred to as Kraft lignin. It is important to note that Kraft lignin can complicate downstream valorization since it includes sulphur in the form of thiol groups.

The efficiency and versatility of the Kraft process have led to a drastic drop in the market share of sulfite pulping which is the 2nd most common chemical pulping method after the Kraft process.[121] Sulfite pulping solubilizes the hemicellulose to oligomeric with some monomeric compounds. Sulfite pulping can work in PH- neutral, acidic or alkaline conditions, the conditions of which are controlled by using (bi) sulfite salts. The sulfonate groups that are present enhance water solubility irrespective of the pH. Lignosulfonates are produced, which are highly condensed and contains higher contents of sulfur (4-8 wt %) than Kraft lignin. Through processes such as ultrafiltration, precipitation or extraction, isolation of these lignosulfonates can be done while being conjugated to counter ions that include Na^+, Mg^+, NH_3^+, and Ca^{2+}.[119]

Soda pulping is another traditional process of pulping, which is also related to Kraft pulping. However, Na_2S is not implemented in soda pulping, and alkaline depolymerization (NaOH) occurs less efficiently due to lack of a primary nucleophile.[118] Even so, other competing reactions have a more considerable extent of occurrence. This pulping method has historically been used to produce pulps from biomass such as straw, flax, miscanthus, bagasse, sugarcane among others which are non-woody. Apart from lower contents of lignin, non-woody biomass also has high portions of alkali-labile linkages of ester and a structure that is more open. Addition of anthraquinone (AQ) can be applied to increase the efficiency of soda pulping. Unlike Kraft and sulfite pulping, soda (-AQ) pulping results in the acquisition of sulfur-free lignin with a low content of β-O-4. Cleavage of ether linkages is reduced by AQ, while at the same time, degradation of carbohydrates is limited. The isolation of the soda (-AQ) lignin can be done through precipitation.[31]

Aqueous alkaline pretreatment is a pretreatment process, which is related to soda pulping. However, unlike soda pulping, it involves a lower severity of treatment due to the lack of need to completely remove the lignin in a biorefinery. An alkaline pretreatment liquor (APL) can be obtained by extracting ca. 55 wt% of initial lignin. The residual solids can be washed with water, removing another 35wt% of lignin. The APL features monomeric phenols like vanillic acid, ferulic acid and p-coumaric acid obtained from breaking ester linkages through hydrolysis. [122]

Complementing the NaOH-based techniques, there are other various methods of alkaline fractionation that make use of ammonia. While preserving the carbohydrates, liquid ammonia can redistribute or solubilize lignin. Due to its high volatility, ammonia can also be easily recovered. Ammonia fiber expansion/explosion (AFEX) is the most well-known of these techniques.[123] It involves reacting wet/moist biomass with ammonia. The reaction is done under elevated pressure. Both hydrolysis and ammonolysis (of low LCC) are induced, and heat is generated. Ester linkages are also produced thus leading to partial solubilization of the lignin polymer. The explosive and rapid pressure released subsequently evaporates ammonia thus redistributing hemicellulose and lignin thus opening up the structure of the biomass. Even though AFEX aids to the subsequent extraction of lignin (with an alkaline or organic solution), it cannot by itself fractionate the constituents of biomass, as it typically leaves in place the hemicellulose and only extracts about 50-60% of the lignin.[123] Small phenolic monomers are also produced, which include acids (e.g., p-coumaric and vanillic acid), aldehydes (e.g., syringaldehyde and vanillin) and amides.

Anhydrous ammonia pretreatment (AAP) is another process closely related to AFEX.[91] Since liquid anhydrous ammonia can penetrate cellulose fibers, it is a known cellulose swelling agent. A cellulose-ammonia complex which leads to an altered crystalline structure is formed after ammonia is removed. The restructured cellulose (a C_{III} polymorph) is more readily hydrolyzed enzymatically than the natural polymorphic form (C_I). The AAP biomass should have a very low content of moisture since water prevents the C_{III} polymorph formation. The C_{III} polymorph formation is one of its significant dissimilarity with AFEX, the second being the absence of an explosive decompression stage. The constituents of solubilized biomass (mainly lignin) are instantly extracted by the ammonia retained under high pressure in the liquid state. This is termed extractive ammonia pretreatment (EAP). It has been shown that EAP enables extraction of 44% of lignin from corn stover while causing just minor degradation of the lignin structure.[124] Alternatively, controlled ammonia evaporation after AAP, followed by lignin extraction under mild conditions using aqueous NaOH solution can be performed. Higher quantities of lignin can be extracted (up to 60%) from corn stover using this variant. This usually is a more practical option with respect to lignin extraction and ammonia recovery. For both processes, β-O-4 linkagess remain unchanged with nitrogen being incorporated as hydroxycinnamoyl amides (feruloyl and coumaroyl amides).[91]

Apart from the AFEX and AAP, ammonia recycled percolation (ARP) is another technique of ammonia-based delignification. In this process which is flow-through, continuous extraction of lignin is carried out by a solution of aqueous ammonia.[125] Delignification of up to 80% for corn stover can be achieved. Apart from the high degrees of delignification, extraction of a large portion of the hemicellulose from the biomass is possible (usually up to 50-60%).[126] Precipitation of the solubilized lignin can be accomplished through solvent evaporation. Carbohydrate impurities can be contained in considerable amounts in the resulting precipitate (as high as 20%).[127] These impurities can be removed, almost entirely through a procedure that involves hydrolysis catalyzed by a mild acid. Further, it has been demonstrated that preservation of β-O-4 linkages can be possible during percolation of ammonia. However, this was only shown for low delignification levels. Like any other ammonia-based method of fractionation, a small nitrogen amount is incorporated into the lignin and carbohydrates.

1.4.2.3 Other delignification techniques

Ionic liquids (ILs) influence the choice of fractionation approach whether ionosolv pulping or ionic liquid dissolution. Often, the extraction of hemicellulose (ionosolv pulping) or dissolution of the lignocellulose substrate (IL dissolution) depends of the IL applied.[128] In ionosolv pulping, aqueous-organic or just plain organic solutions are used as anti-solvents to help precipitate the cellulose from the product mixture. This step is done prior to precipitation of lignin, which precipitates the solubilized cellulose, and leads to decrystallizing cellulose which is beneficial in helping in the downstream processing.

Usually, the cellulose fraction is preserved as solid pulp while hemicellulose and lignin are solubilized. Partial β-O-4 cleavage often occurs during IL-based fractionation which is then succeeded by repolymerization. The degree of β-O-4 cleavage significantly depends on the anion type including phosphate, acetate, or sulfate that can function as nucleophiles. The contributions of cations are negligible. Sulfur anions such as sulfamate, sulfonate, and sulfate can also be added to help incorporate sulfur. The other considerations in choosing the nature of ionic liquids include their ease of recuperation, toxicity, and cost.

Biomass fractionation can also be facilitated by use of mechanical pretreatment which is crucial in separation of milled wood lignin (MWL).[129] MWL begins with a thorough ball milling at about 24 °C which is then succeeded by extraction of lignin using water or dioxane or any other organic solution.[31] The low temperature and lack of chemical catalyst ensures that the resulting lignin has a structure that is virtually unchanged from its native counterpart. On an industrial scale, the isolation of MWL is not effective in spite of the high β-O-4 content. However, MWL generally leads to a low delignification (< 35%) unless the time and milling spans several days or weeks. As such, the technique is suitable for experimental and analytical purposes due to its ability to provide a surrogate lignin that maintains its native structure.

Lastly, lignin can be isolated from raw lignocellulosic biomass through pulp bleaching through oxidative delignification. The process can be done either in alkaline conditions (using lime or aqueous NaOH), or acidic conditions (in paracetic acid or in acetic acid). Oxygen or hydrogen peroxide can be used as oxidants. Eventually, the oxidation processes, which usually occurs simultaneously to base and acid-catalyzed and acid catalyzed reactions, transforms lignin into a low molecular weight mixture that consist of aliphatic carboxylic acids which are ring-opened structures, quinone, and phenolics. However, the structure of the vast majority of the lignin extracted through bleaching processes is still unknown but at least some of the lignin likely undergoes condensation. [31]

Even though all the above-mentioned methods focus on the extraction of lignin from biomass, as lignin-first pretreatment processes, the quality of isolated lignin from such methods is usually low. The low quality refers to the low amount of monolignols can be produced by depolymerization of the resulting isolated lignin. This is mainly because the β-O-4 linkages are not entirely preserved which results in partial repolymerization (condensation) of extracted lignin.[130] Reductive catalytic fractionation, which is presented next, is a method that provides high-quality lignin resulting in maximal production of monolignols (based on cleavage of lignin's ether linkages).

1.4.2.4 *Reductive Catalytic Fractionation (RCF)*

Reductive catalytic fractionation has recently emerged as an effective lignin-first biomass fractionation strategy to produce phenolic mono-, di-, and oligo- mers in high yields, which can be used as polymer additives or platform chemicals for production of aromatics or other chemcials.[131] RFC involves the use of heterogeneous redox catalyst to facilitate reductive depolymerization which helps the solvolytic extraction of lignin.[132] This two-step method starts initially with extracting lignin fragments from biomass through solvolysis. These fragments are then stabilized through hydrogenolysis or hydrogenation reaction with a hydrogen donor over a redox-active catalyst.[133] The advantage of the simultaneous extraction and stabilization is in producing near-theoretical yields of monomers based on ether bond cleavage resulting in a relatively narrow product distribution. Furthermore, the sugars are separated in a processable solid form which can be used in pulp and paper industry or for further upgrading as a compatible feedstock of conventional biorefineries.[134] This also enables the separate processing of sugars and lignin for upgrading, which translates to easier downstream processing.

Although similar to organosolv pulping, RCF does not produce a precipitate of high molecular weight lignin but instead releases lignin oil that is highly depolymerized. Studies on RCF are so far performed in liquid-phase batch reactors where the main factors which affect the delignification process include temperature (180°C to 250 °C), reaction time (2 to 12 hr), pressure (5 to 15 MPa), solid catalyst to biomass ratio, reductant (H_2 or hydrogen donor), and choice of the polar protic solvent.[135]

To increase the lignin extraction, harsher RCF conditions can be used but it also increases the solubilization of carbohydrates that might result in formation of degradation products. Milder conditions extract less lignin, but preserve carbohydrates in solid form.[136] The solvents are often alcohols such as isopropanol, ethanol, and methanol. Hemicellulose removal and high extraction of lignin is enhanced by addition of water and co-catalysts such as metal triflates and H_3PO_4.[131] Typical catalysts include supported noble metals or base metals such as nickel or copper. Substituted methoxyphenols are the main products obtained when lignin depolymerization combines with RCF. The choice of catalyst can dramatically affect the structure of produced phenolic chemicals. For example, Ru/C and Pd/C yield either propyl- or propanol-substituted lignin products, respectively.[137] A mixture of the spent catalyst and (holo)cellulose pulp is the remaining solid after the process. Therefore, RCF process viability is mainly determined by the isolation of catalyst from pulp mixture which is a difficult process. Even though the catalyst-pulp isolation can be facilitated by a catalyst bracket and ferromagnetic catalysts, however, separation of catalyst from pulp still remains the main drawback of this technique.[138]

The type of feedstock has a great impact on the effectivity of RCF process. For example hardwoods result in higher yields compared to softwoods or herbaceous biomass. The reason is that the ratio of S to G units is higher in hardwoods, therefore the relative amount of β-O-4 and other ether linkages, which are easier to cleave, is higher compared to C-C bonds.[139] Herbaceous biomass also contains low S/G ratios with higher content of the hydroxycinnamic acids (HCAs), ferulic, and p-coumaric acids.[46] The p-coumaric acids are typically pendant to the syringyl units in lignin, whereas ferulic acid polymerizes through a broad suite of linkages to the lignin polymer, and also to hemicellulose by covalent bonds.[140] In this regard, studies on the RCF of different feedstock reported yields of 50%, 27%, and 21% for birch (hardwood), miscanthus (herbaceous), and pine/spruce (softwood), respectively.[133,141]

Despite the diversity of pretreatment techniques, there is a need for developing a method that can fractionate components of lignocellulosic biomass and preserve each of them with high-quality that can be available for high-yield upgrading.[31] As mentioned in this section, the existing methods either target providing highly digestible sugars or extracting lignin. Taking into account that biorefineries are mainly based on valorizing sugars, lignin is the component that has been sacrificed for intensifying biorefinery productivity through targeting high yields of carbohydrates extraction. Although lignin can be potentially a valuable source of biofuels and bioproducts because of its aromatic structure, the complexity of extracted lignin makes it difficult for further upgrading. Understanding lignin chemistry and its structural alterations during different pretreatments will play a key role in development of new pretreatment methods for full valorization of lignocellulosic biomass. Before continuing with challenges in existing pretreatment methods and the objectives of this thesis, the chemistry of lignin alteration during pretreatment is presented in the next section.

1.5 Lignin chemistry

As mentioned before, the lignin structure is heavily modified due to a combination of depolymerization and repolymerization reactions. However, the specific mechanisms involved greatly depend on the pretreatment process. Below, I briefly present an overview of chemical reactions identified during the lignocellulose or lignin processing based on their effect on β-O-4 ether linkages, which is the most prominent and generally most reactive motif in lignin. The chemistry of the pretreatments that are discussed in this section includes acid-catalyzed, base-catalyzed, oxidative, thermal, and reductive extraction techniques.

1.5.1 Acid-catalyzed lignin

Acidic conditions can catalyze the hydrolysis of ether bonds in carbohydrate polymers. Consequently, it is commonly used to solubilize and depolymerize the cellulose as well as hemicellulose fraction. In contrast with alkaline environments, acidic conditions do not necessarily enhance the extraction and solubilization of lignin. Though the acid medium tends to break the ether bonds of lignin and induce depolymerization, it also catalyzes irreversible repolymerization reactions that can lead to redeposition of extracted lignin.[100]

The cleavage of β-O-4 ether linkages is regarded as the most important reaction in acid-catalyzed lignin deconstruction (Figure 13). One mechanism suggests that removal of the OH-group on the α-position and consequently formation of an intermediate benzylic carbenium ion (**1**) is the first step in acidolysis of β-O-4 linkages.[142] Depending on the acid, the carbenium ion can go through a pathway of two enol-ether structures (**2a** and **2b**). In presence of water, the acid-labile enol-ethers go through a hydrolysis to C_3-ketone-substitued phenolics (**3**), which is known as Hibbert's ketones, and C_2-aldehyde-substituted phenolics (**4**).[142] Carbenium ions along with C_3-ketone-substitued phenolics and C_2-aldehyde-substituted phenolics can go through a repolymerization reaction which results in a condensed lignin.[143]

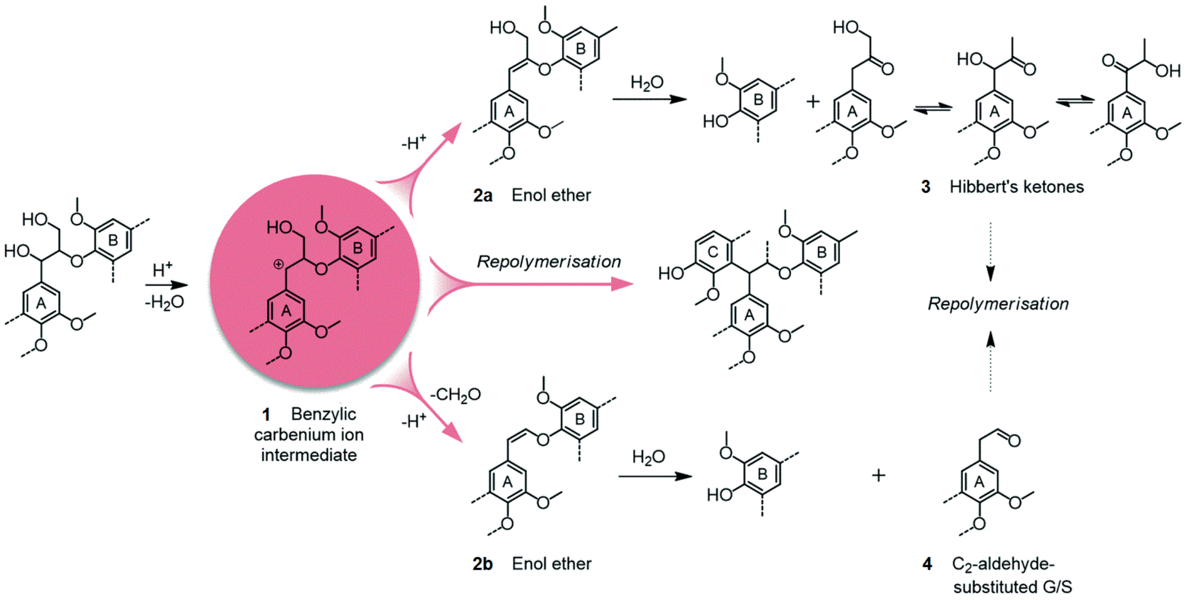

H$^+$
-H$_2$O
1 Benzylic carbenium ion intermediate
-H$^+$
2a Enol ether
H$_2$O
3 Hibbert's ketones
Repolymerisation
Repolymerisation
-CH$_2$O
-H$^+$
2b Enol ether
H$_2$O
4 C$_2$-aldehyde-substituted G/S
A
B
C
HO
OH

Figure 13 Acid-catalyzed lignin chemistry

Adopted by permission from **ROYAL SOCIETY OF CHEMISTRY**: **Chemical Society reveiws**, Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading, W. Schutyser, T. Renders, S. Van den Bosch, S.-F. Koelewijn, G. T. Beckham and B. F. Sels**, COPYRIGHT** (2018)

1.5.2 Base-catalyzed lignin

Because of its medium-polarity, lignin is insoluble in water. However, its solubility in water is increased in basic conditions due to the phenolic OH-group's deprotonation. Consequently, alkaline medium is commonly utilized in delignification processes where extraction of lignin is the main goal such as industrial pulping processes.[144] Other than enabling the fragmentation of lignin through cleavage of β-O-4 motifs and lignin-carbohydrate bonds, alkaline medium also contributes to the solubilization of these fragments.[31] β-O-4 reactivity also depends on the type of linked units. A typical base-catalyzed lignin extraction is shown in Figure 14 based on phenolic (**6**) and non-phenolic (**5**) β-O-4 linkages. Non-phenolic units lead to a slow transformation rate to epoxide intermediate (**7**) while its transformation to phenolic-units (**6**) occurs at higher rates.[145] In presence of a suitable leaving group such as –OH in α-position, phenolic-unit is then converted to quinone methide (**8**) which is likely to undergo a nucleophilic attack.

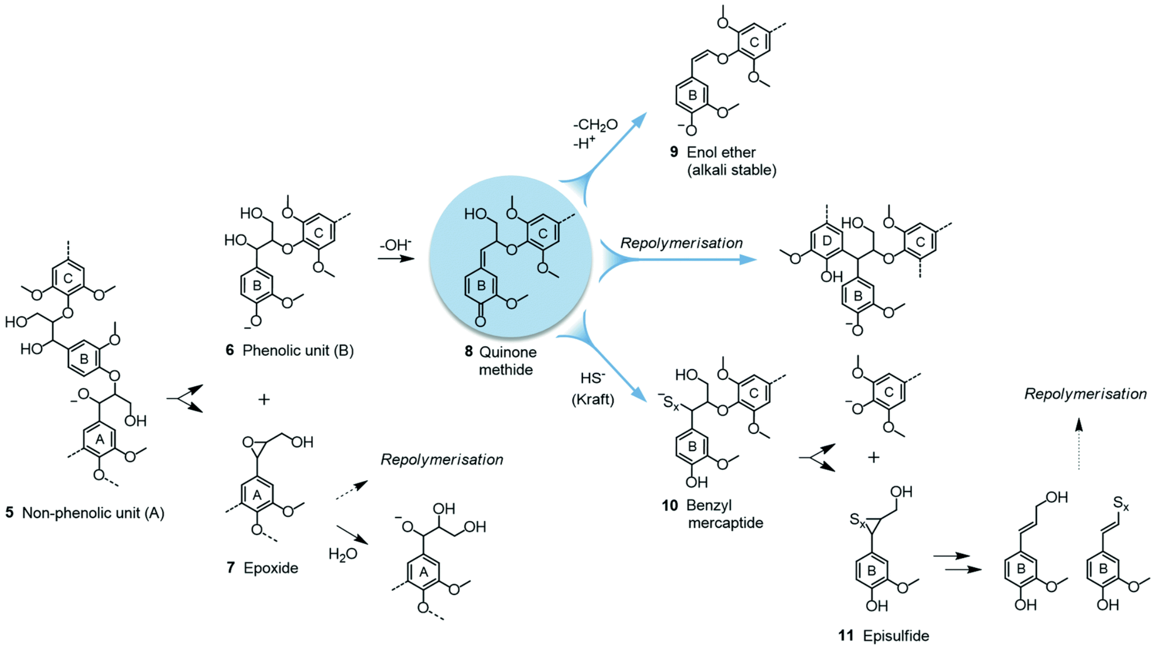

5 Non-phenolic unit (A)
6 Phenolic unit (B)
-OH⁻
8 Quinone methide
-CH₂O
-H⁺
9 Enol ether (alkali stable)
Repolymerisation
HS⁻ (Kraft)
10 Benzyl mercaptide
7 Epoxide
H₂O
Repolymerisation
11 Episulfide
Repolymerisation
A
B
C
D
HO
OH
Sx

Figure 14 Alkaline lignin chemistry. Phenolic rings are labeled with A–D.

Reprinted by permission from **ROYAL SOCIETY OF CHEMISTRY**: **Chemical Society reveiws**, Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading, W. Schutyser, T. Renders, S. Van den Bosch, S.-F. Koelewijn, G. T. Beckham and B. F. Sels**, COPYRIGHT** (2018)

Provided that either –OH or –OR groups are present at the α-position, the first leads to the transformation of phenolic units into a quinone methide. Quinone methide intermediates are susceptible to nucleophilic attacks, which are attributable to such groups' propensity to restore their aromatic structrure.[146] When lignin is depolymerized in the presence of sulfur (e.g. during the Kraft process) episulfide intermediates can form (**11**), which can lead to nucleophilic attack and β-O-4 cleavage in the presence of strong anions such as HS^- (which are common in Kraft pulping).[26] These sulfur containing anions can be converted to coniferyl alcohol which can consequently undergo degradation or condensation.[118] Once carbon-carbon bonds are formed, there is a possibility of quinone methide reacting with a lignin nucleophilic group. As a result, the reaction leads to repolymerization. A third pathway can happen that involves the removal of the γ-CH_2OH group that is promoted by the restoration of the aromaticity of the quinone methid. In cases like Soda pulping or similar processes, this pathway happens even without a strong nucleophile. Furthermore, the results of this pathway are formaldehyde and enol ether motifs (**9**) which are stable in alkali media. Ultimately, while the formaldehyde instigates repolymerization through condensation, the ether bond is cleaved more slowly.[146]

1.5.3 Oxidative lignin depolymerization

Generally, lignin is removed in pulp bleaching to increase the quality of paper products. During the bleaching process, remaining lignin (2–6 wt%), which is strongly colored, is removed to increase the pulp's brightness.[147] Molecular oxygen, chlorine, peroxyacids, hydrogen peroxide, ozone, and chlorine oxide are some of the oxidants that are utilized in the process.[148] Often, lignin oxidation is initiated by electrophilic reactions. This implies that some of the electrophilic species such as OH^+ (found in peroxyacids), Cl^+ (found in chlorine solutions), and oxygen attack the C_β, ortho or para positions. Oxidation with molecular oxygen produces water as its main side product. Nevertheless, despite being readily available, oxygen is a weak oxidizing agent when in its normal state. Consequently, oxidation process under aerobic oxidation regularly involves alkaline conditions for the purpose of ionizing phenolic hydroxyls in lignin to initiate the process.[148]

Alkaline lignin oxidation progresses through a radical reaction mechanism. The process is started by a phenolate ion that undergoes oxidation into phenoxy radicals. The exact and or exhaustive mechanisms for depolymerization of lignin under alkaline aerobic oxidation are not yet fully known. However, the formation of peroxy anions are thought to be a result of oxygen addition to phenoxy structures in C_β, para, or ortho position.[149]. The transformation pathways are found to result into either: formation of oxiranes, formation of p-quinones (**14**) through cleavage of C_4-C_α bonds, formation of muconic acid derivatives (**15**) by cleavage of aromatic rings, or formation of phenolic aldehydes (**12**) via cleavage of Cα–Cβ bond (Figure 15).

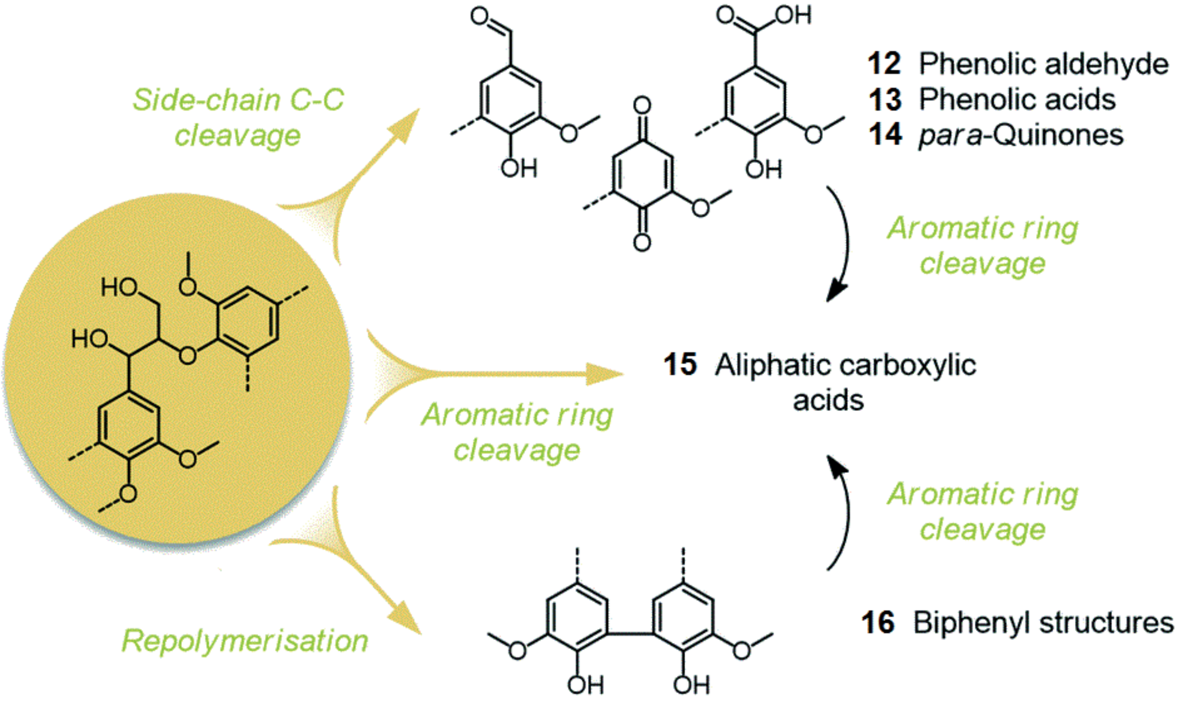

Side-chain C-C cleavage
12 Phenolic aldehyde
13 Phenolic acids
14 *para*-Quinones
Aromatic ring cleavage
15 Aliphatic carboxylic acids
Aromatic ring cleavage
Aromatic ring cleavage
Repolymerisation
16 Biphenyl structures
OH
O
HO
HO

Figure 15 Oxidative lignin chemistry

Adopted by permission from **ROYAL SOCIETY OF CHEMISTRY**: **Chemical Society reveiws**, Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading, W. Schutyser, T. Renders, S. Van den Bosch, S.-F. Koelewijn, G. T. Beckham and B. F. Sels**, COPYRIGHT** (2018)

On the other hand, some other studies have suggested that aerobic alkaline lignin oxidation may not primarily involve oxygen addition, but rather proceeds via second oxidation of the phenoxy radicals.[150] The process is thought to generate a cinnamaldehyde intermediate followed by retro-aldol cleavage of the C_α–C_β bond.

In summary, lignin oxidation happens by either fragmentation of the side-chain which prevents aromatic rings to be cleaved, or the aromatic ring is opened which leads to aliphatic carboxylic acid generation. However, aromatic radicals sourced from the cleavage of C_4–C_α or C_α–C_β bonds are unstable in alkaline conditions. Consequently, they can further transform into aliphatic carboxylic acids (**15**) by a process of aromatic ring cleavage.[148] Mostly, depolymerization of oxidative lignin continues through carbon-carbon bond cleavage routes rather than ether-linkage cleavage. Considering the instability of intermediates in oxidative media, lignin condensation through radical coupling can rapidly occur, which notably produces biphenyl structures (**16**).[149]

1.5.4 Thermal lignin depolymerization

Thermal degradation methods for lignin and lignin derivatives have also been extensively explored. These methods are generally described as pyrolysis or liquefaction approaches. However, due to the complexity involved, a detailed mechanistic understanding of such processes has not yet been achieved. Even general questions such as whether C-C and C-O bonds cleavage reactions occur through homolytic bond cleavage with radicals being created that later condense, or whether the cleavage is concerted naturaly.[151] One example of this debate, is the study of the thermal cleavage of the β-O-4 linkage that has been proposed to primarily undergo both concerted cleavage and homolytic bond dissolution pathways.[152,153] Most of the studies exploring this mechanism have differed in their assumptions and conditions such as lignin model compound, temperature, or pressure. As such, it is crucial to propose that ultimate conclusions in regard to how β-O-4 linkages are thermalized is highly dependent on the techniques being employed.[154] To address a challenge of this complexity, researchers are developing computational techniques and building libraries of lignin intermediates and motifs for the purpose of simulating several reactions networks.[155] Hopefully, the development will contribute to elucidating some of the underlying phenomena in thermal lignin chemistry.

1.5.5 Reductive lignin depolymerization

Frequently, reductive strategies have been applied to lignin depolymerization processes. As mentioned in section 1.4, these strategies always involve either H_2 or an H-donor and a redox catalyst. These processes specifically focus on transforming the β-O-4 and α-O-4 as well as side-chain hydroxyl groups. In all cases, the removal of benzylic OH-groups (i.e. OH_α), cleavage of ether linkages by hydrogenolysis, and the expulsion of OHγ-group is observed.[156] Significantly, small oligomeric particles and substituted methoxyphenols (**17**) seem to be generated by these reactions (Figure 16). Importantly, additional hydrogenolysis occur as secondary reactions, which can lead to formation of cycloalkanes or cyclohexanols.[31] Hydrogenation plays an especially important role, in quenching carbonyl and alkenyl reactive groups that are key condensation precursors. In contrast to acidic and base media, these quenching reactions prevent repolymerization. However, little or no carbon-carbon bond cleavage occurs during these reductive processes. Therefore, the degree of depolymerization is predicted by the amount of cleavable inter-unit ether linkages that are present in the lignin polymer prior to reductive processing.[157]

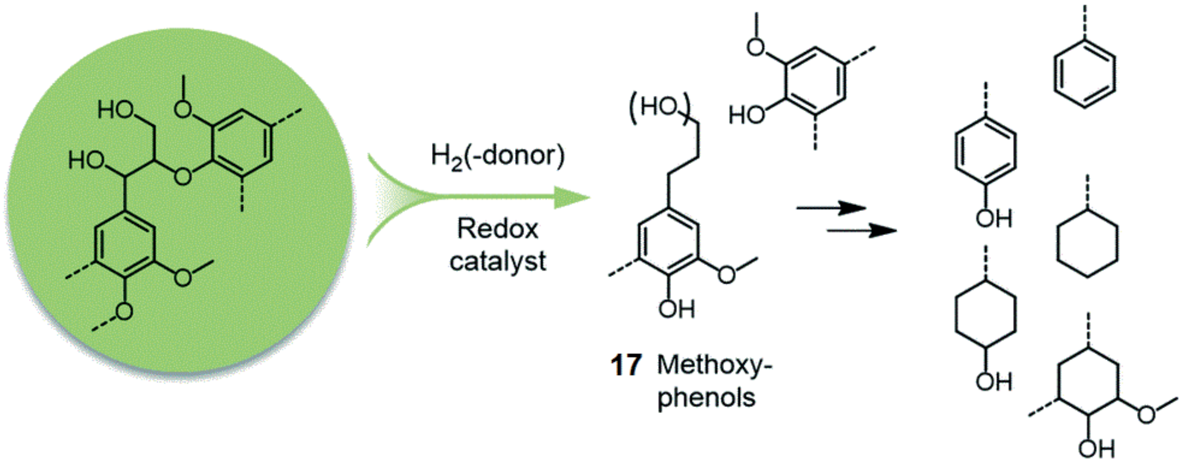


Figure 16 Reductive lignin chemistry

Adopted by permission from **ROYAL SOCIETY OF CHEMISTRY**: **Chemical Society reveiws**, Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading, W. Schutyser, T. Renders, S. Van den Bosch, S.-F. Koelewijn, G. T. Beckham and B. F. Sels**, COPYRIGHT** (2018)

1.6 Challenges in lignocellulosic biomass pretreatment for lignin valorization

Current biomass deconstruction and valorizing schemes, which mainly include pulp and paper processes and emerging biorefineries, generally feature a lignin separation and modification stage, as these processes view lignin as an impediment to the upgrading of the cellulose and hemicellulose fractions.[75] Kraft and Sulphite pulping are the dominant technologies in pulp and paper processing. On the other hand, many biorefinery processes involve treating the raw biomass with mineral acids at high temperature in water, ionic liquids, or various organic solvents. While these strategies are effective at removing lignin, they negatively affect its depolymerization into its constituent monomers post-separation. During the extraction, the benzylic alcohols of lignin can be easily protonated and eliminated, producing reactive benzylic carbocations that can undergo a subsequent electrophilic aromatic substitution with nearby electron rich guaiacyl and syringyl subunits.[142] Some studies have also depicted the formation of unsaturated guaiacyl or syringyl propene intermediates that similarly condense.[158] Once these C–C bonds are formed, their stability leads to low monomer yields after extraction and depolymerization by hydrogenolysis (generally < 5-10% of the original Klason Lignin content).[159] As mentioned in this chapter, the separation of lignin from cellulose and hemicellulose is essential prior to their use in many applications, but this is particularly important in pulp and paper processes (where pure cellulose is required) and before the enzymatic hydrolysis of cellulose (where lignin can suppress yields of glucose). Therefore, to valorize lignin, it is vital to develop a fractionation strategy that efficiently separates it from the cellulose and hemicellulose components of biomass while preventing its condensation. This section provides a brief review of challenges in valorizing lignin in respect to the current status of research and includes discussion on preserving high-quality lignin, developing a fractionation method with high isolation yields, and obstacles in quantification of chemical groups in lignin's structure.

1.6.1 Preventing structural lignin degradation during pretreatment

The primary challenge in lignocellulosic biomass pretreatment is to develop an industrially feasible process that causes minimum alterations in the structure of isolated lignin.[134] This should be achieved by:

1- Preventing significant lignin depolymerization which will keep the original structure of lignin largely untouched.

2- Preventing the formation of reactive intermediates leading to lignin repolymerization.

3- Genetically modifying the biomass.[31]

The primary approach in respect to the first principle is to preserve the β-O-4 linkages in lignin. Therefore, methods based on harsh conditions either acidic or alkaline, are not suitable for this purpose. For example, classic pulping (alkaline) or carbohydrate hydrolysis (acidic), that severely cleave β-O-4 linkages, generate reactive intermediates which will go through an irreversible repolymerization. On the other hand, milder conditions such as ammonia-based methods tend to limit the cleavage of the β-O-4 linkages. The second option is to create a media that increases the solubilization or saccharification of biomass, like using GVL as a solvent. This way utilizing harsh conditions such as high temperatures or using strong acid/alkaline media can be avoided.

The second approach is to prevent the reactive intermediates from being formed and leading to carbon-carbon bonds which are highly stable. One way is to physically remove these intermediates from the reaction media by operating the reaction in a flow-through mode rather than batch mode. The drawback in flow-through mode is the higher solvent-to-solid ratio compared to batch mode.[31] This results in energy-intensive downstream processing due to solvent recovery and product purification of a more diluted product stream. The intermediates can be also removed chemically by quenching. For example in RCF, the reducing agent transforms these intermediates to end products that are stable.

The genetic modification of biomass targets designing lignins that can undergo a lower degree of repolymerization. For this purpose, the lignin structure should contain less branching, less lignin-carbohydrate cross-linking, higher hydrophilicity and also a higher ratio of S to G units.[160] On the other hand, genetic modification can have adverse impacts on growth rates of the plant which affects the efficiency of the biorefinery. Therefore, deep understanding of characteristics and also comprehensive lifecycle assessment of the genetically modified plants is necessary for further development of this approach.[161]

1.6.2 Fractionation efficiency

The efficiency of pretreatment (or fractionation) is defined by the yield, purity, and quality of all the isolated fractions including cellulose, hemicellulose, and lignin. This translates as the amount of extracted component compared to its content in the input biomass, the amount of impurities in the final isolated fraction and complexity of required steps for further upgrading. As discussed previously, the existing pretreatment methods mainly focus on the yield and quality of either extracted lignin or carbohydrates. Ideally, an efficient fractionation method should maximize the efficiency and quality of all three fractions or to provide a right balance, rather than omitting one component to increase the yields of other components.

1.6.3 Identifiation and quantification of lignin's structural features

Structural features of lignin such as the content of β-O-4 (or ether) linkages or OH-groups are indicators of the lignin's reactivity. Thus, measuring such structural features is extremely useful for evaluating the quality of isolated lignin.

2D ^{1}H-^{13}C Heteronuclear Single Quantum Coherence Nuclear Magnetic Resonance Spectroscopy (2D-HSQC NMR) is the technique which is used most often to evaluate the β-O-4 content.[146] However, this method should be used in a comparative context rather than an indicator of the absolute amount of these, due to the difficulties of existing technique for performing quantitative analyses.[162] There are destructive analytical methods that can be used as complementary to 2D-HSQC NMR such as thioacdolysis[163] or nitrobenzene oxidation[164]. These methods also aim to calculate the amount of β-O-4 linkages (reactivity of lignin) and the possible yield of volatile products (monomers).

Applying each technique to a range of different isolated lignins can provide valuable comparative results.[127,165,166] Even though such comparative studies present similar observations on the quality of isolated lignins, but on the other hand, it is difficult to provide a generalized comparison due to the numerous variables involved in the fractionation process and the analytical method. It is worth mentioning that even several biomass samples from the same species can have structural differences.[77,167] In general, the quality of extracted lignin is highly dependent on the fractionation method, its severity and the source of lignocellulosic biomass.

Finding a method for simultaneous identifcaition and quantification of lignin's structural features by NMR can play an important role for providing comparative results based on valid quantification of an isolated lignin's quality. The main reason for the failure of standard 2D-HSQC NMR to provide quantitative results is that the obtained cross-peaks are not proportional to the concentration of chemical groups in the sample due to the different senistivities of the different chemical groups depending on their position in an oligomeric structure.[162] These differences are caused by resonance offsets, T2 relaxation differences, imperfect pulses, homonuclear coupling, and coupling constant deviations.[31,168] A suggested pulse sequence for quantitative HSQC NMR (Q-HSQC NMR) has been shown to restrain the $^1J_{C\text{-}H}$ dependence of signals.[169] However, this suggestion still fails to provide accurate quantification due to the errors from T2 relaxations and resonance offsets.[170]

1.7 Objectives

The general aim of the work in this thesis is to develop a method for valorizing lignin in a process that is compatible with existing biorefineries. In this chapter, I have explained the various methods for pretreatment of lignocellulosic biomass and a review of the literature on proposed chemical mechanisms involved in each process. Based on that, the typical challenges in lignin valorization was presented which include preserving the strucutural features of lignin during the fractionation process, developing a high-efficiency fractionation method, and addressing difficulties in performing quantitative NMR analysis on isolated lignin. The goal of my Ph.D. studies was based on exploring these issues to provide a proper solutions. Therefore, the objectives of this thesis are divided into three parts as presented below.

1.7.1 Preventing lignin condensation during the pretreatment process

As mentioned above, lignin condensation could be prevented by applying three principles which can limit the formation of inter-unit carbon-carbon linkages and repolymerization (condensation) of lignin. Chapter 2 presents a study on the possibility of avoiding the condensation of lignin by preventing the formation of C-C linkages by adding formaldehyde, which acts as a protection group, during the pretreatment of lignocellulosic biomass. This leads to the reaction of formaldehyde with an α,γ-diol structure on the side chain of extracted lignin to form a 1,3-dioxane structure, consequently stabilizing the lignin. The results show a near theoretical yield of monomers from the stabilized lignin upon depolymerization by hydrogenolysis. However, the undesirable reaction of formaldehyde with lignin subunits results in partial methylation of produced monomers. This translates to the distribution of the yield among a broader range of final products which is not favorable due to the more complex downstream processing required for further upgrading. Furthermore, formaldehyde also reacts with cellulose and hemicellulose. In the case of cellulose, enzymatic hydrolysis yields are dramatically lower compared to conventional methods. However, the yields are partially recovered by acid-treatment of cellulose fraction prior to enzymatic hydrolysis. In the case of hemicellulose, the reaction product is a form of xylose identified as diformylxylose. This product is characterized and can be used as feedstock for further upgrading.

Considering the drawbacks of our primary study on stabilization of lignin, we continued by exploring the utilization of various protecting groups in order to develop a deeper understanding of the chemistry involved in this process and to optimize the process towards higher efficiency and narrower product distribution. These protecting groups include ketones, boronic acids, alkyl carbonates, and other aldehydes. The results show that linear aldehydes can stabilize lignin during pretreatment with near theoretical monomers yield upon depolymerization. Furthermore, we achieved a narrower product distribution in the case of propionaldehyde. Interestingly, using a genetically modified biomass (third principle) results in more than 80% selectivity to one single monomer.

1.7.2 Developing a high-efficiency fractionation process

Developing a fractionation process for large-scale production of lignin's biopolymers is essential for the implementation of such processes in biorefineries and could facilitate the development of new upgrading pathways. Providing a high-efficiency method requires high isolation yields for each fraction along with their high upgrading quality. Based on our discovery on stabilization of lignin and optimization of this process, chapter 3 presents a protocol to fractionate lignocellulosic biomass into its three main constituents in large laboratory-scale with high extraction yields. The results show extraction yields higher than 95% for each biopolymer. Furthermore, each fraction is characterized and their upgrading quality is studied with enzymatic hydrolysis in case of cellulose, and hydrogenolysis for lignin. The hemicellulose fraction is obtained with structural alteration due to its reaction with aldehyde during pretreatment. However, a hydrolysis step can reverse this reaction to produce xylose and furfural.

1.7.3 Simultaneous identification and quantification of structural features of lignin by HSQC NMR

Currently, the depolymerization of the aldehyde-stabilized lignin is done by hydrogenolysis over noble metal catalysts. Even though this destructive method provides valuable information on the amount of the produced monomers, it does not provide sufficient information on the chemical structure of the isolated lignin. Furthermore, the product distribution and overall monomer yields are highly dependent on the process conditions such as reaction temperature, reaction time, and the catalyst.

Chapter 4 presents a non-destructive method for the prediction of the potential yield of monomers from the depolymerization of isolated lignin. A new 2D heteronuclear single quantum coherence nuclear magnetic resonance (2D-HSQC NMR) spectroscopy technique can be used for simultaneous identification and quantification of chemical groups in the structure of lignin. Standard 2D-HSQC NMR spectroscopy does not provide accurate quantification due to the limiations of traditional HSQC methods. However, these quantification limitations can be avoided by extrapolation to time-zero ^{13}C HSQC (HSQC$_0$) for series of spectrums, acquired for different repetition times. The prediction results of total monomers yields and the product distribution with experimental results within a few percentage points as determined by gas chromatography (GC) after hydrogenolysis.

Chapter 2 Protecting groups for stabilization of lignin during pretreatment

This chapter was adapted from the following articles with the permission of all co-authors and the journal.

***Postprint version of the article**: Shuai, Li, Masoud Talebi Amiri, Ydna M. Questell-Santiago, Florent Héroguel, Yanding Li, Hoon Kim, Richard Meilan, Clint Chapple, John Ralph, and Jeremy S. Luterbacher. "Formaldehyde stabilization facilitates lignin monomer production during biomass depolymerization." Science 354, no. 6310 (2016): 329-333. (DOI: 10.1126/science.aaf7810)*

***My contribution**: Providing the results for optimization of the pretreatment for spruce, F5H-Poplar, and parts of Beech wood biomass. Purification of the monomers. Calculations related to H_2 consumption during hydrogenolysis.*

***Postprint version of the article**: Lan, Wu, Masoud Talebi Amiri, Christopher M. Hunston, and Jeremy S. Luterbacher. "Protection Group Effects During α, γ-Diol Lignin Stabilization Promote High-Selectivity Monomer Production." Angewandte Chemie 130, no. 5 (2018): 1370-1374. (DOI: 10.1002/anie.201710838)*

***My contribution**: Screening some of the protecting groups. Optimization of hydrogenolysis towards narrower product distribution in regards with reaction temperature, reaction pressure, reaction time, and catalyst for F5H-Poplar and Birch.*

2.1 Utilization of formaldehyde for lignin stabilization

Extracting a soluble and uncondensed lignin substrate during biomass pretreatment could facilitate the production of lignin monomers and be compatible with current biorefinery strategies. The major challenge in developing such a process is to prevent interunit C–C coupling during extraction, which involves the condensation of a benzylic cation and an electron-rich aromatic ring (Figure 17).[171]

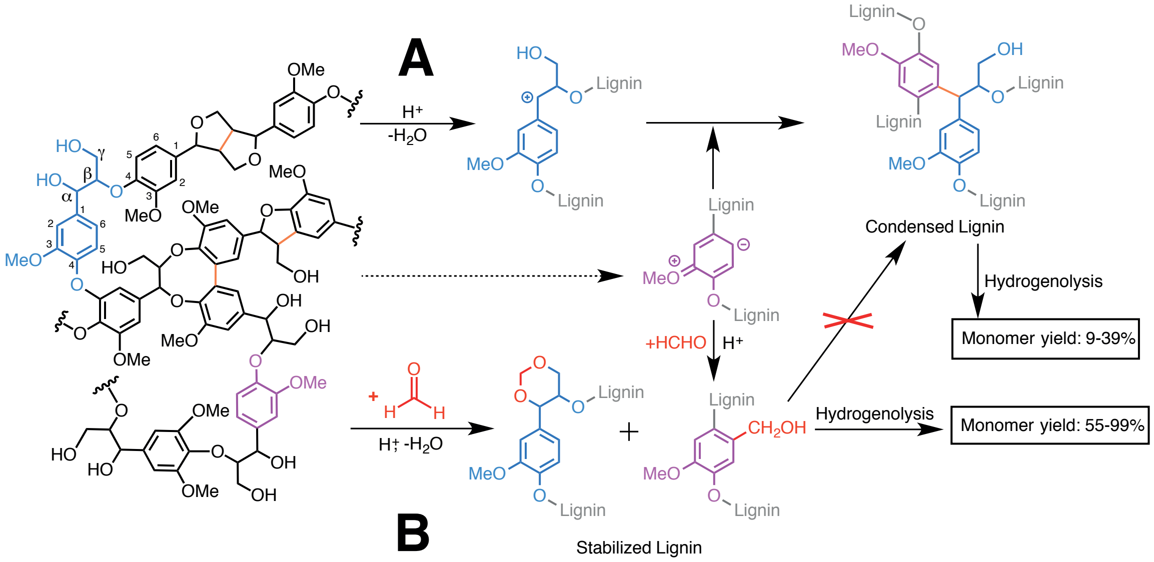


Figure 17 Lignin monomer production by extraction followed by hydrogenolysis.
(A) Lignin extraction, condensation and hydrogenolysis in a standard acidic process, and (B) lignin extraction, stabilization with formaldehyde and hydrogenolysis. Monomer yields (based on Klason lignin on a molar basis) are based on extracted lignin (for runs where 80±2 wt% of Klason lignin was extracted). The bonds highlighted in orange represent inter-unit C–C linkages.

Our strategy was to attempt to block these reactive positions with a protection agent during pretreatment. We report here that using formaldehyde (FA) to stabilize lignin during extraction leads to near-theoretical yields of lignin monomers after hydrogenolysis of the extracted product. These yields were 3-7 times higher than those obtained when using the same method without FA (Figure 17). Evidence suggests that FA hinders the formation C–C linkages through two mechanisms. First, in an acidic, water-deficient environment, FA reacts to form a stable six-membered 1,3-dioxane (acetal) structure with the 1,3-diols (and their α- and γ-hydroxyl groups) on lignin side-chains (Figure 17B), thereby blocking the formation of benzylic cations. Second, the electron-rich positions at the para or meta positions to methoxyl groups on the aromatic ring reacted with the positively charged, protonated FA to form hydroxymethyl groups, blocking these reactive positions.

We used a model lignin dimer (Figure 18A, veratrylglycerol-β-guaiacyl ether, VG, 1), representing the predominant β–O–4-linked units with their characteristic available α- and γ-hydroxyl groups for initial tests.[92] Without FA, **1** depolymerized to form **3** and **4** through **2**, as was consistent with previous report [171] (Figure 18A and 18C). The low yield of **3** (40%) compared to **4** (95%) after 7 hours was likely due to condensation reactions similar to those observed with real lignin. In the presence of FA, **1** was rapidly converted to **5** and **6** (Figure 18B and 18D). Comparison of 2D heteronuclear single-quantum coherence nuclear magnetic resonance (HSQC NMR) spectra of pure **1** and the derived products obtained in the presence of FA (Figure 18E and 18F) revealed the formation of the 1,3-dioxane structure (FA-derived carbon is highlighted in red).

Following the acid treatment of **1**, as in a typical biomass fractionation process, a hydrogenolysis step was performed to produce monomers.[92] Using the products from the reaction without FA, **3** was quantitatively converted to veratrylpropyl compounds **7** (Figure 18A and Figure A1B, A1C and A1D in appendix A), with the low yields confirming the loss of over half of these **3**-derived products due to condensation (43% yield). In contrast, hydrogenolysis of the product mixture obtained in the presence of FA produced 85% for veratrylpropyl compounds **8** and yields of 87% for guaiacyl compounds **9** (Figure 18B and Figure A1 in appendix A). This substantial increase in yield with FA addition (85% vs. 43%) mirrors the stabilizing effect that FA has on real lignin (Figure 17).

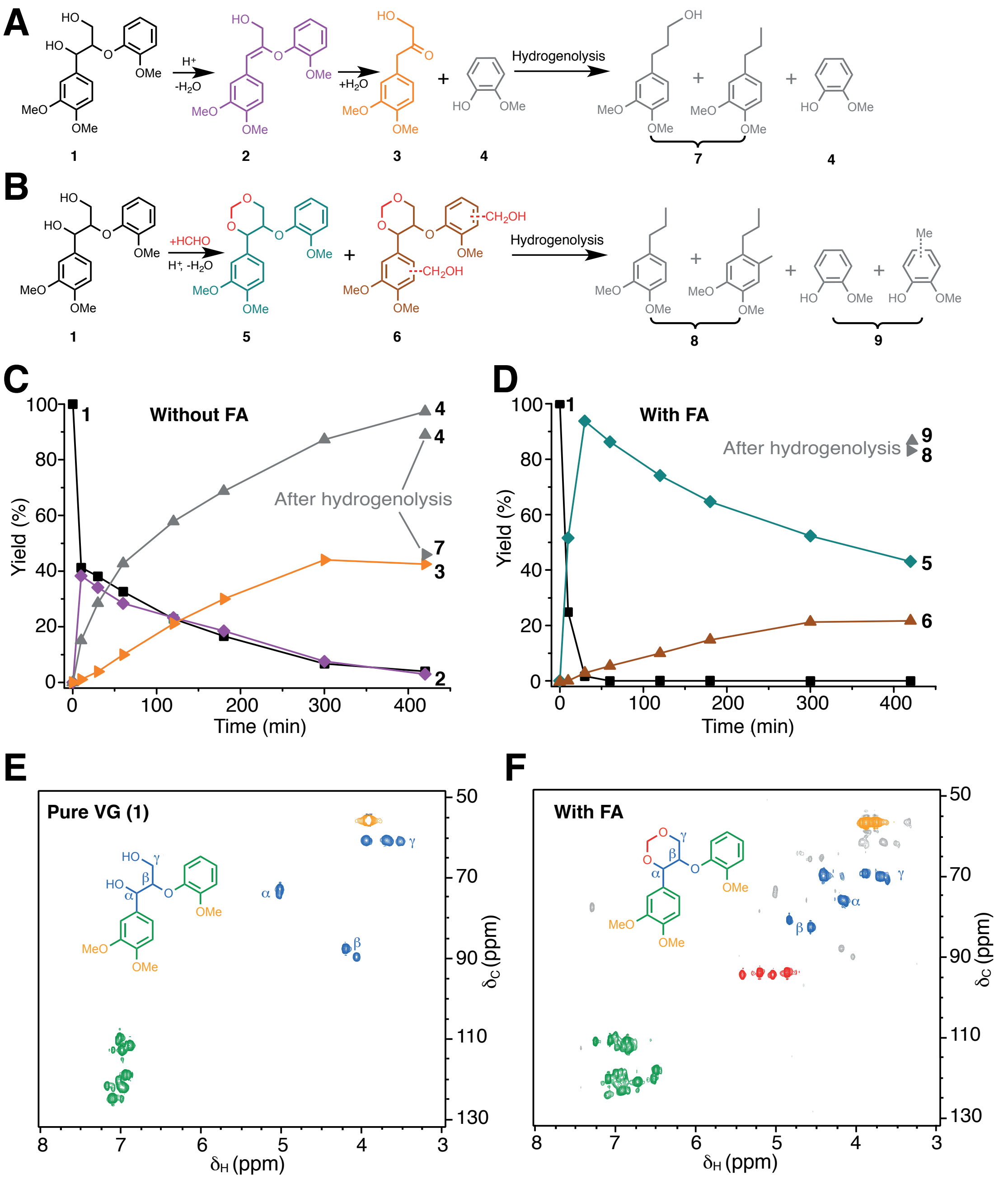


Figure 18 Acid-catalyzed depolymerization of veratrylglycerol-β-guaiacyl ether (VG). (A) Without formaldehyde (FA); (B) with FA; (C) time course of reaction A; (D) time course of reaction B; (E) 2D HSQC NMR spectrum of pure VG 1; (F) 2D HSQC NMR spectrum of the product mixture resulting from reaction B after 30 min.

We observed the hydroxymethylation of aromatic rings with FA via the additional ring methylation in several VG reaction products (**8** and **9**), and when reacting several lignin monomers with FA followed by hydrogenolysis (Figure A2 in appendix A). Methylation appears to act as a secondary lignin stabilization mechanism as demonstrated by condensation studies using vanillyl alcohol, which can polymerize like lignin but cannot form a dioxane structure with FA.[92]

We used beech wood (Fagus grandifolia) to test the effect of FA during the extraction of real lignin using similar reaction conditions. 2D HSQC NMR of lignin extracted in the presence of FA (Figure 19A) revealed the presence of the 1,3-dioxane structures observed with model compounds (Figure 18). Without FA addition, no dioxane structure signals were observed, but signals characteristic of the lignin side-chains had disappeared (Figure 19A). This disappearance was consistent with the results of previous studies of lignin condensation.[142] The color of the lignin obtained in the presence of FA was significantly lighter than that obtained in the absence of FA, serving as a qualitative indicator of reduced condensation (Figure A3 in appendix A). Lignin monomer identification and quantification are described extensively in the appendix A.[92] Five major lignin monomers were identified in the products without FA addition, or from the direct hydrogenolysis of native lignin (Figure 19 and Figure A4 in appendix A), that were all consistent with previous reports.[172,173] With FA addition, five additional methylated monomers were identified (Figure 19, A4 and A5 in appendix A). The addition of FA did not significantly affect lignin removal, but final monomer yields after hydrogenolysis of the extracted lignin were remarkably different (Figure 19B). Without FA addition, hydrogenolysis of the extracted lignin resulted in a monomer yield of 7%, which was consistent with previous studies.[133,174] With FA addition, the highest yield was 47% after hydrogenolysis of the extraction liquor (Figure 19B). We attribute the increased monomer yield obtained with FA primarily to the same stabilization mechanism observed with the model lignin dimer (Figure 17 and 18). This yield (47%)

was comparable to the yield of 48% obtained using direct hydrogenolysis [133,173], and to the yields obtained using established lignin monomer analysis methods, including derivatization followed by reductive cleavage (DFRC)[111] and nitrobenzene oxidation (NBO) (Figure A7 in appendix A).[175]

We applied this method to softwood (spruce, Picea abies) and obtained a yield of 21%, which was comparable to the yield (21%) obtained by direct hydrogenolysis. These relatively low monomer yields (~20%) are consistent with previous reports using softwoods[173,176] (14, 27) and are likely due to the increased quantities of interunit C–C linkages in softwoods compared to hardwoods.

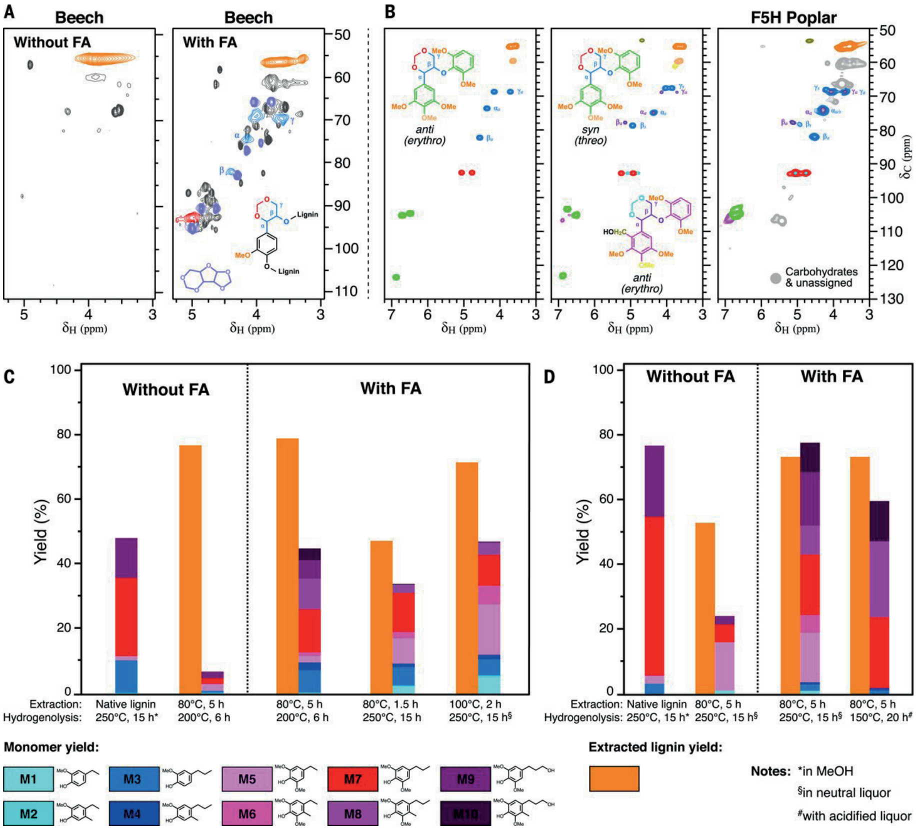


Figure 19 Production of lignin monomers from lignocellulosic biomass.
(A) 2D HSQC NMR spectra of lignin extracted from beech wood in the absence and presence of FA. Signals attributed to diformyl-xylose (Figure A6) are highlighted in light purple. (B) The same extraction by using high-syringyl F5H-poplar (after a precipitative purification of the lignin, right), and assignments from authentic syringyl model compounds for the produced dioxanes. (C) Lignin monomer yields from beech wood (Table A1, entries 1 to 5). (D) Lignin monomer yields from high-syringyl F5H-poplar (Table A1, entries 14, 15, 16, and 22).

Our strategy for improving lignin monomer yields is to hinder the formation of C–C linkages during lignin extraction. Thwarting the production of such linkages by altering the native lignin could further increase yields. Overexpression of the ferulate 5-hydroxylase gene in poplar (Populus spp., F5H-poplar) yields lignin that has a high content of syringyl units (98.3%) and reduced native inter-unit C–C linkages (87% β-O-4 linkages, 3.5% β-1 linkages and 10% β-β linkages, Figure S8 in appendix A).[111] Using FA stabilization, we obtained lignin monomer yields of up to 79.25% after hydrogenolysis of the extracted substrate. This yield was comparable to that obtained by direct hydrogenolysis of F5H-poplar (77.70%) (Figure 19C) In contrast, a monomer yield of only 24.53% was obtained from native lignin without FA addition. Furthermore, monomer selectivities from the isolated soluble F5H-poplar lignin were quantitative, which was confirmed by gel permeation chromatography of the lignin before and after hydrogenolysis. After hydrogenolysis, the broad lignin oligomer peak (Mw ≈ 2900 g/mol) had disappeared and only a single monomer peak remained (Figure A9 in appendix A). Selectively hydrogenating only the extracted upgradeable lignin, as opposed to recalcitrant lignin and other biomass derivatives, could facilitate continuous processing and increase the selective use of H2 by avoiding the hydrogenation of recalcitrant lignin or other biomass fractions. These advantages illustrate the potential benefits of controlling lignin biosynthesis for improved lignin valorization. FA-extracted lignin can also be converted to monomers at 150°C if acid is present to break the 1,3-dioxane ring structure (Figure 19C). Direct hydrogenolysis of F5H-poplar at 150°C in neutral conditions leads to negligible monomer production. In acidic conditions, direct hydrogenolysis leads to lower yields (48% vs. 60%) than with FA extraction, and can also lead to the partial loss or reaction of the polysaccharide fraction of biomass (Table A1 in appendix A).[92]

In contrast to direct biomass hydrogenolysis, extracting a soluble stabilized lignin fraction before its hydrogenolysis is compatible with most polysaccharide depolymerization approaches. With

FA present, 40% of xylan was dehydrated into furfural and 50% reacted with FA to form diformyl-xylose (structure given in Figure 19A, and Figure A6 in appendix A), which could be converted back to xylose under aqueous acidic conditions (Figure 20, Section A2.3 in appendix A). High yields of glucose (80%) were obtained after enzymatic hydrolysis of the leftover solids (Section A2.3 in appendix A).[92] Therefore, the treatment of beach wood under acidic conditions in the presence of FA leads to overall yields of 47% lignin monomers (theoretical yield), 80% glucose, 40% furfural and 50% xylose (90% overall yield from xylan) (Figure 20A). With F5H-poplar, yields of 79% lignin monomers, 76% glucose, 14% furfural and 75% xylose (89% overall yield from xylan) were obtained.

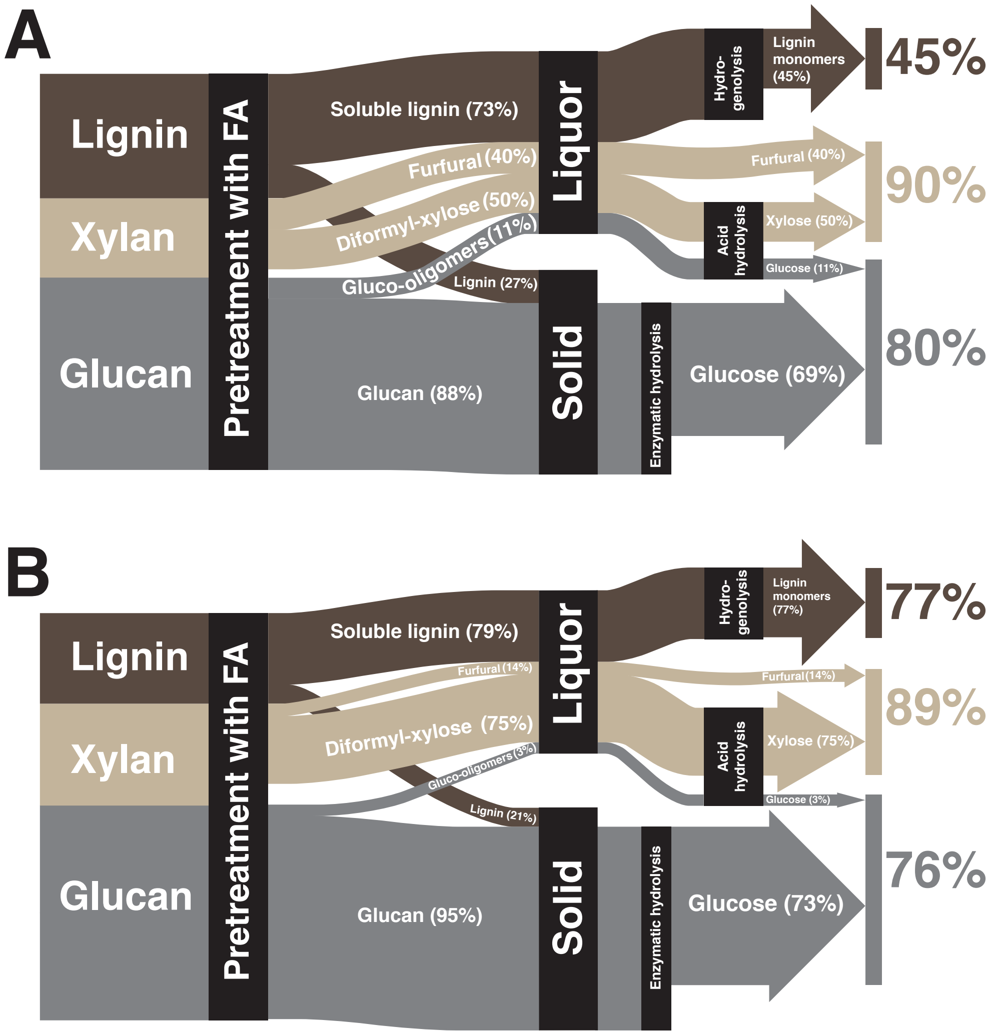

A
Lignin
Xylan
Glucan
Pretreatment with FA
Soluble lignin (73%)
Furfural (40%)
Diformyl-xylose (50%)
Gluco-oligomers (11%)
Lignin (27%)
Glucan (88%)
Liquor
Solid
Hydro-genolysis
Lignin monomers (45%)
45%
Furfural (40%)
Acid hydrolysis
Xylose (50%)
Glucose (11%)
90%
Enzymatic hydrolysis
Glucose (69%)
80%
B
Lignin
Xylan
Glucan
Pretreatment with FA
Soluble lignin (79%)
Furfural (14%)
Diformyl-xylose (75%)
Gluco-oligomers (3%)
Lignin (21%)
Glucan (95%)
Liquor
Solid
Hydro-genolysis
Lignin monomers (77%)
77%
Furfural (14%)
Acid hydrolysis
Xylose (75%)
Glucose (3%)
89%
Enzymatic hydrolysis
Glucose (73%)
76%

Figure 20 Mass balance of polysaccharide (glucan and xylan) and lignin fractions
(A) Beech wood and (B) F5H-poplar wood during integrated biomass depolymerization with formaldehyde addition.

These results suggest that lignin upgrading could be easily integrated into current biorefinery schemes, especially considering that FA is a relatively inexpensive bulk chemical that can easily be produced from biomass-derived syngas or methanol, either sourced biologically or from lignin methoxyl groups.[173] Furthermore, only 2-13% of FA was consumed for stabilizing lignin, while the remaining FA could be recovered (Section A2.4 in appendix A). Studies with 13C-labeled FA demonstrated that the FA that was consumed in lignin ended up either as methyl groups on the lignin aromatic rings, or as methane after hydrogenolysis (Section A2.4 in appendix A). Though processes exist for converting methane back to FA, we estimated that the cost of FA would be at most 20% of the cost of monomers (Section A2.4).

One issue in this technique is the alkylation of aromatic rings by formaldehyde and unselective hydrogenolysis which complicates the product mixture leading to the production of 6-10 major monomers. Threfore, screening various protective groups can primarily help us to shape a better understanding of the chemistry involved in this reaction, and subsequently to optimize the process towards lower energy consumption and providing a narrow product distribution.

2.2 Protecting group optimization towards high-selectivity monomer production

We investigated the effect of adding different diol protection reagents during lignin extraction process on the final lignin monomer yield after hydrogenolysis, as shown in Figure 21. As a control, we performed the same extraction without a protection reagent. This control was performed in triplicate and the average yield was 14 wt% with a 95% confidence interval between 13-15 wt%. Therefore, any experiment resulting in a yield value outside of this range was very likely due to the effect of protection reagents. The first group of reagents that we tried was aldehydes that form cyclic acetal with the α,γ-diol of the lignin side chain. All these aldehydes except for hydroxymethylfurfural led to a statistically significant improvement in monomer yield. In the case of acetaldehyde and propionaldehyde, the monomer yields were 37 and 42 wt%, respectively, which were comparable to that obtained with formaldehyde (46 wt%). Interestingly, no alkylated aromatic monomers were detected, resulting in higher selectivity to ethylsyringol and ethylguaiacol. Ketal groups, especially acetonide, is one of the most common diol protection group in carbohydrate chemistry. Therefore, acetone, butanone and some other related reagents that protect 1,2 and 1,3-diol to form cyclic ketals were also tested.

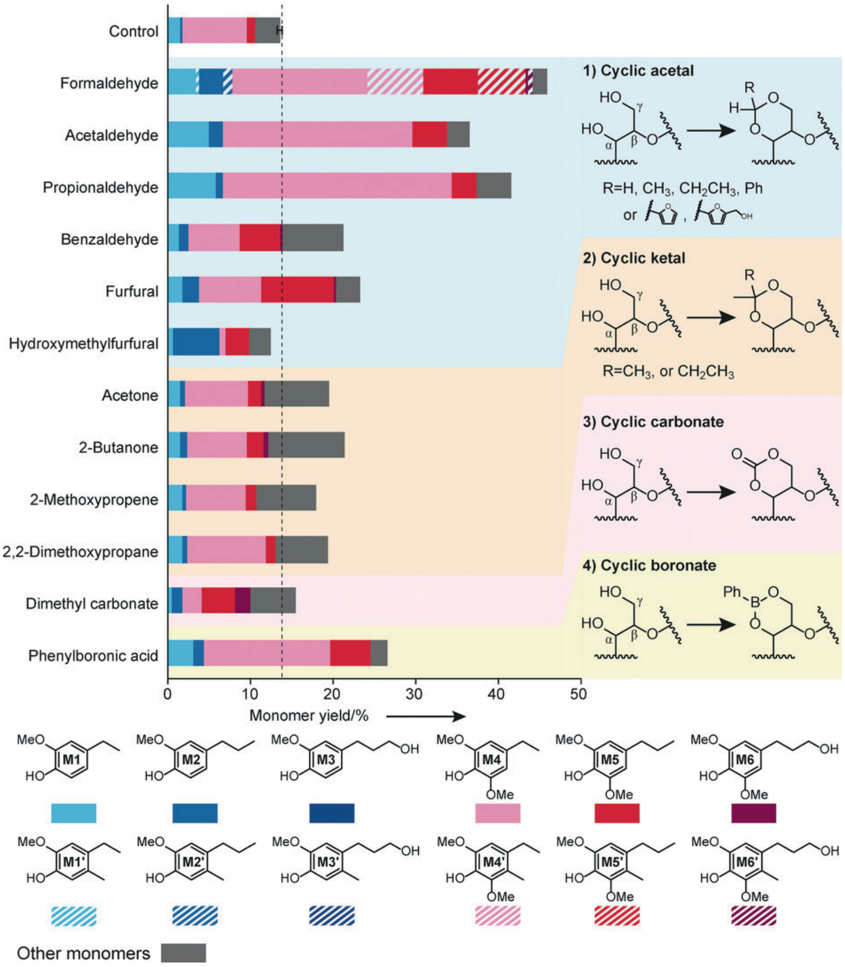


Figure 21 Hydrogenolysis of birch lignin extracted with the addition of different protection reagents. Extraction condition in dioxane: 80 °C, 5 h; hydrogenolysis condition: Ru/C, 250 °C, 40 bar of H2, 15 h. Monomer yields are given based on Klason lignin content.

In all of these experiments some protective effect was detected during lignin mild acid extraction condition, resulting in moderate but statistically significant monomer yield improvements after the subsequent hydrogenolysis. We also tested carbonate and boronate protection group that were known to react with alcohol groups under alkaline condition. The use of dimethyl carbonate and phenylboronic acid as protection reagents resulted in monomer yields of 16 and 27 wt%, respectively (in both cases more than that the control experiment). This study demonstrated that all diol protection groups were effective in stabilizing lignin to some degree but aldehydes were the most effective. Both acetaldehyde and propionaldehyde showed especially remarkable activity in protecting the α,γ-diol position and preventing condensation during lignin extraction and pretreatment conditions.

To better understand the reaction mechanism between acet- or propionaldehyde and lignin, we performed reactions using the common lignin dimeric model compound veratrylglycerol-β-guaiacyl ether (VG) and acet- or propion-aldehyde. The yields of ethylidene acetal and propylidene acetal protected VG were 90% and 86%, respectively, after a 30 min reaction, slightly lower than that of methylene acetal protected VG (92%). The formation of the acetal protected structures were confirmed by NMR (Figure 22, A and B).

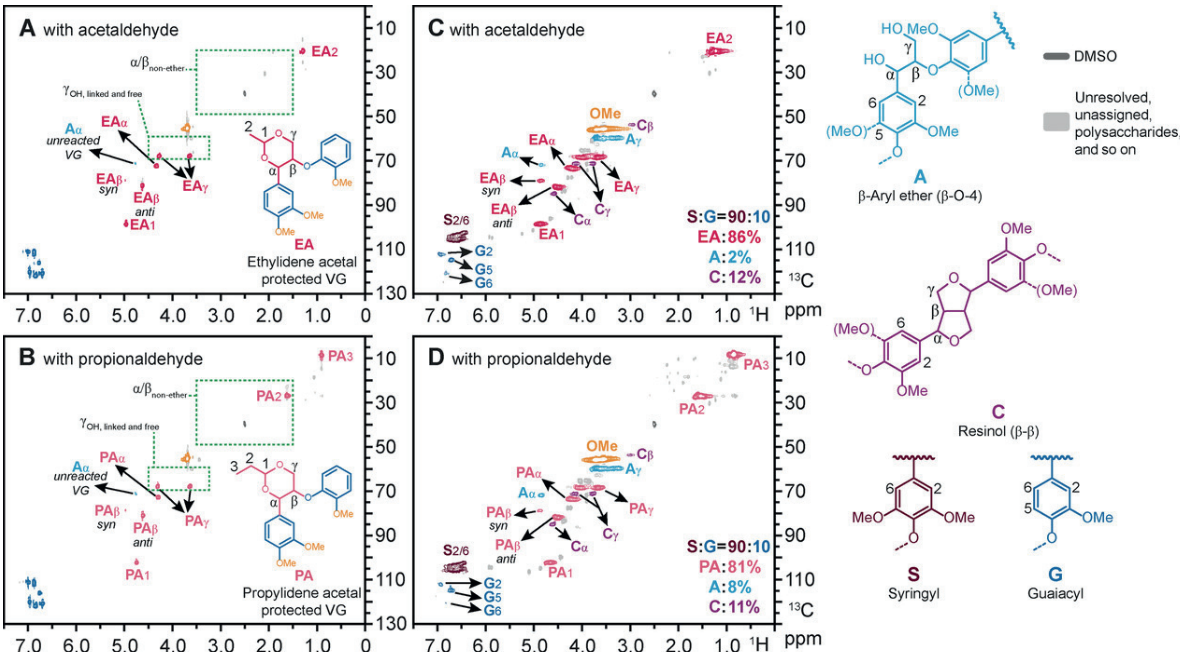


Figure 22 HSQC spectra of the product mixture

Resulting from the reaction between veratrylglycerol-β-guaiacyl ether and A) acetaldehyde, B) propionaldehyde, or isolated birch lignin extracted in the presence of C) acetaldehyde or D) propionaldehyde. The contour areas were integrated to calculate the S:G ratio and interunit linkage distributions. S:G equals (S2/6)/2:(G2); AAα, Aα, plus Cα equals 100 %. Green dashed lines indicate the regions of hydroxypropyl and hydroxyethyl correlations. Key: dimethyl sulfoxide (DMSO), veratrylglycerol-β-guaiacyl ether (VG).

According to the HSQC spectra, the major compound of the mixture was the acetal protected VG with trace amounts of unreacted VG. A detailed inspection of the HSQC and COSY spectra revealed no peaks originating from –CH_2OH/–or CH_2CH_2OH in their corresponding region (green dash line), indicating that no hydroxyethylation or hydroxypropylation occurred during the reaction. After a subsequent hydrogenolysis reaction, no products with additional ethyl/propyl side chains on the aromatic ring were detected, confirming the results obtained with native lignin in birch (Figure 21). Lignin extracted from birch in the presence of both acet- and propionaldehyde were also prepared for NMR characterization (Figure 22 C and D) and confirmed the structure observed with the model compound. Contours corresponding to each structure were integrated to calculate the S/G ratio and the percentage of different inter-linkages. Results showed that more than 90% of the β-O-4 structure was protected by an ethylidene or propylidene acetal group, which was consistent with the model compound study.

The absence of aromatic ring alkylation when acet- or propionaldehyde were used as protection reagents instead of formaldehyde reduced the amount of products by half opening the door to high selectivity lignin upgrading (as shown in Figure 23).

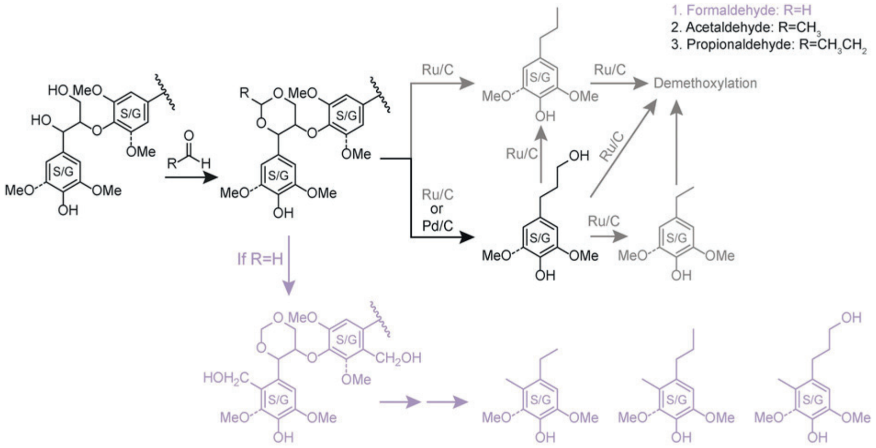


Figure 23 Reaction mechanism for the hydrogenolysis of aldehyde-stabilized lignin.
When using formaldehyde during lignin extraction, alkylation on the aromatic rings occurred and the corresponding products were generated. When using Ru/C as the catalyst, propanolphenolic monomers were largely converted to ethylphenolic monomers and demethoxylation of syringyl units occurred; when using Pd/C as the catalyst, these reactions were mostly suppressed and the major products were propanolphenolic monomers.

As shown in the previous section of this chapter, hybrid poplar in which the ferulate 5-hydroxylase (F5H) gene was overexpressed contained a much higher proportion of syringyl units in the lignin (more than 95%)[150], resulting in higher content of cleavable β-O-4 ethers. We therefore used F5H poplar as a feedstock to get higher monomer yields and a higher product selectivity to syringyl monomers. The results of hydrogenolysis of lignin extracted with the addition of propionaldehyde are shown in Figure 24. Using the same extraction conditions as for formaldehyde, a 54% monomer yield was obtained with propionaldehyde compared to 75% for formaldehyde. With formaldehyde, long extraction time benefited lignin extraction. However, considering that propylidene acetal is less stable than methylene acetal, a longer extraction time could have favored the partial cleavage of the protection groups and led to some lignin repolymerization. Therefore, we reduced the lignin extraction time, which increased the monomer yield to 65 wt% for a 2.5 h treatment. At these conditions about 20% of guaiacyl monomers were obtained, which is higher than what would be expected based on the native lignin (~5%). This discrepancy was likely due to the demethoxylation of syringyl units during hydrogenolysis over Ru/C. We also assumed that ethylphenols were mainly converted from propanol-phenols under the applied hydrogenolysis condition over Ru/C by cleaving the $–CH_2OH$ group (Figure 23). Model compound studies using POHS and propylsyringol (PS) as reactants for hydrogenolysis demonstrated the cleavage of $–CH_2OH$ and demothoxylation under the same reaction conditions. We lowered the temperature and/or time of hydrogenolysis in order to suppress these demethoxylation and CH_2OH cleavage reactions and achieve higher product selectivity (Figure 24). When decreasing the reaction time from 15 h to 5 h, a 73 % monomer yield was obtained. Additionally, the POHS and POHG products were favored at lower temperature: 48% selectivity to POHS was reached at 200 °C for 15 h over Ru/C and the fraction of the syringyl products increased to 89%.

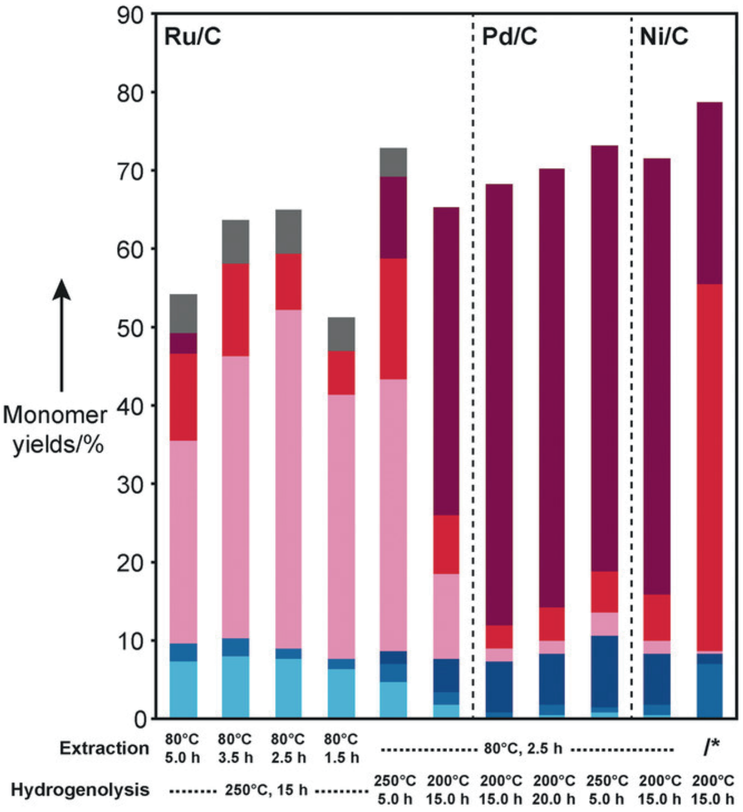


Figure 24 Monomer yields (on the basis of Klason lignin) and product distribution
Resulting from the hydrogenolysis of propylidene-acetal-protected lignin extracted from F5H poplar. Key: see Figure 21 for color-coding of molecules; *direct hydrogenolysis.

To further increase the selectivity to POHS/POHG, Pd/C was used as the hydrogenolysis catalyst due to its low alcohol hydrogenation activity.[177,178] By running hydrogenolysis at 200 °C for 15 h, a yield of 68% monomers was obtained with a 82% selectivity to a single POHS product (92 % to POHS+POHG). A higher monomer yield (73%) was achieved but slightly lower selectivity (74% toPOHS) at 250 °C for 5 h (Figure 24). HSQC spectra of the hydrogenolysis product mixtures further confirmed the monomers structure and also demonstrated full conversion of β-O-4 linkages of the lignin polymer. This result is in agreement with previous work by Sels and co-workers that had also reported a high selectivity to POHG and POHS when treating birch sawdust mixed with Pd/C in

MeOH to directly upgrade native lignin.[137] Here, we are able to achieve equivalent with extracted lignin but also significantly higher selectivities to a single product (POHS) when using F5H-poplar. Ni has often been proposed as a non-noble metal hydrogenolysis catalyst for lignin producing near theoretical yields of mainly propylphenolic monomers during direct hydrogenolysis of untreated wood. When high-syringyl poplar was used during direct hydrogenyolsis of untreated wood with Ni/C in MeOH a relatively low selectivity was obtained (79% monomer yield with a 59% selectivity to PS and only 29% to POHS). Interestingly, when propionaldehyde-protected extracted lignin was used, a selectivity of 78% to POHS was achieved with an only slightly diminished yield of total monomers (72%) (Figure 24). The acetal protection group is likely causing a suppression of alcohol hydrogenation over Ni but a further study of the mechanism is still underway.

We also performed lignin extraction with both acet- and propionaldehyde and hydrogenolysis with Pd/C with birch and spruce (Figure 25). In all cases more than 91% selectivity to POHS and POHG monomers were achieved. Total monomer yields were those that are typically expected for both wild type species based on their fraction of inter-unit ether linkages: 48% and 20% for birch and spruce, respectively (with acetaldehyde extraction yields were within 10% of the yield from direct hydrogenolysis of untreated wood[133]). The relatively low yield from spruce is consistent with the low β-aryl ether content in softwood lignin as reported elsewhere.[179] However, due to softwoods having only guaiacyl linkages, the selectivity to the single POHG product was 91%.

Besides lignin depolymerization, recovery of carbohydrates was also evaluated for this process. Unlike direct hydrogenolysis of native wood, which causes irreversible conversion of some of the carbohydrates, this method generated clean cellulose or acetal protected carbohydrates (mostly xylose) that could be easily deprotected by aqueous acidic treatment. Up to 78% of xylose was obtained under 120 °C with 6 wt% H_2SO_4 solutions within 10 min (Figure 26). The leftover solids after pretreatment could undergo a typical enzymatic hydrolysis process without any further acid treatment, generating an 80% yield of glucose. The high enzymatic hydrolysis without any further acid treatment was a significant improvement over our previous work with formaldehyde, where the resulting cellulose required a harsh acid treatment to remove formyl groups that caused significantly reduced yields (~10%) during enzymatic hydrolysis. We suspect that the binding of propionaldehyde to cellulose was much more easily reversible than that of formaldehyde and was likely removed even under the mild hydrolysis conditions (50°C, pH 5).

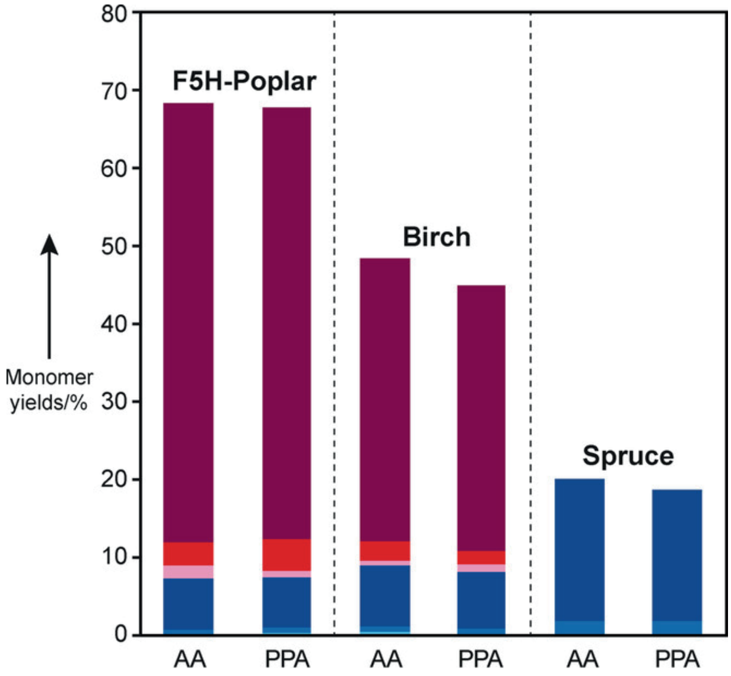


Figure 25 Monomer yields (on the basis of Klason lignin) and product
Resulting from the hydrogenolysis of lignin extracted from F5H poplar, birch, and spruce with addition of acetaldehyde (AA) and propionaldehyde (PPA). The extraction conditions were 80 °C, 2.5 h for F5H poplar and birch, and 3.5 h for spruce. The hydrogenolysis was performed at 200 °C, with 40 bar of H2 for 15 h with Pd/C . Key: see Figure 21 for color-coding of molecules; ferulate 5-hydroxylase (F5H).

In summary, we developed a high yield lignin extraction and depolymerization process that is compatible with most biorefinery schemes and can, depending on the biomass source, produce a single product at high selectivity. Because lignin is extracted under typical pretreatment conditions, both xylose and glucose can be produced at high yields as well leading to mostly single products from all three major biopolymers. By adding acet- or propionaldehyde to a lignin extraction and biomass pretreatment process, the α,γ-diol in lignin was protected by an ethylidene/propylidene acetal group without aromatic ring alkylation. The use of these protection groups led to near theoretical monomer yields during the following hydrogenolysis and high selectivity to one or two propanolphenolic monomers. When using F5H-poplar as a feedstock, 82% and 78% selectivity of POHS was obtained with Pd/C and Ni/C, respectively. To our best knowledge, these are the highest selectivities obtained to a single product from hardwood lignin depolymerization and the first time a non-precious metal was used for high-selectivity lignin hydrogneolysis with molecular hydrogen.

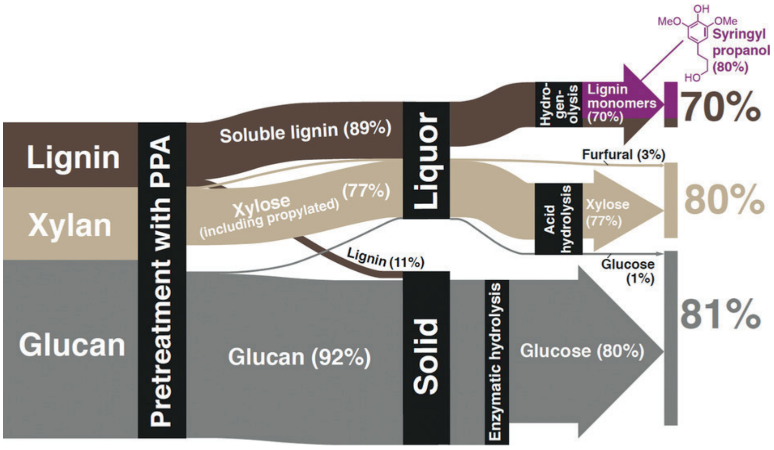

OH
MeO
OMe
HO
Syringyl propanol (80%)
Lignin
Xylan
Glucan
Pretreatment with PPA
Soluble lignin (89%)
Xylose (including propylated) (77%)
Lignin (11%)
Glucan (92%)
Liquor
Solid
Hydro-gen-olysis
Lignin monomers (70%)
70%
Furfural (3%)
Acid hydrolysis
Xylose (77%)
80%
Glucose (1%)
Enzymatic hydrolysis
Glucose (80%)
81%

Figure 26 Mass balance of polysaccharides (glucan and xylan) and lignin fractions

Chapter 3 A protocol for fractionation of lignocellulosic biomass

This chapter was adapted from the following article with the permission of all co-authors and the journal.

***Postprint version of the article**: Masoud Talebi Amiri, Graham R. Dick, Ydna M. Questell-Santiago, and Jeremy S. Luterbacher. "Fractionation of Lignocellulosic Biomass to Produce Uncondensed Aldehyde-Stabilised Lignin." Nature Protocols (Manuscript accepted)*

***My contribution**: Preliminary study on finding a compatible method for fractionation of lignocellulosic biomass. Developing the method for formaldehyde-stabilized lignin procedure along with full characterization and yields calculation for each fraction.*

In the previous chapter, we outlined methods to protect and upgrade lignin using aldehydes. However, the lignin was systematically upgraded while still largely mixed with the pretreatment liquor or only partially purified. Here we detail two optimized protocols for isolating and purifying both the formaldehyde and propionaldehyde stabilized lignins in solid form without degrading them by adapting the procedures introduced in the previous chapters. Key to the development of these strategies was the determination of the conditions to neutralize the reaction solution and the appropriate solvent blends to facilitate the precipitation and purification of the lignin. These isolated lignins are ideal substrates for further processing or upgrading studies as they have retained their full upgrading potential and can be processed without further effects from the cellulose and hemicellulose fractions. The optimized protocols also describe methods for recovering highly digestible cellulose (in the case of the propionaldehyde fractionation) and stabilized xylose thereby delineating a methodology to truly fractionate lignocellulosic biomass. Of the biopolymers (as determined by biomass composition analysis[180]), ≥95 wt% were recovered as solid cellulose (cellulose-rich solids: ~77-81% glucan, ~2-10%

xylan, 87-95% of native glucan, and 6-24% of native xylan), stabilized xylose (solid or liquid depending on the purity and aldehyde type, 66-86% of native xylan), and solid stabilized lignin (typically representing 103-133% of the Klason lignin in the original biomass with the extra mass arising from the contribution of acid soluble lignin and extraneous aldehyde functionalization) (Figures 27).

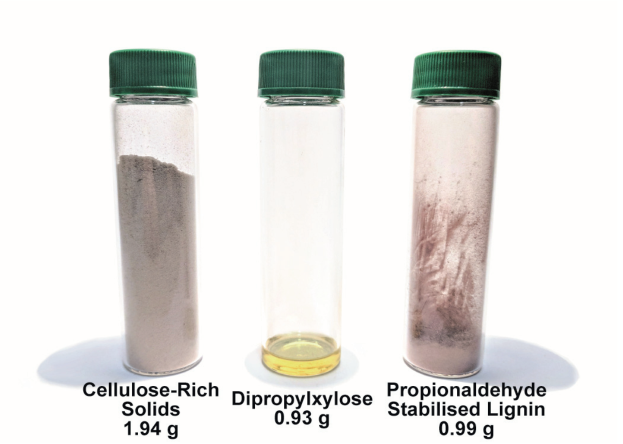

Cellulose-Rich Solids 1.94 g
Dipropylxylose 0.93 g
Propionaldehyde Stabilised Lignin 0.99 g

Figure 27 The products of the propionaldehyde-based fractionation protocol.
The products of the propionaldehyde-based fractionation protocol. Birch wood (extracted and dried, 4.5027 g) was fractionated using the propionaldehyde fractionation protocol into highly digestible cellulose-rich solids (left), dipropylxylose (centre), and propionaldehyde stabilized lignin (right).

3.1 Assessing lignin valorization strategies

To evaluate both the extraction and quality of the resulting lignin, we wanted to define a metric to benchmark the total yields of monomers obtained from the extracted and purified lignin for a specific biomass source. In the previous chapter, we have referred to yields of monomers from lignin as a wt% of the Klason Lignin content of the original biomass as determined during the sulphuric acid mediated biomass composition analysis. The monomers' yields are calculated by reconstituting their molecular weight to their pre-hydrodeoxygenated state (e.g. for M1, the nominal molecular weight is 152.19 g·mol-1, but it is produced from a subunit of lignin with a molecular weight of 196.20 g·mol-

1, see Equipment Setup: Gas Chromatography in Appendix B), summed, and then divided by the Klason Lignin content. By doing so, we, and many others, have reported yields between 40 and 55 wt%, which is generally assumed to be the theoretical maximum for wild hardwoods.[181] However, as illustrated below (Figure 28), we have also noted that this measure can vary substantially across biomass sources. Additionally, the Klason lignin measurement doesn't include the acid soluble lignin fraction, which can be substantial, but is also based on an imprecise UV-absorption measurement.[182] This leads to fluctuating yields between species and even more significantly across genera (Figure 28).

In the past few years, it has been established that the direct hydrogenolysis of the native biomass provides close to the maximum yield of obtainable monomers by cleaving the native lignin interunit ether bonds and leaving the interunit C–C bonds intact. Here, we refer to this procedure as "direct hydrogenolysis", but it's also known as "reductive fractionation" or "lignin first".[31] We argue that this procedure provides the most accurate measure of the theoretical monomer yield from lignin for a given biomass source because it depolymerizes the native lignin's ether linkages before any condensation can occur.[179] We propose that the quality of a given lignin extraction procedure is best determined by comparing the total amount of monomers produced after hydrogenolysis of the extracted lignin to the total amount of monomers produced from the same quantity of native biomass by direct hydrogenolysis. This comparison leads to a rapidly obtained and easily understood metric for determining isolated lignin quality.

3.2 Advantages, limitations, and alternative methods

When we subjected the isolated stabilized lignin to hydrogenolysis and, as discussed above, compared the isolated monomer yields to those resulting from the direct hydrogenolysis of native lignin, we observed yields from isolated lignin that were within ≥88% of those obtained by direct hydrogenolysis without any extraction or fractionation (Figure 28). This comparison demonstrates that we were able to isolate and purify a lignin fraction that contained almost all the original native lignin and could be upgraded at near theoretical yields. In parallel, the isolated cellulose could be enzymatically depolymerized to produce ≥99 mol% glucose when using propionaldehyde-based stabilization (as compared to the compositional analysis). As previously detailed, cellulose isolated in the presence of formaldehyde has poor digestibility (enzymatic hydrolysis yields below 20%) but could be improved if an acid treatment is performed to remove acetal or formyl species. Furthermore, ≥66% of the xylose could be recovered as the aldehyde stabilized derivative. All of these balances are summarized in Figure 28 and in Tables 4-9 for both fractionation protocols.

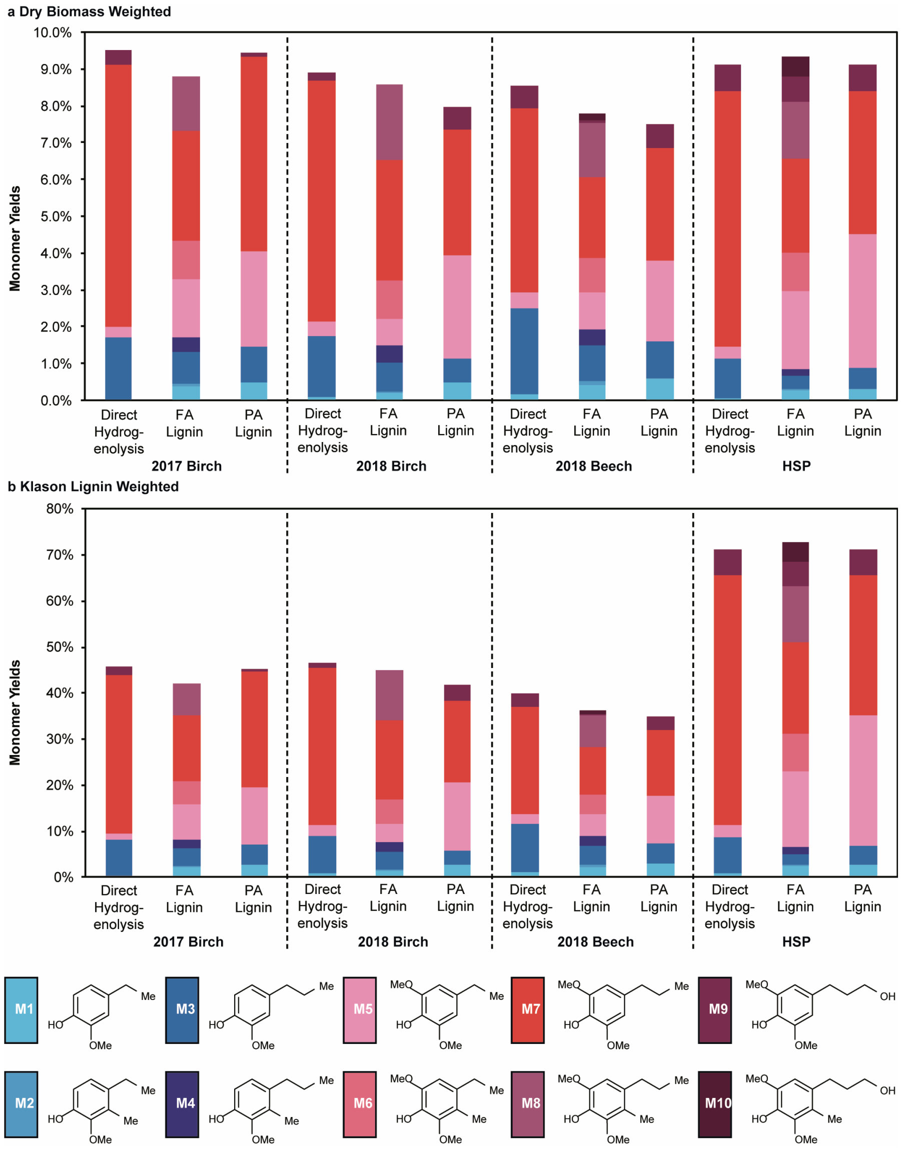


Figure 28 Hydrogenolysis Data for the Formaldehyde and Propionaldehyde Stabilized Lignins Compared with the Direct Hydrogenolyses of the Feedstock Biomass. These two charts compare the monomer yields from the hydrogenolysis of the raw biomass (Direct Hydrogenolysis), formaldehyde stabilized lignin (FA), and propionaldehyde stabilized lignin (PA) for three biomass sources: 2017 Birch, 2018 Birch, 2018 Beech, and High Syringyl Poplar (HSP). The direct hydrogenolysis represents the highest possible yield of monomers for these biomass sources and was performed on biomass that had not been extracted or dried. The formaldehyde and propionaldehyde stabilized lignins were fractionated from extracted and dried biomass, which typically lowers their yields by ~1-2% when Klason weighted and ~0.25-0.4% for whole biomass weighted (See Figure 1 in Appendix B). Each data point represents one experiment

The main emphasis of these protocols is to isolate high-quality, uncondensed bench stable lignin using aldehyde stabilization. The only alternative for producing standard uncondensed lignin is to use the cellulolytic lignin isolation method, which requires extensive ball-milling, enzymatic treatment, and chemical processing of the wood.[183] The sum of all the necessary operations can exceed 5 days and be very difficult to scale. In contrast, the protocols presented here are easily scalable to enable the production of gram quantities of bench stable lignin. They also only require inexpensive chemicals, common laboratory equipment, a rudimentary understanding of synthetic organic chemistry, and 6-7 hours for the isolation of the stabilized lignin (more for the full fractionation procedure).

The main limitations of these protocols stem from their reliance on organic solvents, which are often toxic and/or flammable; formaldehyde and propionaldehyde, which are toxic; and acids, which are corrosive. Consequently, these procedures require a sufficiently ventilated workspace and appropriate protective equipment. However, these precautions and requirements are no different from those of many common chemical reactions or industrial processes. Additionally, these procedures have been performed exclusively on hardwood and softwood biomass sources. Given the similarity of lignin structure across several biomass phyla, we believe that these protocols should function well on other lignocellulosic biomass sources, but have no experimental evidence to support it.

3.3 Overview of the procedure

In appendix B, we have detailed the formaldehyde and propionaldehyde protocols that lead to the full fractionation of lignocellulosic biomass into its three constituent biopolymers as distinct, readily-upgradable fractions. Figure 29 and Figure 30 present the general scheme of the formaldehyde and propionaldehyde protocols, respectively. To aid in their comprehension, we have provided a brief summary of these procedures here.

First, the extraction is performed in which the biomass is heated with the stabilizing aldehyde and hydrochloric acid in 1,4-dioxane at elevated temperatures for 3-3.5 hours. Once complete, the cellulose-rich solids are collected by filtration and washed with 1,4-dioxane and methanol to remove any residual lignin or stabilized-sugars. The filtrate is then set aside, and the cellulose-rich solids are treated with either dilute sulphuric acid in the case of formaldehyde or saturated sodium bicarbonate in the case of propionaldehyde to cleave any residual acetals on the substrate. This material is then washed with de-ionized water and acetone and then collected.

The filtrate that was set aside previously is then neutralized through the addition of either a saturated sodium bicarbonate solution in the case of formaldehyde or solid sodium bicarbonate in the case of propionaldehyde. At this point, these procedures diverge substantially. To recover the formaldehyde-stabilized lignin, the dioxane of the neutralized filtrate is first removed by evaporation. This results in the precipitation of the lignin, which is then collected by filtration, washed with de-ionized water, and dried. To recover the formaldehyde stabilized xylose, the filtrate is extracted with hexanes, and the hexanes fraction is concentrated in vacuo to afford the sugar as a yellow oil. To recover the propionaldehyde-stabilized lignin, the neutralized filtrate is filtered to remove the bicarbonate, and the filtrate is concentrated in vacuo to form a brown oil. This oil is then diluted with ethyl acetate and added dropwise to a stirred solution of hexanes resulting in the precipitation of a solid. This solid is collected by filtration and triturated with diethyl ether to afford the propionaldehyde-stabilized lignin. To recover the propionaldehyde-stabilized xylose, the hexanes filtrate and diethyl ether supernatant are combined, concentrated in vacuo, and purified by chromatography to afford the sugar as a yellow oil.

To determine the quality of the cellulose-rich solids, enzymatic hydrolyses were performed on the materials in a pH 5 citrate buffer and compared against their sulphuric acid mediated compositional analyses. The quality and purity of the extracted lignins were determined through both ^{1}H-NMR and hydrogenolysis. The purity of the stabilized xyloses was determined by ^{1}H-NMR.

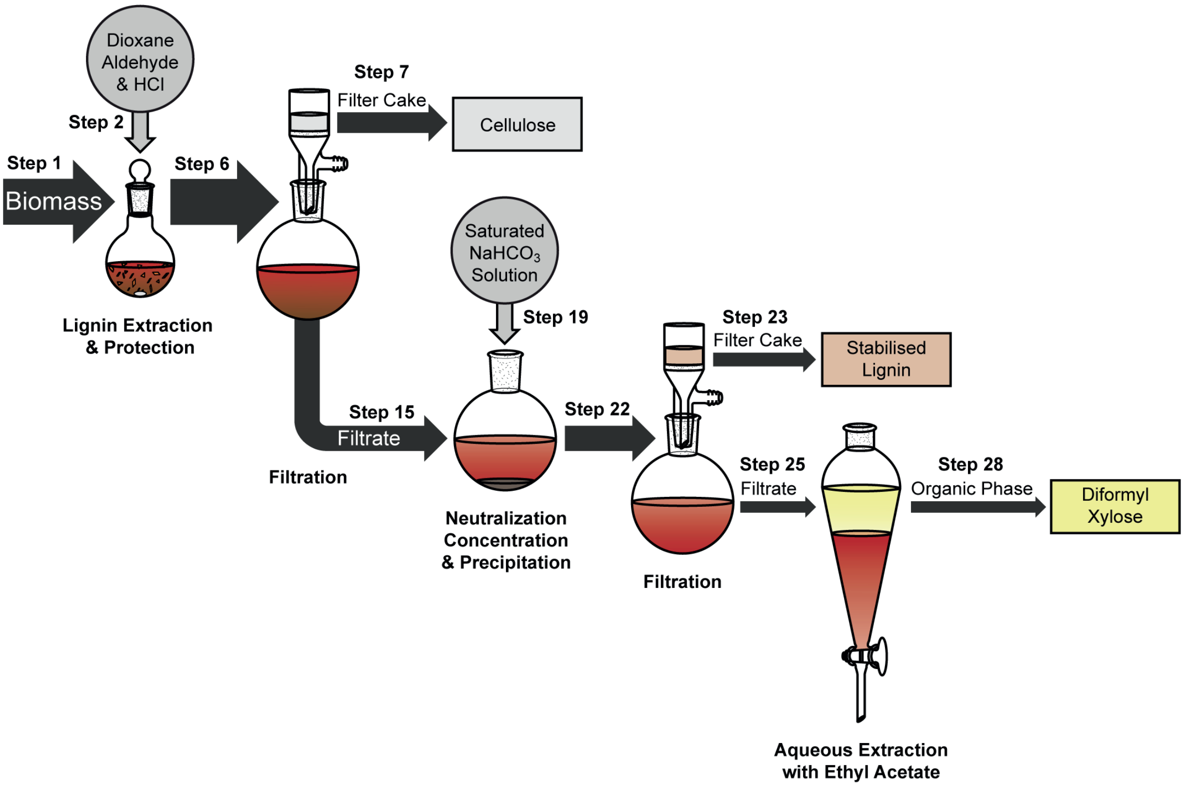

Dioxane
Aldehyde
& HCl
Step 2
Step 1
Biomass
Step 6
Lignin Extraction
& Protection
Step 7
Filter Cake
Cellulose
Filtration
Step 15
Filtrate
Saturated
NaHCO$_3$
Solution
Step 19
Neutralization
Concentration
& Precipitation
Step 22
Step 23
Filter Cake
Stabilised
Lignin
Filtration
Step 25
Filtrate
Step 28
Organic Phase
Diformyl
Xylose
Aqueous Extraction
with Ethyl Acetate

Figure 29 The Formaldehyde Biomass Fractionation Protocol
An overview of the formaldehyde biomass fractionation protocol, which yields cellulose-rich solids, formaldehyde-stabilized lignin, and diformylxylose. The arrow widths are in proportion to the mass of the fraction being isolated for 2018 Birch. Some of the later purification steps are eliminated for clarity.

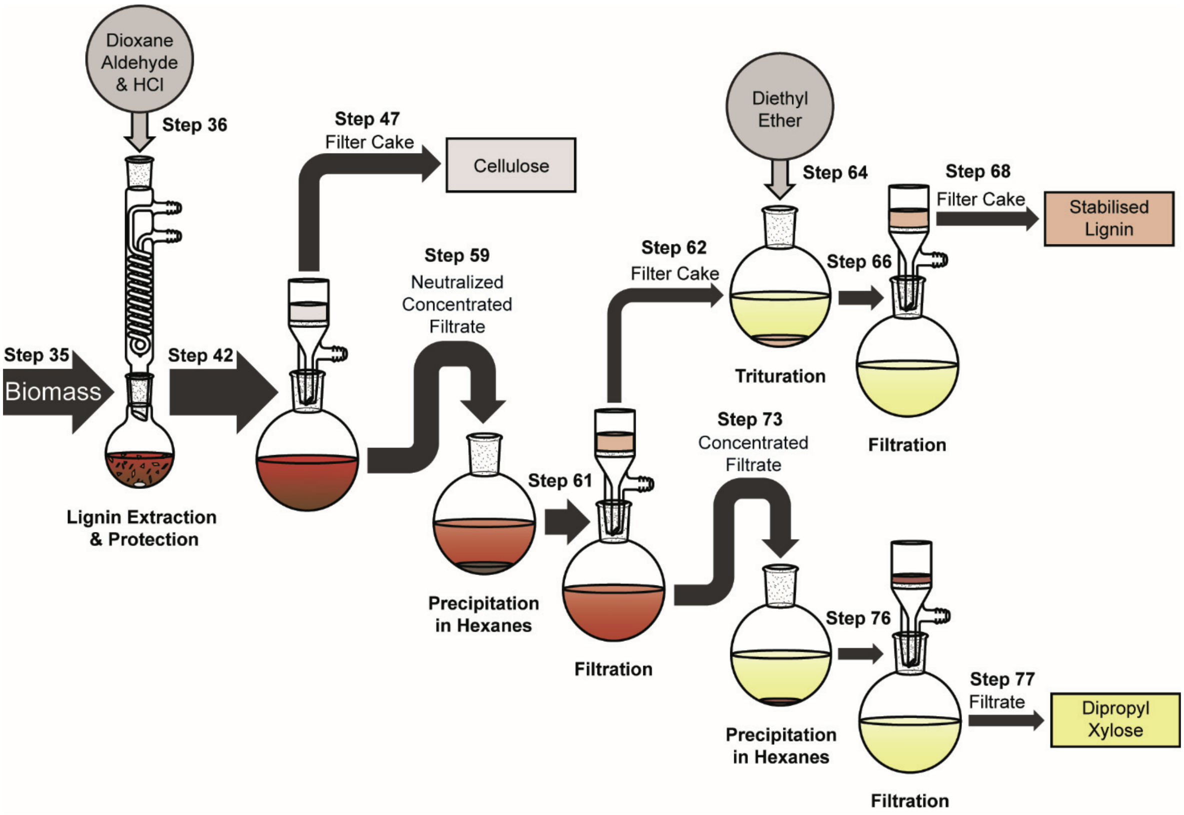

Dioxane
Aldehyde
& HCl
Step 36
Step 35
Biomass
Step 42
Lignin Extraction
& Protection
Step 47
Filter Cake
Cellulose
Step 59
Neutralized
Concentrated
Filtrate
Precipitation
in Hexanes
Step 61
Filtration
Step 62
Filter Cake
Diethyl
Ether
Step 64
Trituration
Step 66
Filtration
Step 68
Filter Cake
Stabilised
Lignin
Step 73
Concentrated
Filtrate
Precipitation
in Hexanes
Step 76
Filtration
Step 77
Filtrate
Dipropyl
Xylose

Figure 30 . The Propionaldehyde Biomass Fractionation Protocol
An overview of the propionaldehyde biomass fractionation protocol, which yields highly digestible cellulose-rich solids, propionaldehyde-stabilized lignin, and dipropylxylose. The arrow widths are in proportion to the mass of the fraction being isolated for 2018 Birch. Some of the later purification steps for the dipropylxylose are eliminated for clarity.

3.4 Anticipated results

3.4.1 Formaldehyde biomass fracrtionation protocol

After completion of the entire protocol, we anticipate the collection of three separate biomass fractions: cellulose-rich solids, formaldehyde stabilized lignin, and diformylxylose. For birch wood, we expect that the cellulose-rich solids will appear as a fluffy, beige, fibrous powder representing 43.0

wt% (2.1373 g) of the raw, unextracted biomass. Enzymatic hydrolysis of this cellulose will yield 33.8 wt% as glucose and 3.6 wt% as xylose representing 46.3% of the glucan and 9.6% of the xylan in the raw biomass (note: the wt% of glucose and wt% of xylose were calculated as the dehydrated glucan and xylan respectively). The formaldehyde stabilized lignin will be isolated as a grey powder representing 24.0 wt% (1.2150 g) of the raw, unextracted biomass after correcting for the formaldehyde stabilization. Hydrogenolysis of this powder will yield 34% of its mass as monomers (after correction for hydrodeoxygenation) for an overall yield of monomers of 8.57 wt% versus dry biomass (8.01 wt% versus the raw, unextracted biomass). The diformylxylose will be isolated as a yellow oil representing 13.2 wt% (0.8631 g, corrected for the formaldehyde stabilization and converted to xylan) of the raw, unextracted biomass and 76% of the xylan.

For beech wood, we expect that the cellulose-rich solids will appear as a fluffy, beige, fibrous powder representing 35.3 wt% (1.7548 g) of the raw, unextracted biomass. Enzymatic hydrolysis of this cellulose will yield 40.5 wt% as glucose and 4.1 wt% as xylose representing 44.1% of the glucan and 9.6% of the xylan in the raw biomass (note: the wt% of glucose and wt% of xylose were calculated as the dehydrated glucan and xylan respectively). The formaldehyde stabilized lignin will be isolated as a light brown powder representing 21.4 wt% (1.0846 g) of the raw, unextracted biomass after correcting for the formaldehyde stabilization. Hydrogenolysis of this powder will yield 33% of its mass as monomers (after correction for hydrodeoxygenation) for an overall yield of monomers of 7.78 wt% versus dry biomass (7.26 wt% versus the raw, unextracted biomass). The diformylxylose will be isolated as a yellow oil representing 13.0 wt% (0.8517 g, corrected for the formaldehyde stabilization and converted to xylan) of the raw, unextracted biomass and 77% of the xylan. For both biomass samples, ^{1}H-NMR and ^{13}C-NMR spectra of the diformylxyloses and HSQC spectra of the stabilized lignins are provided in the Supplementary Information. For a more detailed presentation of this data, please see **Tables 4-9**.

3.4.2 Propionaldehyde biomass fracrtionation protocol

After completion of the entire protocol, we anticipate the collection of three separate biomass fractions: cellulose-rich solids, propionaldehyde stabilized lignin, and dipropylxylose. For birch wood, we expect that the cellulose-rich solids will appear as a fluffy, grey, fibrous powder representing 39.1 wt% (1.9425 g) of the raw, unextracted biomass. Enzymatic hydrolysis of this cellulose will yield 77.4 wt% as glucose and 10.1 wt% as xylose representing 96% of the glucan and 24% of the xylan in the raw biomass (note: the wt% of glucose and wt% of xylose were calculated as the dehydrated glucan and xylan respectively). The propionaldehyde stabilized lignin will be isolated as a purplish-brown powder representing 18.5 wt% (0.9853 g) of the raw, unextracted biomass after correcting for the propionaldehyde stabilization. Hydrogenolysis of this powder will yield 38% of its mass as monomers (after correction for hydrodeoxygenation) for an overall yield of monomers of 7.97 wt% versus dry biomass (7.47 wt% versus the raw, unextracted biomass). The dipropylxylose will be isolated as a yellow oil representing 10.7 wt% (0.9257 g, corrected for the propionaldehyde stabilization and converted to xylan) of the raw, unextracted biomass and 66% of the xylan.

For beech wood, we expect that the cellulose-rich solids will appear as a fluffy, grey, fibrous powder representing 37.9 wt% (1.8998 g) of the raw, unextracted biomass. Enzymatic hydrolysis of this cellulose will yield 82.1 wt% as glucose and 9.3 wt% as xylose representing 96% of the glucan and 23% of the xylan in the raw biomass (note: the wt% of glucose and wt% xylose were calculated as the dehydrated glucan and xylan respectively). The propionaldehyde stabilized lignin will be isolated as a purplish- powder representing 20.6 wt% (1.0976 g) of the raw, unextracted biomass after correcting for the propionaldehyde stabilization. Hydrogenolysis of this powder will yield 32% of its mass as monomers (after correction for hydrodeoxygenation) for an overall yield of monomers of 7.49 wt% versus dry biomass (6.99 wt% versus the raw, unextracted biomass). The dipropylxylose will be isolated as a yellow oil representing 10.3 wt% (0.9011 g, corrected for the propionaldehyde stabilization and converted to xylan) of the raw, unextracted biomass and 68% of the xylan. For both biomass samples, ^{1}H-NMR and ^{13}C-NMR spectra of the dipropylxyloses and HSQC spectra of the stabilized lignins are provided in the Supplementary Information. For a more detailed presentation of this data, please see **Tables 4-9**.

Table 4. Composition of Biomass Used for the Fractionation Protocols.

Biomass Type	**Ash**[a]	**Hydration**[a]	**Extractives**[a]	**Klason Lignin**[a]	**Acid Soluble Lignin**[a]	**Glucan**[a]	**Xylan**[a]	**Galactan**[a]	**Arabinan**[a]	**Mannan**[a]
Birch[b]	-	5.70%	2.60%	19.60%	-	-	-	-	-	-
Birch[c]	0.15%	6.09%	3.30%	17.94%	2.81%	31.39%	16.14%	1.46%	0.56%	0.26%
Beech[c]	0.17%	6.48%	3.06%	20.05%	2.05%	32.40%	15.13%	1.37%	0.50%	0.54%
HSP[d]	-	6.4%	4.4%	12.8%	6.4%	34.6%	14.3%	-	-	-

[a] Fractions are presented a wt% of the raw biomass. [b] Year 2017. [c] Year 2018. [d] Data taken from Lan, W., *et. al.* Angew. Chem. Int. Ed. 57, 1356–1360 (2018). Abbreviations: High Syringyl Poplar (HSP)

Table 5. Yield of Monomers from the Direct Hydrogenolysis of the Biomass on a Dry Basis.

Biomass Type	**M1**[a]	**M3**[a]	**M5**[a]	**M7**[a]	**M9**[a]	**Total**[a]
Birch[b]	0.00%	1.69%	0.29%	7.15%	0.39%	9.49%
Birch[c]	0.09%	1.64%	0.42%	6.53%	0.23%	8.91%
Beech[c]	0.16%	2.32%	0.43%	4.99%	0.63%	8.54%
HSP	0.05%	1.06%	0.34%	6.96	0.70%	9.11%

[a] Yield is presented as a wt% of the total raw biomass on a dry basis and is corrected for any mass lost with respect to the monomers' initial structures in the native lignin polymer. Even-numbered monomers are not seen in the direct hydrogenolysis as there is no hydroxymethylation as that is only a consequence of the formaldehyde pre-treatment. [b] Year 2017. [c] Year 2018. Abbreviations: High Syringyl Poplar (HSP)

Table 6. Yield of Lignin Monomers from the Hydrogenolysis of the Isolated Formaldehyde-Stabilized Lignin Powder.

Biomass Type	**M1**[a]	**M2**[a]	**M3**[a]	**M4**[a]	**M5**[a]	**M6**[a]	**M7**[a]	**M8**[a]	**M9**[a]	**M10**[a]	**Total**
Birch[b]	0.37%	0.08%	0.84%	0.41%	1.57%	1.04%	3.00%	1.45%	0.00%	0.00%	8.77%
Birch[c]	0.18%	0.05%	0.80%	0.44%	0.73%	1.03%	3.29%	2.04%	0.00%	0.00%	8.57%
Beech[c]	0.42%	0.10%	0.96%	0.43%	1.01%	0.93%	2.21%	1.48%	0.07%	0.16%	7.78%
HSP	0.25%	0.04%	0.35%	0.21%	2.09%	1.05%	2.56%	1.56%	0.67%	0.54%	9.32%

[a] Yield is presented as a wt% of the total raw biomass on a dry basis and is corrected for any mass lost with respect to the monomers' initial structures in the native lignin polymer due to the hydrogenolysis. [b] Year 2017. [c] Year 2018. Abbreviations: High Syringyl Poplar (HSP)

Table 7. Yield of Lignin Monomers from the Hydrogenolysis of the Isolated Propionaldehyde-Stabilized Lignin Powder.

Biomass Type	**M1**[a]	**M3**[a]	**M5**[a]	**M7**[a]	**M9**[a]	**Total**[a]
Birch[b]	0.49%	0.98%	2.59%	5.26%	0.11%	9.43%
Birch[c]	0.47%	0.65%	2.83%	3.39%	0.63%	7.97%
Beech[c]	0.59%	1.00%	2.21%	3.04%	0.64%	7.49%
HSP	0.29%	0.58%	3.64%	3.88%	0.71%	9.11%

[a] Yield is presented as a wt% of the total raw biomass on a dry basis and is corrected for any mass lost with respect to the monomers' initial structures in the native lignin polymer due to the hydrogenolysis. [b] Year 2017. [c] Year 2018. Abbreviations: High Syringyl Poplar (HSP)

Table 8. Composition and Enzymatic Hydrolysis of the Extracted Cellulose from the Aldehyde Biomass Fractionation Protocols.

		Compositional Analysis					**Enzymatic Hydrolysis**	
Biomass Type	**Fractionation Protocol**	**Glucan (wt%)**	**Xylan (wt%)**	**Hydration (wt%)**	**Klason Lignin (wt%)**	**Acid Soluble Lignin (wt%)**	**Glucan (wt%)[a]**	**Xylan (wt%)[a]**
Birch[b]	Formaldehyde	77.4 %	2.2%	1.7%	4.4%	0.2%	33.8%	3.6%
Beech[b]	Formaldehyde	79.4%	2.6%	2.5%	3.7%	0.3%	40.5%	4.1%
Birch[b]	Propionaldehyde	78.1%	6.0%	5.2%	2.3%	0.5%	77.4%	10.1%
Beech[b]	Propionaldehyde	80.6%	5.6%	3.1%	2.0%	0.4%	82.1%	9.3%

[a] This yield refers to the weight percentage of the glucan and xylan in the cellulose-rich solid that is liberated by the enzymes and observed on the HPLC as glucose or xylose. It is not corrected for hydration allowing for direct comparison to the compositional analysis. [b] Year 2018

Table 9. Masses and Yields of Isolated Fractions from the Fractionation Procedures[a].

Biomass Type	**Fractionation Protocol**	**Cellulose-Rich Solids**	**Protected C_5-Sugars[b]**	**5-Hydroxymethyl Furfural**	**2-Furfural**	**Stabilized Lignin**
Birch[c]	Formaldehyde	2.1373g (43.0%)	0.8631g (13.2%)	0.0101g (0.3%)	0.0549g (1.5%)	1.2150g (24.0%)
Beech[c]	Formaldehyde	1.7548g (35.3%)	0.8517g (13.0%)	0.0137g (0.4%)	0.0665g (1.8%)	1.0846g (21.4%)
Birch[c]	Propionaldehyde	1.9425g (39.1%)	0.9257g (10.7%)	0.0083g (0.2%)	0.0521g (1.4%)	0.9853g (18.5%)
Beech[c]	Propionaldehyde	1.8998g (37.9%)	0.9011g (10.3%)	0.0122g (0.3%)	0.0616g (1.7%)	1.0976g (20.6%)

[a] The yields in this table are all wt% and are presented as the reconstituted aldehyde-free versions of their original biopolymers. (e.g. xylose to xylan, 5-hydroxymethylfurfural to glucan). [b] Diformylxylose from the formaldehyde fractionation protocol and dipropylxylose from the propionaldehyde fractionation protocol. [c] Year 2018.

Chapter 4 Lignin monomer yield prediction by quantitative 2D-HSQC NMR

In previous chapters, we demonstrated that we could protect lignin against the typical structural modifications as well as extract and isolate it very efficiently from biomass. We have also discussed that the amount of ether linkages in lignin both native and extracted could/should be predictive of the yields that can be obtained after hydrogenolysis. This prediction requires an accurate quantification of the absolute amount of ether linkages in a given sample. As mentioned in the section 1.6.3, such accurate quantification is not possible with standard HSQC methods. Another important issue is the differences in relaxation times for different parts of the chain in lignin's polymeric structure. Therefore, a single data acquisition is not sufficient for providing accurate quantification.

A recently developed method for quantitative NMR analysis of metabolites can be used simultaneous identification and quantification of structures in metabolite mixtures by repetition of a pulse sequence between the first ^{1}H excitation pulse and the acquisition point.[177] Errors due to different relaxation times are eliminated by extrapolation to a zero relaxation time (HSQC$_0$) for a series of spectrums, acquired with different repetition times. The most accurate of these methods is known as gradient-selective HSQC$_0$, which provides more accurate quantitative results by reducing the t_1 noise which can affect the peak intensities at low concentrations.[185] The t_1 noise, which generally has higher intensity compared to the normal thermal noise, is a ridge of noise around the large peaks in parallel to the F_1 axis and is specific to 2D NMR spectra.

In this work, we sought to use gradient-selective $HSQC_0$ on isolated lignin samples to quantitatively determine their structural features and conclusively link them to monomer yields and distributions measured by GC-FID obtained after a subsequent hydrogenolysis (Figure 31). By isolating lignin in the presence and absence of aldehydes, we were able to analyze lignins with a wide range of ether contents, including lignins with near-native features.

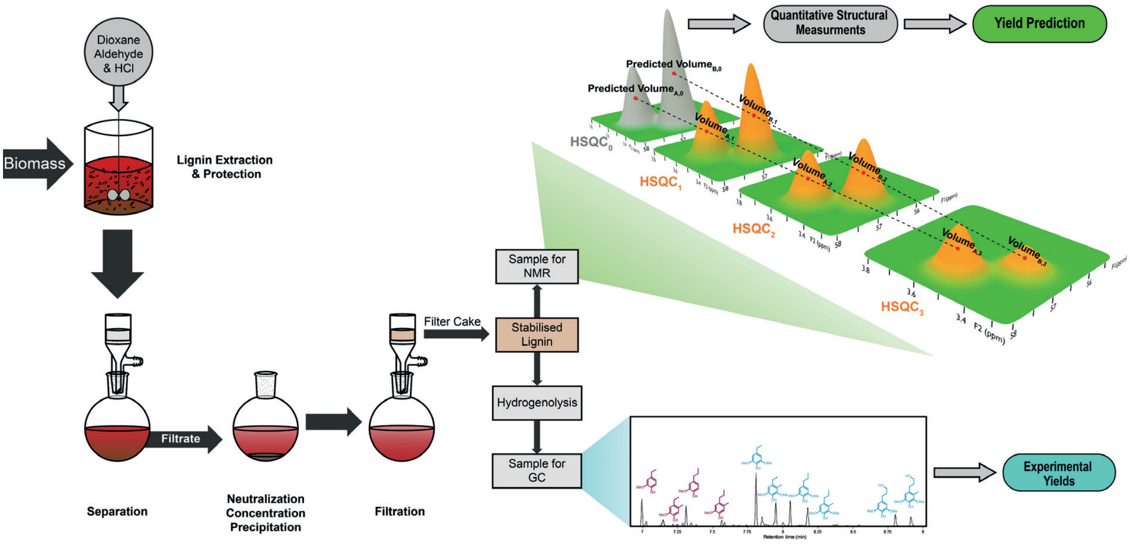

Dioxane
Aldehyde
& HCl
Biomass
Lignin Extraction
& Protection
Separation
Filtrate
Neutralization
Concentration
Precipitation
Filtration
Filter Cake
Sample for NMR
Stabilised Lignin
Hydrogenolysis
Sample for GC
Quantitative Structural Measurments
Yield Prediction
Experimental Yields
Predicted Volume$_{B,0}$
Predicted Volume$_{A,0}$
Volume$_{A,1}$
Volume$_{B,1}$
Volume$_{A,2}$
Volume$_{B,2}$
Volume$_{A,3}$
Volume$_{B,3}$
HSQC$_0$
HSQC$_1$
HSQC$_2$
HSQC$_3$
F2 (ppm)
F1 (ppm)
Retention time (min)

Figure 31 Experimental sequence
Including preparation of lignin samples, prediction of monomer yields by HSQC NMR, and the validation of results with experimental yields

To provide a large range of lignin samples to analyze, we isolated lignin using several extraction techniques described elsewhere, including: formaldehyde-stabilized,[186] propionaldehyde-stabilized,[187] mild dilute acid-catalyzed (MDAC),[188] and Organosolv (unstabilized) lignin. A detailed description of these various isolation methods is provided in the Appendix C but briefly, formaldehyde and propionaldehyde stabilized lignins are prepared by extraction in dioxane water solutions (typically 8.5:1) with HCl and the aldehyde. The aldehyde rapidly reacts with the diol structure on the β—O—4 linkage to form either methylidene acetal structures (in formaldehyde-stabilized lignin, Figure 32a) or propylidene acetal structures (in propionaldehyde-stabilized lignin, Figure 32b). The presence of formaldehyde can also lead to the addition of hydroxymethyl groups to lignin's aromatic ring (Figure 32a). Because the acetal prevents degradation, lignin can be almost completely extracted (>90 wt%) while preserving its ether linkage structure leading to very clean 2-D-HSQC spectra (Figures 32a and 32b). Mild dilute acid-catalyzed lignin extraction uses a similar solvent system (9:1 dioxane:water) but about half the acid concentration, no aldehyde and shorter extraction times (45 min vs. 3 hr). This methods leads to lignin with significant remaining ether linkages (Figure 32c) but typically can only extract small fractions of the native lignin (<20 wt%). As the extraction times are extended, more lignin is recovered but the ether linkage content of the lignin drops rapidly due to condensation. Finally, lignin can be extracted under similar conditions as for the aldehyde-stabilized lignins but without aldehyde addition. These conditions once again lead to near complete extraction of the lignin but absence of a protection group resulted in significant condensation and, thus, much lower ether signals in the 2-D NMR spectrum (Figure 32d).

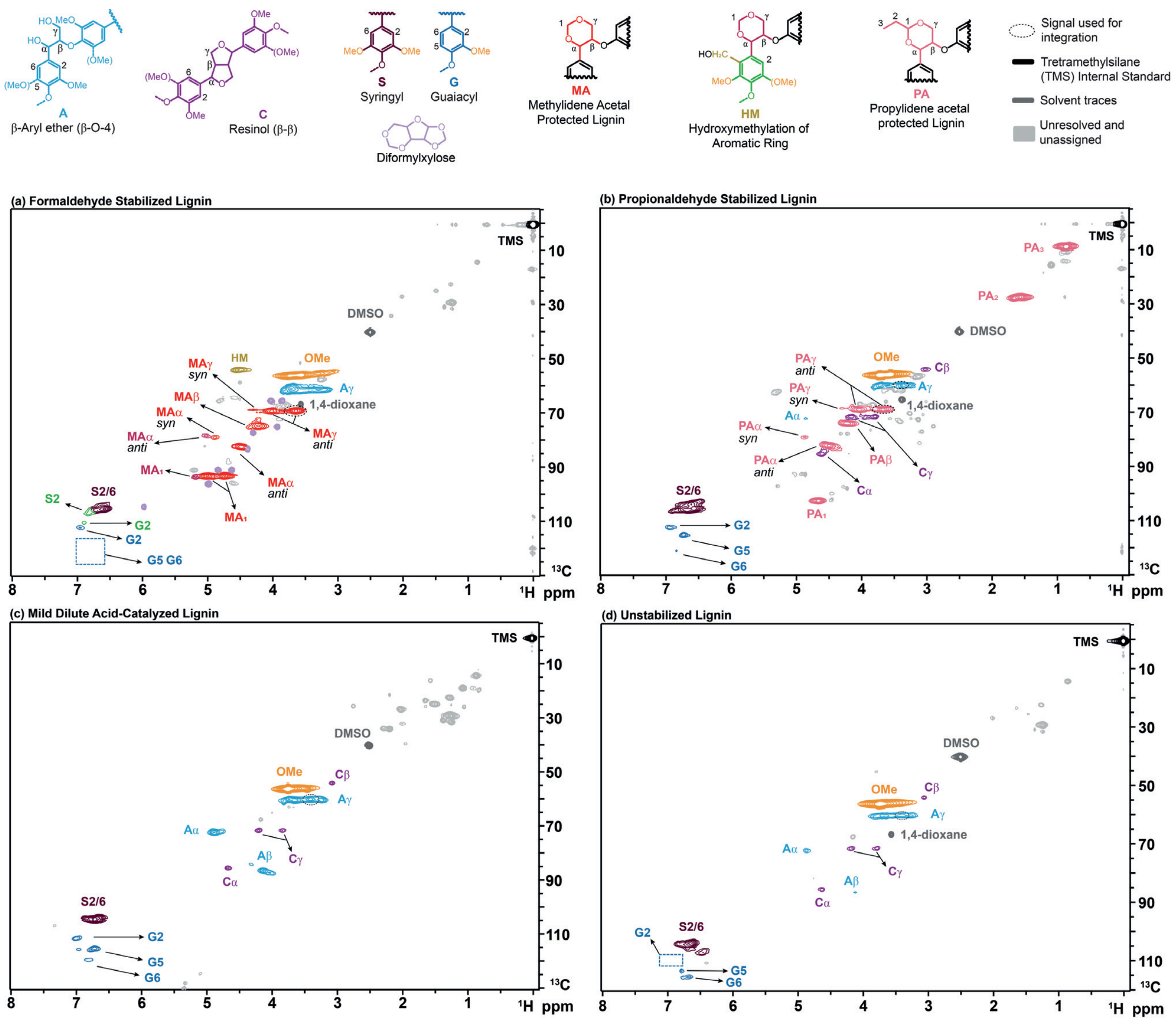


Figure 32 2D-HSQC NMR spectra of isolated lignins.
(a) Formaldehyde-stabilized lignin (b) Propionalde-hyde-stabilized lignin (c) Mild dilute acid-catalyzed lignin (d) Organosolv (unstabilized) lignin.

Gradient-selective $HSQC_0$ was performed by taking three $HSQC_i$ spectra ($i = 1, 2, 3$), which were similar to those shown in Figure 32, with different T2 relaxation times, integrating the volumes of the cross-peak correlations of interest, and calculating a hypothetical cross peak volume for zero relaxation time using a logarithmic extrapolation (See Figure 31 for a depiction of this process and Equation S3 in the Appendix C). The extrapolation is shown for formaldehyde-stabilized lignin in

Figure 33. The extrapolated cross-peak volumes were directly proportional to the number of moles of each chemical group in the sample. The added accuracy brought by extrapolation is well illustrated by the interpolated volumes for the three carbons found in the β—O—4 linkage (C_α, C_β, and C_γ) shown in Figure 32. The amount of moles of C_α, C_β, and C_γ are all the same because they are part of the same chemical functionality and have the same amount of C—H bonds (C_γ has two C—H bonds but only one peak of the doublet is considered during integration due to the t_1 noise, see Section S3.2.5 of the Appendix C). Therefore, their volume should have the same value. However, in $HSQC_1$, $HSQC_2$ and $HSQC_3$, the measured volumes showed small deviations, even in the log scale, which illustrates the accuracy issues of HSQC NMR (Figure 33). In comparison, the extrapolated led to almost no deviation between volumes for these three signals, which is what is expected based on the lignin's chemical structure. The extrapolated cross-peak volume for the formyl group's carbon leads to a higher projected value because the effective number of C—H bonds contributing to the volume is twice that of the aforementioned chemical groups.

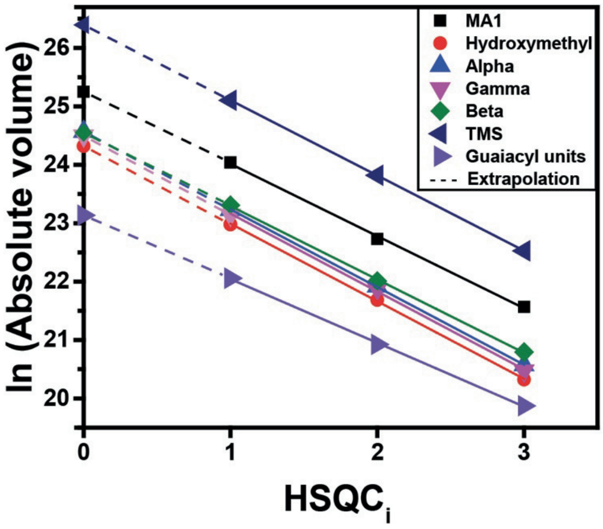


Figure 33 Extrapolation of 2D HSQCi (i=1, 2, 3) integrated peak volumes (Vi), to find V0 values. This plot corresponds to the formaldehyde-stabilized lignin shown in Figure 32a.

Therefore, this method allows us to accurately predict the quantity of protected and unprotected ether bonds as well as the number of resinol (C—C) linkages. We then assumed that each mole of either protected ($n_{\text{Protected}}$) or unprotected ether bond ($n_{\text{Unprotected}}$) would be broken during hydrogenolysis to form one mole of monomers and used this assumption to calculate a predicted a total number of moles of monomers ($n_{\text{Total Monomers}}$) and the resulting monomer yield after hydrogenolysis (Equation 1).

$$n_{\text{Total Monomers}} = n_{\text{Protected}} + n_{\text{Unprotected}} \qquad \text{(Equation 1)}$$

Where the number of moles of protected ether bonds are calculated by taking the average of the amount of the various chemical groups that form this protection group ($n_{\mathrm{protected},\alpha}$, $n_{\mathrm{protected},\beta}$, $n_{\mathrm{protected},\gamma}$, $n_{\mathrm{protection\ group}}$):

$$n_{\mathrm{Protected}} = \mathrm{Average}(n_{\mathrm{protected},\alpha}, n_{\mathrm{protected},\beta}, n_{\mathrm{protected},\gamma}, n_{\mathrm{protection\ group}}) \qquad \text{(Equation 2)}$$

These protected ether bonds correspond to the structures of MA and PA in the case of formaldehyde-stabilized and propionaldehyde-stabilized respectively (Figure 32). The number of moles of unprotected ether bonds correspond ($n_{\mathrm{Unprotected\ Lignin}}$) to the corresponding number of β-aryl ether structures in Figure 32 and are calculated as shown below:

$$n_{\mathrm{Unprotected\ Lignin}} = \mathrm{Average}(n_{\mathrm{unprotected},\alpha}, n_{\mathrm{unprotected},\beta}, n_{\mathrm{unprotected},\gamma}) \qquad \text{(Equation 3)}$$

Furthermore, the quantification of the syringyl and guaiacyl units (along with their hydroxymethylated form in the case of formaldehyde-stabilized lignin) was used to predict the expected monomer distribution. In doing so, we assumed that all these units were equally distributed throughout the oligomers and were equally likely to be connected to ether or C—C linkages. The detailed equations used for the prediction of monomers yield and their distribution are given in Section S3.2.4 of the Appendix C.

The assumption that the total number of monomers after hydrogenolysis can be predicted by the number of β—O—4 bonds assumes that these ether bonds (either protected or unprotected) are almost only found in oligomers that are largely free of C—C linkages. In addition, this assumption also requires that these oligomers must be long enough to ignore end effects (because an oligomer of length n with only β—O—4 linkages should yield $n+1$ monomers). As detailed by Hodge et al. [189], if an oligomer contains randomly distributed ether and C—C linkages, one has to consider that each monomer that is produced had to have been surrounded by two ether linkages, which leads to the following correlation for monomer yield prediction:

$$\text{Monomers yield (mol\%)} = \frac{(\mathrm{n}-2) * (\beta - \mathrm{O} - 4\ \text{content})^2}{\mathrm{n}} + \frac{2 * (\beta - \mathrm{O} - 4\ \text{content})}{\mathrm{n}} \qquad \text{(Equation 4)}$$

The β—O—4 content and monomers yield given above are molar ratios and n is the number of monolignols in a given chain of lignin polymer (chain length). For large n, this formula becomes:

$$\text{Monomers yield (mol\%)} \approx (\beta - \mathrm{O} - 4\ \text{content})^2 \qquad \text{(Equation 5)}$$

Because, native lignin is often assumed to have a large chain length the aforementioned theoretical monomer yield for lignin based on ether bond cleavage is often calculated based Equation 5.

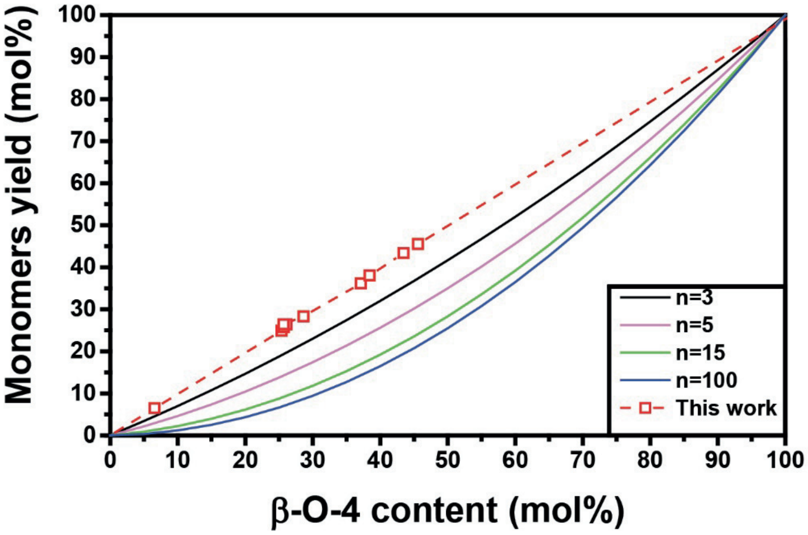


Figure 34 Comparison of different monomer yield models based on extracted lignin.
The model used in this work is based on Equation 1, whereas the other models (for varying chain length) are based on the model proposed by Hodge et al.189 and are given by Equation 1, whereas the other models (for varying chain length) are based on the model proposed by Hodge et al.189 and are given by Equation 4.

Nevertheless, the total monomer yields predicted here based on the simple assumption that one β—O—4 linkage produces one monomer after hydrogenolysis showed excellent agreement with experimental results (within 3%) (Figure 34 and 35). The more complex model that that assumes randomly distributed C—C and ether linkages within all oligomers (Equation 4) performed far worse compared to experiments, regardless of the chain length that was chosen (Figure 34). This comparison suggests that, for extracted lignin, β—O—4 and C—C linkages are not randomly distributed. In fact, our results point to the presence of oligomers that are largely formed with just β—O—4 linkages and oligomers that are condensed and contain almost only C—C linkages. In contrast, the accurate predictions of maximum hydrogenolysis yield from native lignin based on ether linkage content in past studies using Equation 5 suggest that, linkages are randomly distributed in native lignin. Together, these observations suggest that lignin units that are linked by at least one interunit C—C linkage are likelier to condense during extraction compared to lignin units, which would favor the formation of separate groups of condensed oligomers and oligomers containing mostly β—O—4 linkages.

As previously mentioned, a chain consisting of only β—O—4 linkages with n linkages should results in $n + 1$ monomers, while our model neglects these end effects and assumes the production of just n monomers. Past measurements by gel permeation chromatography had shown that formaldehyde-stabilized lignin had an approximate average chain length of about $n = 15$ units.[5] Therefore, in such a case, the model presented in Equation 1 should underestimate the monomer yield by about 7%. The fact that this systematic error does not occur could be explained by the fact that the hydrogenoylsis yield of the ether linkages is a bit lower than 100%, which could be compensating the underestimation.

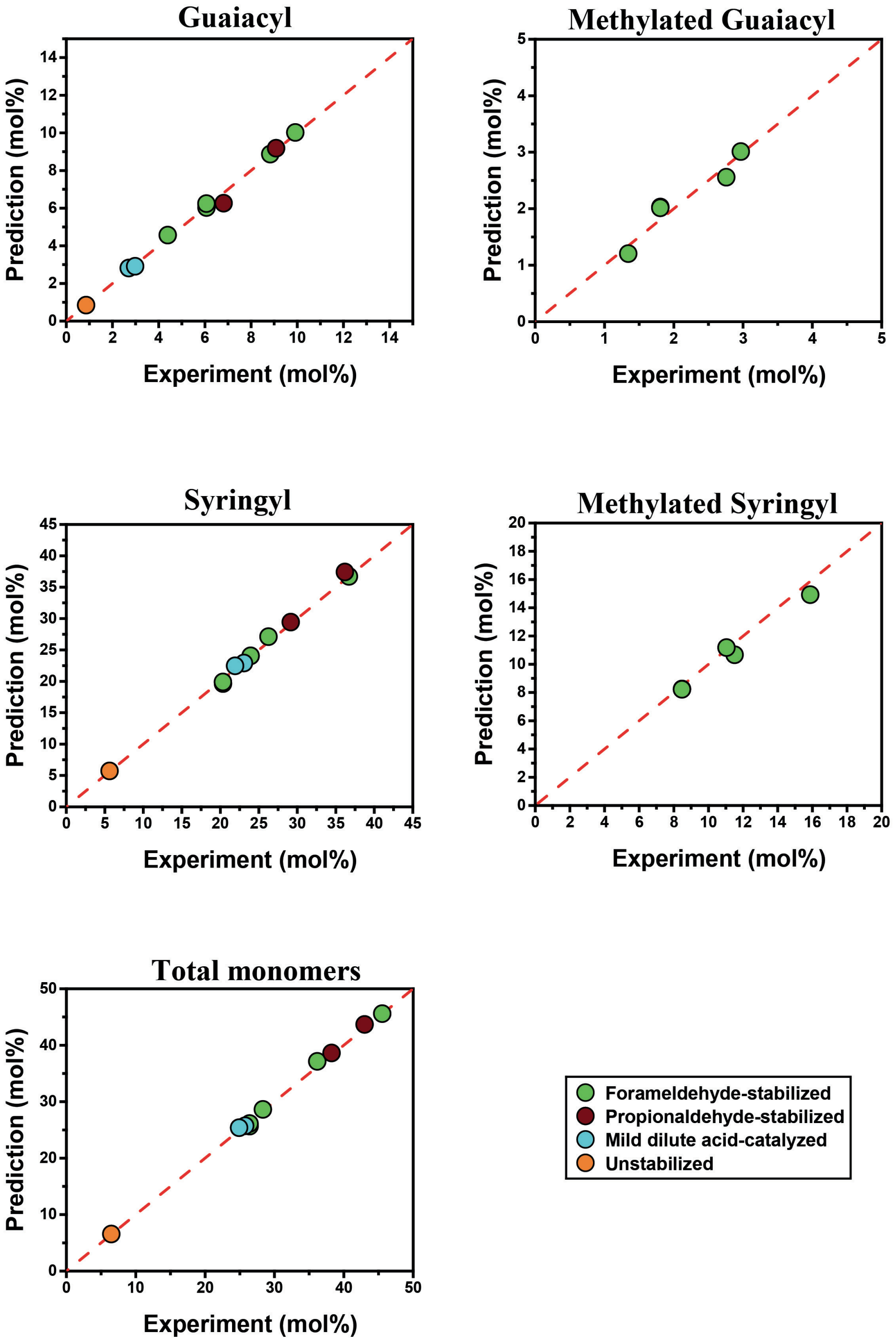


Figure 35 Predicted yields versus experimental yields.
The yields are in moles per moles of lignin units.

As previously mentioned, we could also use our model to predict the distribution of individual monomers (Figure 35). In doing so, we observed a trend where the groups that were present in smaller quantities led to more deviations between their predictions and experimental values, presumably due to error in the cross-peak integration and, to a lesser extent, quantification after hydrogenolysis. The predicted quantities of syringyl monomers, which were most abundant, showed similar accuracy to the total monomer quantities (< 4%). In comparison, the predicted quantities of methylated units based on the quantity of hydroxymethylated units detected in the HSQC spectrum showed much larger deviations (< 8%) probably because they were 3 times less prominent than syringyl units. Nevertheless, all predictions remained accurate across a wide range of monomer yields (6% to 46%) for both stabilized and non-stabilized isolated lignins.

Such successful predictions clearly demonstrate that the structure of lignin is a key determining factor controlling its ability to be depolymerized. Furthermore, we can now provide a non-destructive method to very accurately predict lignin upgrading yields based on its 2-D HSQC NMR spectra. The very clear relationship between structure and upgradeability could also further guide the development of new structure-function relationships for lignin and, as demonstrated here, insights into the distribution of functionalities along the lignin polymer.

Chapter 5 Conclusion

5.1 Summary of the work

As mentioned in the objectives, the primary goal of the work in this thesis was to develop a method for valorizing lignin in a process that is compatible with existing biorefineries. Therefore, after presenting a review of conventional pretreatments of lignocellulosic biomass and their mechanisms, some of the challenges faced for lignin valorization by these processes were discussed (Chapter 1). Subsequent of this thesis is research which has been done to provide a better understanding and also to facilitate finding a proper solution for the discussed challenges.

In chapter 2, we developed a high-yield lignin extraction and depolymerization process that is compatible with most biorefinery schemes. The process was optimized towards narrower product distribution by both using a genetically modified biomass, and optimizing the catalyst's nature and process parameters so that a single product could be produced at high selectivity.

In chapter 3, a fractionation method was presented for isolating each of the three biopolymers of lignocellulosic biomass and/or their derivatives. Furthermore, this method was able to produce isolated fractions of the three main constituents of lignocellulosic biomass at a 100-g scale with high extraction yields and retain full upgradability. Such a method paved the way for further process development and scale up as well as making large quantities of uncondensed, bench-stable lignin available for further upgrading studies.

Finally, in chapter 4, we showed the effectiveness of a non-destructive analysis method based on 2D-HSQC NMR for simultaneous identification and quantification of the structural features of lignin. In contrast with standard 2D-HSQC NMR, this method could be used to very accurately quantify the concentration of lignin's chemical groups. This accurate measurement led to very accurate predictions of monomer yields which were in agreement with the experimental results within 3% error.

5.2 Outlook and future research

The work presented here led to new insights for the valorization of not only lignin but also lignocellulosic biomass in a broader sense. By providing a method for production of fully upgradable lignin at large scale, an unprecedented quantity of lignin can be used in many studies notably to find pathways for further upgrading. For example, the produced monomers from depolymerization of aldehyde-stabilized lignin can be studied for further upgrading by oligomerization and/or polymerization reaction for production of higher value-added bioproducts. Moreover, the only depolymerization method used in this work is hydrogenolysis which is a reductive process. The availability of uncondensed bench-stable lignin allows for the possibility of developing other high yield depolymerization methods including those using oxidative, acid/base, thermal, or solvolytic chemistry. This could lead to a wider range of products that can be produced from stabilized lignin. In addition, this form of lignin could be used to develop theoretical knowledge of lignin transformation including understanding the kinetics and mechanism of stabilization and depolymerization reactions. This can be done by first studying the stabilization kinetics by using model compounds and further evaluating their stability and deprotection kinetics during depolymerization compared to the isolated stabilized lignin. Finally, another interesting topic would be to further study the aldehyde-treated carbohydrates that are produced with this method, especially hemicellulose derivatives. Studying the upgrading of aldehyde-functionalized xylose could lead to new compounds and routes for the valorization of this fraction. So far, it has been used to produce xylose and furfural by hydrolysis. However, there is a very high potential for studies on this subject, as these product have not been explored in depth in the past.

In the last part of this thesis I presented work for adapting a novel quantitative 2D HSQC NMR method for identification and quantification of structural features in isolated lignin. This method could be used to predict the highest depolymerization yields that can be achieved from a sample alongside with the distribution of lignin's chemical groups. In contrast with destructive methods such as performing a full depolymerization for which the yields are dependent on reaction conditions, the prediction results for this NMR based method is independent of such conditions. Therefore, this method could be used for rapid evaluation of lignin quality no matter how it was produced.

Another aspect that was introduced in this thesis was that the use of genetically modified biomass could be used to increase monomers yield and selectivity. This concept allowed us to address the challenge of lignin degradation that is mentioned in section 1.6. We showed the possibilities of the use of such sources of biomass in chapter 2. Therefore, studies on different routes for genetically modifying the biomass can be an interesting topic to further increase the final bioproduct yields. That being said, cultivating and harvesting such sources of biomass can have unexpected environmental impacts. This issue should be addressed by setting strict regulations and procedures for testing the effects of cultivation of transgenic biomass. Furthermore, the large-scale cultivation must be controlled by continuous observation, as a measure to prevent adverse environmental impacts. Once these potential consequences are better understood, we will be able to understand whether or not the added benefits of using these forms of biomass are worth the potential risks.

Along with studies on different upgrading routes and kinetic studies, considering the novelty of this method, the unprecedented results, and its compatibility with conventional biorefineries, there is a lot of potential for further research on the topics of intensifying and optimizing the process by experiments, or process simulation. Combination with other upgrading routes and performing life cycle assessment studies could illuminate the wider environmental and economic implications of integrating stabilization into biorefineries. The research presented here enabled us to perform the fractionation of lignocellulose at kilogram-scale, which presages the potential for pilot-scale production and further implementing industrial-scale units.

Appendix A Appendix for chapter 2

A.1 Materials and methods

A.1.1 Chemicals and materials

All commercial chemicals were analytical reagents, and were used without further purification. 5% Ru on Carbon, guaiacol (2-methoxy-phenol, 98%, **4**), ethylguaiacol (4-ethyl-2-methoxy-phenol, >97%), propylguaiacol (4-propyl-2-methoxy-phenol, >99%), syringol (1,3-dimethoxy-2-hydroxybenzene, 99%), propylbenzene (98%), pyridine (anhydrous, 99.8%), decane (>99%), Zinc nanopowder (<50 nm particle size, >99% trace metals basis), BSTFA (N, O-bis(trimethylsilyl)trifluoroacetamide, >99%), sodium bicarbonate, and hexane were purchased from Sigma Aldrich. Vanillin (4-hydroxy-3-methoxybenzaldehyde >99%), ethyl vanillin (3-ethoxy-4-hydroxybenzaldehyde >97%), acetyl bromide (>98%) were purchased from Arcos Organic. Vanillic acid (4-hydroxy-3-methoxybenzoic acid, >98%), syringaldehyde (4-hydroxy-3,5-dimethoxybenzaldehyde, >98%), syringic acid (4-hydroxy-3,5-dimethoxybenzoic acid, >98%) were purchased from Alfa Aesar. Propylsyringol (4-propyl-2,6-dimethoxy-phenol, 95%), ethylsyringol (4-ethyl-2,6-dimethoxy-phenol, 95%), and propanol-syringol (3-(4-hydroxy-3,5-dimethoxyphenyl)-1-propanol 95%) were purchased from Chemspace. Methanol (>99%), tetrahydrofuran (THF, >99%), 1,4-dioxane (99%), N-Methyl-N-(trimethylsilyl) trifluoroacetamide (98%), and dichloromethane (>99%) were purchased from ABCR GmbH. Propanol (>98%) and vanillyl alcohol (4-(hydroxymethyl)-2-methoxy-phenol, 98%) were purchased from TCI chemicals. Fuming hydrochloric acid (37 %), sulfuric acid (95-97%), chloroform, and acetonitrile were purchased from VWR. Formaldehyde solution (37%) was purchased from Roth AG. Chloroform-d and THF-d_8 were purchased from Armar Chemical. Veratrylglycerol-β-guaiacyl ether [3-(3,4-dimethoxyphenyl)-2-(2-methoxyphenoxy)-propane-1,3-diol], >97%, **1**) was purchased from AstaTech. Beech wood (*Fagus sylvatica*) was harvested in Zollikofen (Switzerland) in October 2014 and air-dried for storage. Prior to experiments beech chips were milled to pass through a 0.45-mm screen. Water was purified using a Millipore Milli-Q Advantage A10 water purification system to a resistivity higher than 18 MΩ.cm.

The production of *F5H*-overexpressing transgenic hybrid poplar (*Populus tremula* × *P. alba*) was previously described.[190] In March 2014, three-year-old trees were coppiced (cut near ground level), sawn into ~30-cm lengths in the field, and stored in milk crates in a walk-in freezer (25 °F) for ~1 month. Stem material was then dried at 45 °C for up to 7 days (depending on diameter) and stored at room temperature until the bark was peeled using a spokeshave. Dried, bark-free stem material was knife-milled to pass through a ¼" screen by Hazen Research (Golden, CO). The whole plant cell wall gel-state Nuclear Magnetic Resonance (NMR) samples were prepared as previously described.[191] Briefly, the dried cell wall sample was pre-ground for 30 s in a Retsch MM400 mixer mill at 30 Hz, using zirconium dioxide (ZrO_2) vessels (10 mL) containing ZrO_2 ball bearings (2 × 10 mm). The cell walls were extracted with distilled water (ultrasonication, 1 h, three times) and 80% ethanol (ultrasonication, 1 h, three times). The cell walls were dried and finely milled using a Fritsch planetary micro mill PULVERISETTE 7 (Germany) at 600 rpm with zirconium dioxide vessels (20 mL) containing with zirconium dioxide ball bearings (10 mm × 10). Each sample (200 mg) was ground for total 2 h 40 min (interval: 10 min., break: 5 min, repeated 11x). The cell walls were collected directly into the NMR tubes (50 mg for each sample) and gels formed using 0.5 mL DMSO-d_6/pyridine-d_5 (4:1). NMR experiments for the whole plant cell wall gel-state samples were performed as previously reported[191] and are briefly described in Section A1.3.

Compositional data derived from both standard degradative methods and NMR analyses (Figure A8) of the entire lignin component (as well as the isolated lignin fraction) indicated that the lignin from this transgenic poplar is comprised of 98.3% syringyl (S) units, the remainder being guaiacyl (G) units and traces of hydroxyphenyl (H) units; the syringyl level in the wild-type control is 68%. This lignin is more linear and displays both a lower degree of polymerization and reduced inter-unit C–C linkages.[192]

A.1.2 Experimental methods

(1) Lignin model compound reactions

Veratrylglycerol-β-guaiacyl ether (VG): In a 15-mL glass vial, 25 mg of veratrylglycerol-β-guaiacyl ether (VG) was mixed with 900 µL of 1,4-dioxane, 42 µL of an HCl solution (36.5-37 wt%, 18 mg HCl, and 31 µL water), and 100 µL of a formaldehyde (FA) solution (36.5 wt% in water, 40 mg FA and 69 µL water). For the control experiments (without FA), 69 µL of water were added instead of the 100 µL of FA solution. The reaction was conducted in an oil bath at 80 ºC and stirred with a stir bar at 300 rpm. Aliquots of 100 µL of the reaction mixture were sampled at specific times to analyze the products using the methods described in Section A1.3. Briefly, 50 µL of the sampled solution was diluted with 500 µL 1,4-dioxane and 100 µL of an internal standard (8 mg decane dissolved in 5 mL 1,4-dioxane) and used for gas chromatography with flame ionization detector (GC-FID) analysis. In cases where dimers were present, silylation was performed to increase product volatility. In such cases, 50 µL of the sampled solution was mixed with 250 µL N-Methyl-N-trimethylsilyl-trifluoroacetamide (MSTFA), 250 µL pyridine, and 100 µL of the internal standard solution (8 mg decane dissolved in 5 mL 1,4-dioxane) and kept at room temperature for 2 h prior to GC-FID or gas chromatography-mass spectroscopy (GC-MS) analysis.

The intermediate and final reaction products obtained in the absence of FA and identified by GC-MS were consistent with previous reports.[142] However, the identities of the final products were verified by further characterization. The identity of guaiacol was confirmed by comparison of the GC-FID retention time to that of an authentic standard. The two other products were purified by high-performance liquid chromatography (HPLC) using an automated fraction collector and characterized by ^{1}H NMR (Figure A1D an A1E). An Agilent Infinity 1260 HPLC system equipped with UV-Vis Detector and a Zorbax SB-C18 column at 30 °C was used for separation. Water and acetonitrile (ACN) (1 mg/L ammonia formate) were used as the mobile phases for separation in a gradient elution. Optimal solvent gradients and flow were developed for each separation, ranging from 5% to 95% ACN over 30 to 40 min and flows varying between 0.5 and 0.8 mL/min.

The intermediates and products of the reaction performed in the presence of FA were identified by GC-MS (Figure A10) and their identities were confirmed by 2D heteronuclear single-quantum coherence nuclear magnetic resonance (HSQC NMR) (Figure 18F, A1C and A1D). Specifically, the unmethylated products were identical to those purified and authenticated in the case where no FA was added (Figure A1A and A1B), but methylation of other compounds was confirmed by 2D-HSQC NMR analysis of the product mixture (Figure A1C and D). A 2D-HSQC NMR spectrum of the intermediate was obtained after 30 min of reaction with FA (Figure 18F).

Vanillyl alcohol (VA): In a 15-mL glass vial, 100 mg of VA was mixed with 900 µL of 1,4-dioxane, 42 µL of an HCl solution (36.5-37 wt%, 18 mg HCl and 31 µL water), and 100 µL of an FA solution (36.5 wt% in water, 40 mg FA, and 69 µL water). For the control experiment without FA, 69 µL of water instead of 100 µL of the FA solution were added. The reaction was conducted in an oil bath at 80 ºC and stirred with a stir bar at 300 rpm. Aliquots of 100 µL of the reaction mixture were sampled at specific times for analysis using Gel Permeation Chromatography (GPC) and NMR spectroscopy. Specifically, 50 µL of the sampled solution was diluted with 5 mL water and then extracted with 20 mL chloroform three times. The organic phases were combined and the solvent was removed under reduced pressure using a rotary evaporator at 60 ºC. The dried residue was dissolved in deuterated chloroform (chloroform-d) for NMR analysis (Section A1.3) or in THF for GPC analysis.

(2) Lignin extraction

In a 50-mL glass vial, 1 g of air-dried beech wood particles were mixed with 9 mL of 1,4-dioxane, 420 µL of an HCl solution (36.5-37%, 180 mg HCl and 315 µL water), and 1 mL of a FA solution (36.5% in water, 400 mg FA and 690 µL water). For the control experiment without FA, 690 µL of water was added instead of the FA solution. The reaction was conducted in an oil bath set at the specified temperature and stirred by a stir bar at 300 rpm. After the reaction, the slurry was filtered and washed with acetone until the filtrate was colorless. The filtrate was then neutralized by addition of a bicarbonate solution (~420 mg in 5 mL water). The solvent was removed with a rotary evaporator at 60 ºC. The dried residue was then dissolved in 25 mL THF to extract lignin, leaving the salt and carbohydrates as precipitates. 20 mL of the resulting lignin-THF solution was used for hydrogenolysis.

When the 1,4-dioxane mixture was used directly for hydrogenolysis, the slurry after reaction was filtered and washed with 10 mL 1,4-dioxane. The filtrate was then neutralized with a sodium bicarbonate solution (~420 mg in 5 mL water). 1,4-Dioxane was added to the neutralized solution to reach 25 mL and centrifuged for 5 min at 4600 rpm to remove any precipitated salts. 20 mL of the resulting lignin-1,4-dioxane solution was used for hydrogenolysis.

The amount of extracted lignin was determined by subtracting the amount of Klason lignin in extracted residue from the amount of Klason lignin in untreated wood powder. Klason lignin determination is detailed in Section A1.3.

(3) Hydrogenolysis of model compounds and lignin into lignin monomers

In cases where extracted lignin was used as a feedstock, 20 mL of the lignin solution was added to a 50-mL high-pressure Parr reactor along with 100 mg of catalyst (5 wt% Ru/C). The reactor was stirred with a magnetic stir bar and heated with high-temperature heating tape (Omega) connected to a variable power supply controlled by a PID temperature controller (Omega) with a K-type thermocouple that measured the reaction temperature through a thermowell. Once closed, the reactor was purged three times and then pressurized with 40 bar of H_2. The reactor was heated to the desired temperature and then held at that temperature for the specified residence time. After reaction, the reactor was cooled with an external flow of compressed air to room temperature. A sample of the resulting liquid was taken for GC analysis (Figure A11).

In the cases where direct hydrogenolysis was performed on native lignin, 1 g of wood powder was mixed with 20 mL of THF or methanol and 200 mg of the Ru/C catalyst. The remaining procedure was performed as described above.

In the cases where a lignin model compound was used as the feedstock, the solution of 25 mg VG in 1 mL 1,4-dioxane-water (with or without FA) obtained after reaction was neutralized with a sodium bicarbonate solution (42 mg in 500 μL water). THF was added to reach 20 mL. The solution was centrifuged for 5 min at 4600 rpm to remove any precipitated salt and used for hydrogenolysis. 5 mg of Ru/C was added into the reactor, which was then purged and pressurized with 20 bar of H_2. The remaining procedure was performed as described above.

(4) Preparation of methylated lignin monomer standards

All authenticated standards of un-methylated compounds were reacted with FA (25 mg of the authenticated standard was mixed with 18 mg HCl, and 20-40 mg of FA in 1 mL of 1,4-dioxane/water (9:1) at 100 ºC for 2 h) followed by neutralization and hydrogenolysis (with 20 bar H_2 at 200 ºC for 15 h). The GC chromatograms of resultant product mixtures were compared with the GC chromatogram of biomass-derived lignin monomer mixtures, as shown in Figure A2. To obtain pure methylated lignin monomer standards, each resultant mixture was separated by reverse-phase HPLC and the fraction of the targeted methylated monomer was collected using an automated fraction collector.

Separation of methylated lignin monomers was conducted using an Agilent Infinity 1260 HPLC system equipped with UV-Vis Detector and a Zorbax SB-C18 column at 30 °C using water and acetonitrile (ACN) (1 mg/L ammonia formate) as the mobile phases with gradient elution. An optimal combination of solvent gradients and volumetric flow was developed for each separation. Solvent gradients from 5% to 95% ACN over 30 to 40 min were used with flows varying between 0.5 and 0.8 mL/min.

The collected liquid was dried under reduced pressure in a rotary evaporator at 40 ºC and the resultant residue was used for GC, GC-MS, ESI-MS (Figure A12, A13 A14, and Table A2) and NMR analyses (Figure A15, A16 and A17). Comparison between GC and GC-MS spectra of the authenticated methylated lignin monomer standards and the compounds obtained in wood-derived mixtures were used to confirm the identity of these lignin monomers.

(5) Preparation of diformyl-xylose

Xylose (2 g), 1.8 g HCl solution (36.5-37 wt%), and 4.0 g FA were mixed in 100 mL of 1,4-dioxane/water (9:1) and left to react at 80 ºC for 30 min. The resultant solution was neutralized using sodium bicarbonate and dried under reduced pressure with a rotary evaporator at 60 ºC. The resultant residue was distilled at 120 ºC, 0.1 mbar to obtain diformyl-xylose as a yellowish solid. The residue was then extracted with 500 mL hexane three times. The organic phases were combined and dried under reduced pressure with a rotary evaporator at 60 ºC. The resultant yellowish solid was recrystallized in ethanol twice. Finally, a white solid crystal was obtained as the diformyl-xylose standard. The solid was dissolved in 1,4-dioxane for GC-MS analysis and chloroform-d for NMR analysis (Figure A6). The authenticated standard was compared to the molecule identified in the lignin monomer mixture (Figure A12 and A13).

(6) Hydrolysis of diformyl-xylose

Extraction liquors containing diformyl-xylose were diluted 10 times with a 3 wt% H_2SO_4 solution and incubated in an autoclave at 120 ºC for 1 h. The resultant liquor was used for HPLC analysis to determine xylose quantities.

(7) Enzymatic hydrolysis of pretreated substrates

Enzymatic hydrolysis of the pretreated substrates was carried out at 50 °C using a shaking incubator (New Brunswick Scientific, Model 126) at 250 rev/min. The solid substrate resulting from the pretreatment of 1 g of biomass (~0.3-0.4 g glucan) was loaded into a 15 mL glass vial. 1 mL of tetracycline chloride solution (40 mg/mL) was added to prevent microbial growth and prevent consumption of liberated sugars. Acetate buffer (pH=5, bought from Fisher, catalogue No. 25862-0010) was added to make the final liquid volume 12 mL (~12 g). The slurry was then put into the incubator at 50 °C for 1 h and then was neutralized with acetic acid to form a pH=5. After that, the acetate buffer was added to make the final liquid volume 12 mL (~12 g). Cellulase (30 FPU (Filter Paper Units) of CTech2 per gram glucan) was loaded into the glass vial. In some cases, a stir bar was added and vortexed with the substrate to facilitate the defibrillation and dispersion of cellulose. 200 µL of liquid was sampled at specific intervals for sugar yield determination by HPLC.

Note: in all cases, the temperature of the oil bath was set 3 °C higher than the specified reaction temperature to compensate for differences between the wall and interior of the reactor.

A.1.3 Analytical methods

(1) Compositional analysis of biomass and biomass liquors

The compositional analysis of biomass and substrate after lignin extraction followed a standard TAPPI method [193]. The moisture content of biomass was determined by drying wet biomass at 105°C overnight followed by cooling in a desiccator for 2 h. Before conducting composition analysis, 2 g of biomass samples were extracted with an ethanol/water mixture (40 mL) (ethanol/water, 4:1 (v/v)) three times followed by pure water (40 mL) extraction three times. Each extraction was conducted at room temperature with sonication for 30 min. After each extraction, the extractant was removed by centrifugation. The extracted biomass samples were dried at 105ºC overnight and placed in a desiccator for 2 h and ready for use. The mass loss after extraction corrected for moisture content was used to determine the quantity of extractives. Extracted wood particles or the dried pretreated substrate (0.25 - 0.50 g) were loaded into 50-mL beakers with the addition of 7.5 mL of a 72 wt% H_2SO_4 solution. The mixture was left at room temperature for 2 h and stirred with a glass rod every 10 min. Afterwards, the slurry was transferred into a round-bottom-flask and 290 mL of water were added to reach a H_2SO_4 concentration of 3 wt%. The glass bottle was sealed with a screw cap and heated to 120 ºC for 1 h in an autoclave. The resultant solution was filtered and the filtrate was used for sugar analysis by HPLC. The precipitate was washed with water and dried at 105 ºC and weighed to determine Klason lignin. Composition analysis results are presented in Table A3.

For the analysis of pretreatment liquor, the slurry after pretreatment/lignin extraction was filtered and washed with 30 mL water and analyzed by HPLC. To hydrolyze possible oligomers or acetalized sugars into monomeric sugars, 20 μL concentrated H_2SO_4 was added to 1 mL of the filtrate and heated to 120 ºC for 1 h in an autoclave. The resulting mixture was again analyzed by HPLC.

HPLC analysis of the sugars was conducted using an Agilent Infinity 1260 HPLC system equipped with a Refractive Index Detector and a Bio-Rad Aminex HPX-87P column at 80°C using water as the mobile phase and a flow rate of 0.6 mL/min.

Analysis of furfural and HMF was conducted using an Agilent Infinity 1260 HPLC equipped with UV-Vis Detector and a Bio-Rad Aminex HPX-87H column at 80 °C using 5 mM H_2SO_4 in water as the mobile phase and a flow rate of 0.6 mL/min.

(2) Lignin monomer analysis by DFRC method and nitrobenzene oxidation method

(i) Alkaline Nitrobenzene oxidation (NBO)

Extracted biomass (40 mg) was mixed with nitrobenzene (0.4 mL) and 2M NaOH (7 mL) and reacted at 170 ºC for 2 hours in an oil bath. Afterwards, the reactor was cooled in ice water and 1 mL of freshly prepared ethyl vanillin (3-ethoxy-4-hydroxybenzaldehyde **EV**) (5 μmol/mL) in 0.1 M NaOH solution was added to the reaction mixture as an internal standard. The mixture was transferred to a 100 mL separatory funnel and washed with 15 mL of dichloromethane three times. The remaining aqueous layer was acidified with 2 M HCl, until the pH was below 3.0 and extracted with 20 mL of dichloromethane twice and 20 mL of diethyl ether. The combined organic layer was washed with deionized water (20 mL) and dried over Na_2SO_4. After filtration, the filtrate was collected in a 100 ml pear-shaped flask and dried under reduced pressure. For the TMS (trimethylsilyl) derivatization step, NBO-products were washed with pyridine (3 × 200 µL) into a GC vial and BSTFA (150 µL) was added. The mixture was heated to 50 °C for 30 min. The silylated NBO-products were analyzed by GC-MS (Shimadzu GC2010/PARVUM2, IC-1 column) equipped with a fused-silica capillary column (30 m × 0.25 µm film, SHR5XLB capillary column, Shimadzu Co.) to identify the products by the comparison with the peak retention time and mass spectra of the authentic compounds. The identified products were quantified by GC-FID (Shimadzu GC-2014) using the same column and the following response factors (1.03, **V**; 0.96, **VA**; 0.87, **S**; 0.77, **SA**) determined by running mixtures of those pure monomers and internal standard (**EV**) at the same time. Initial column temperature: 150 ºC (held for 10 min), raised at 5 ºC/min to 280 ºC (held for 20 min). Total running time 56 min. Retention times: 6.5 min, **V**; 8.2 min, **EV**; 12.3 min, **S**; 14.4 min, **VA**; 18.3 min, **SA**.

(ii) Derivatization Followed by Reductive Cleavage (DFRC)

10 mg of the extracted biomass sample suspended in acetyl bromide-acetic acid solution (1:4 v/v, 3.0 mL) was added to a 2-dram vial containing a magnetic stir bar. The vial was sealed with a PTFE-lined cap and heated to 50 °C for 2.5 hours, after which the solvent was removed under reduced pressure on a SpeedVac. The resulting residue was treated with ethanol (0.5 mL), followed by removal of the solvent on the SpeedVac (50 °C, 15 min, 35 torr/min, 1.0 torr). The residue was redissolved in a mixture of dioxane/acetic acid/water (5:4:1 v/v, 4 mL) with the addition of nano-powder zinc (approximately 50 mg) and stirred at room temperature for 20 hours. The resultant liquid was transferred to a separatory funnel containing saturated ammonium chloride (15 mL) and the corresponding deuterated internal standards, using dichloromethane to rinse the reaction vessel (3 × 1 mL). The organic phase was separated, and the aqueous phase was extracted with dichloromethane (4 × 10 mL). The combined organic fractions were dried over anhydrous sodium sulfate, filtered, and concentrated *in vacuo*. The residue was then reacted with pyridine/acetic anhydride (1:1 v/v, 5 mL) overnight to assure full acetylation of the product. The solvent was then removed under reduced pressure using a rotary evaporator. The crude oil was loaded onto the LC-Si SPE cartridge using dichloromethane (3 × 1 mL) and eluted with hexanes/ethyl acetate (1:1, 10 mL). The final products were transferred quantitatively to a GC vial and then injected on a GCMS-TQ-instrument (Shimadzu GCMS-TQ8030 triple quadrupole GC/MS/MS). The identified products (coniferyl alcohol diacetate **CD** and sinapyl alcohol diacetate **SD**) were quantified in multiple reaction-monitoring (MRM) mode, determined by running mixtures of those pure monomers and deuterated internal standard

(3) Lignin monomer analysis by GC and GC-MS

To analyze lignin monomers after hydrogenolysis, 1 mL of the resultant solution was directly sampled for analysis without any further treatment other than the addition of 100 µL of prepared internal standard (8 mg decane in 5 mL 1,4-dioxane). The solution (~1.1 mL) was analyzed with a GC (Agilent 7890B series) equipped with an HP5-column and a flame ionization detector (FID). The injection temperature was 300 °C. The column temperature program was: 40 °C (3 min), 30 °C/min to 100 °C, 40 °C/min to 300 °C and 300 °C (5 min). The detection temperature was 300 °C. Sensitivity factors of the products were obtained by using estimates based on the effective carbon number due to lack of commercial standards. Yield calculations and validation experiments for this method are detailed in Section A2.1.

Identification of monomer peaks in the GC-FID chromatograms was performed initially by GC-MS using an Agilent 7890B series GC equipped with a HP5-MS capillary column and an Agilent 5977A series Mass Spectroscopy detector. The peaks in the GC-MS chromatogram appear in the same orders as those in GC-FID chromatogram due to the use of a similar capillary column. The following operating conditions were used: injection temperature at 250 °C, a column temperature program of 50 °C (1 min), 15 °C/min to 300 °C and 300 °C (7 min), and a detection temperature of 290 °C.

(4) NMR analysis of lignin or lignin model compounds

Following the model compound reaction described in Section A1.2, after 30 min of reaction time, the mixture of 25 mg of VG in 1,4-dioxane/water (with or without FA) was neutralized and dried under reduced pressure with a rotary evaporator at 60 °C. The dried sample was dissolved in 600-1000 µL chloroform-d and transferred into NMR sample tubes. NMR spectra were acquired on a Bruker Avance III 400 MHz spectrometer. The chloroform solvent peak was used as an internal reference (δ_C, 77.2 ppm; δ_H, 7.24 ppm). In the case where extracted lignin was analyzed, ~30 mg of extracted lignin was dissolved in 800 µL THF-d_8 and transferred into NMR sample tubes. The analysis was performed as described above.

(5) Gel-NMR (2D HSQC) experiment of ball-milled poplar whole cell wall

NMR spectra were acquired on a Bruker Biospin (Billerica, MA) Avance 700 MHz spectrometer equipped with a 5 mm QCI $^1H/^{31}P/^{13}C/^{15}N$ gradient cryoprobe with inverse geometry (proton coils closest to the sample). The central DMSO solvent peak was used as internal reference (δ_C 39.5, δ_H 2.49 ppm). The ^1H–^{13}C correlation experiment was an adiabatic HSQC experiment (Bruker standard pulse sequence ‘hsqcetgpsisp.2’; phase-sensitive gradient-edited-2D HSQC using adiabatic pulses for inversion and refocusing). HSQC experiments were carried out using the following parameters: acquired from 11.5 to - 0.5 ppm in F2 (^1H) with 1682 data points (acquisition time 100 ms), 215 to -5 ppm in F1 (^{13}C) with 620 increments (F1 acquisition time 8 ms) of 100 scans with a 500 ms interscan delay; the d_{24} delay was set to 0.86 ms (1/8J, J = 145 Hz). The total acquisition time was 11 h. Processing used typical matched Gaussian apodization (GB = 0.001, LB = -0.5) in F2 and squared cosine-bell and one level of linear prediction (32 coefficients) in F1. Volume integration of contours used Bruker’s TopSpin 3.2 (Mac version) software, and was carried out on data that was not subjected to linear prediction.

(6) Gel permeation chromatography (GPC) analysis

The preparation of samples was described in Section 3.1. GPC analysis was conducted using an Agilent Infinity 1260 HPLC equipped with a Refractive Index Detector and an Agilent PLgel MIXED C column at 40 °C using THF as the mobile phase and a flow rate of 1 mL/min.

A.2 Experimental studies

A.2.1 Monomer identification and quantificaiton

(1) Monomer identification

Initial monomer identification was performed by GC-MS. The GC-MS spectra of identified intermediates and monomers from the lignin dimer model VG (**1**) are shown in Figure A10. A GC chromatogram from a beech-derived sample is shown with the identified lignin monomers in Figure A10.

Given the complexity of lignin monomer mixtures, whenever possible, we compared the GC retention time and GC-MS spectra of the lignin monomer mixture with authenticated standards. Authentic standards were available from commercial sources for ethylguaiaicol, propylguaiacol, ethylsyringol, propylsyringol, and propanolsyringol. The characterization of these lignin monomer standards and the comparison of these compounds with those identified in the lignin monomer mixture are shown in Figures A11-A14. The identity of methylated lignin monomers was confirmed using four separate methods. First, their mass spectra, obtained by GC-MS were compared to that of their unmethylated counterparts (Figure A14). These fragmentation spectra were all consistent with the addition of a methyl group on the aromatic ring. Second, the hexane-extracted lignin monomer mixture was characterized with 2D ^{1}H–^{13}C HSQC NMR (Figures A4A and A4C), ^{13}C NMR and ^{13}C-DEPT NMR (Figures A5B and A5D). All spectra were consistent with the addition of methyl groups onto the aromatic rings of the lignin monomers. In particular, comparison of the ^{13}C NMR, and ^{13}C-DEPT NMR spectra confirmed the presence of additional tertiary carbon atoms on the aromatic ring after reaction with FA (Figure A5). Third, all authenticated standards of unmethylated compounds were reacted with FA followed by hydrogenolysis. This procedure largely mirrored the treatment of lignin. As revealed by comparison of MS spectra and GC-FID retention time, this treatment systematically led to the formation of the same methylated monomers as those observed with real lignin (Figures A2 and A11). Finally, the five methylated lignin monomers were prepared and purified along with their isomers (Figures A14, A16 and A17). The comparison of isomers confirmed the substitution of the proton at the *para* position with a methoxyl group. However, not all isomers were detected in the wood-derived mixtures. The 2D and 1D NMR spectra of the prepared compounds are shown in Figures A15, A16 and A17 and confirmed the appearance of methyl group on the aromatic ring. The GC-FID residence times and MS spectra of these purified, authenticated molecules were identical to those of the molecules identified in the wood-derived mixtures (Figures A12 and A13). Finally, high-resolution electrospray ionization mass spectrometry (ESI-MS) was performed for the synthetized and purified standards of lignin monomers and model compound reaction intermediates. ESI-MS spectra were recorded by direct introduction of the sample at a 3 μL/min flow rate into an LTQ-Orbitrap high-resolution mass spectrometer (Thermo, San Jose, CA, USA), equipped with a conventional ESI source. The working conditions were the following: a spray voltage of 3.1 kV, a capillary voltage of 45 V, and a capillary temperature 220 °C. Their masses were compared to the theoretical masses (Table A2) further confirming their identification.

In addition to lignin monomers, diformyl-xylose was also synthetized (Section A1.2), characterized, and similarly compared to the molecules detected in wood-derived mixtures (Figure A18) confirming its identity.

(2) Monomer quantification and yield calculations

Due to the difficulty in obtaining or producing a large quantity of monomers, we used a quantification based on an internal standard (decane) and the effective carbon number (ECN) method. As stated in Section A1.3, 100 µL of an internal standard solution (8 mg decane in 5 mL 1,4-dioxane) was added to 1 mL of the solution to be analyzed. The resulting solution (~1.1 mL) was analyzed by GC-FID. The monomer yield was calculated based on the area of the monomer and the area of decane in the GC chromatogram. The detailed calculation was as follows:

$$n_{decane} = \frac{W_{decane\ in\ sample}}{MW_{decane}} \tag{A1}$$

$$n_{monomer} = \frac{A_{monomer\ in\ sample}}{A_{decane\ in\ sample}} \times n_{decane} \times \frac{ECN_{decane}}{ECN_{monomer}} \tag{A2}$$

$$W_{monomer} = n_{monomer} \times MW_{monomer} \tag{A3}$$

$$Y_{monomer} = \frac{W_{monomer} \times V}{W_{extracted\ lignin}\ or\ W_{Klason\ lignin}} \times 100\% \tag{A4}$$

In the equations,

$W_{decane\ \text{in sample}}$ (mg): the weight of decane used as an internal standard in each analyzed sample;

MW_{decane} (mg mmol^{-1}): the molecular weight of decane (142 mg mmol^{-1});

n_{decane} (mmol): the molar amount of decane in each analyzed sample;

$n_{monomer}$ (mmol): the molar amount of monomer in each analyzed sample;

$A_{monomer\ in\ sample}$: the peak area of monomer in the GC-FID chromatogram;

$A_{decane\ in\ sample}$: the peak area of decane in the GC-FID chromatogram;

ECN_{decane}: the effective carbon number (10) of decane;

$ECN_{monomer}$: the effective carbon number of the lignin monomer molecule;

$W_{monomer}$ (mg): the molecular weight of a β-O-4 bonded monomer guaiacyl glycerol (196 mg mmol^{-1}) or a β-O-4 bonded syringyl glycerol (226 mg mmol^{-1})) in the analyzed sample;

$Y_{monomer}$: the yield of monomer based on the weight of extracted lignin or original (native) Klason lignin;

$W_{extracted\ lignin}$ (mg): the weight of extracted lignin (Section A1.3);

$W_{Klason\ lignin}$ (mg): the weight of Klason lignin in the original biomass (Section A1.3); and

V (mL): the total volume of sample, 1 mL of which was used for GC analysis.

Yields were calculated based on Klason lignin and do not include acid-soluble lignin (ASL), as is consistent with a vast majority of previously reported lignin yields.[92,133,173] In other work, yields were based on lignin determined by the acetyl bromide-soluble lignin method, which gave almost identical numbers to Klason lignin.[179] ASL lignin was determined by UV absorbance using sensitivity values measured for other types of biomass than the one here, and thus are not nearly as accurate as Klason lignin measurements. For this reason, ASL is not usually used in yield calculations. However, F5H-poplar has a slightly higher proportion of ASL compared to wild-type poplar and other woods. If ASL was included in our yield calculations, our beech wood yields, which are typical of near-theoretical yields attainable from wild-type native lignin, would be around 40% (instead of 47%). In the case of F5H-poplar, the inclusion of ASL lignin in the yield calculation would result in yields around 52% (instead of 79%). Therefore, including ASL lignin in the yield calculation does not change the conclusion that reduced interunit C–C linkage levels in F5H-poplar lead to a significant increase in final monomer yields.

(3) Validation of monomer quantification

The ECN rule has been widely used to quantify carbon-containing products based on their response in GC-FID in cases where an authentic standard compound is not available or available only in limited quantitie. ECN is the sum of the contributions made by the individual carbon atoms modified by their functional group contributions. Due to the structural variations of different compounds, the accuracy can vary. In order to verify the accuracy of this rule as applied to our lignin monomer quantification, the ECNs of six different lignin model compounds were measured based on their known quantity and their GC response was compared to that of decane. This ECN calculated based on the FID response using decane as an internal standard was compared to the theoretical ECN based on this empirical rule

(Figure A19). Based on this, carbons connected only to hydrogen and carbon atoms add 1 unit to the ECN, carbon atoms in methoxy groups (ethers) are predicted to not add to the FID response, carbons connected to primary hydroxyl groups add 0.4-0.6 unit to the ECN, and carbons connected to phenolic hydroxyls typically add 0.5 unit to the ECN.[194] The only inconsistency we found is that carbon connected to phenolic hydroxyl groups in lignin monomers (compounds **1**–**5**) did not contribute to the ECN (compared to 0.4 in previous reports[194]). The contributions of carbon in lignin monomers to ECN used here are summarized in Table A4. Based on this adjusted ECN rule, the ECNs calculated experimentally matched those based on the ECN rule with errors below 1% for all compounds (Figure A19). This demonstrated the high accuracy of using decane as an internal standard and using the factor $ECN_{monomer}/ECN_{decane}$ (Equation A2) to quantify lignin monomers. Based on these results, we followed the empirical rule with the modified phenolic carbon connected to the hydroxyl group contribution of 0 for calculating the ECN of biomass-derived lignin monomers. All ECNs of lignin monomers ($ECN_{monomers}$) used for quantification are listed in Table A5.

To further verify that quantification with the ECN method and GC-FID was accurate, we compared the quantification using this method with the integration values of different protons in 2D NMR spectra, as shown in Figure A5A. In both cases, molar amounts of hydrogen in different functional groups are normalized to the total number of moles of hydrogen present in methoxy groups. When the mixture was characterized by GC-FID, the molar quantities of each compound are known. Therefore, the total number of moles of hydrogen in each functional group was also known. These various quantities were normalized so that the total amount of hydrogen in methoxy groups was set to 100 [unit]. With NMR signals, we set the integration value of hydrogen in methoxy groups to 100 [unit] (Figure A5C), and calculated the relative integration value of other protons using the MestReNova software. The small difference between the resulting calculation based on our ECN method and the NMR quantification results (<5%) further validated the accuracy of the ECN quantification of lignin monomers (Table A6).

A.2.2 Model Compound Studies with Vanillyl Alcohol

To evaluate whether blocking the negatively charged positions of lignin with hydroxymethyl groups contributes to reducing the condensation of lignin, vanillyl alcohol (VA), a model compound with only an α-hydroxyl group (without any γ-hydroxyl groups), was used to exclude the stabilization effect of the six-membered 1-3-dioxane ring structure. As with lignin, VA can polymerize under acidic conditions (Figure A20). Because of the complexity of the polymerized products resulting from treatment in acidic conditions (in the presence and absence of FA), Gel-Permeation Chromatography (GPC) and Nuclear Magnetic Resonance (NMR) were used to qualitatively evaluate a product's degree of condensation. GPC chromatograms show that the addition of FA led to a lower average molecular weight for the products (Figure A20C). After a reaction time of 5 h, monomers and dimers were still the major products with the addition of FA, while they were hardly present in the products obtained without FA. NMR spectra show that without FA the condensation of VA resulted in a decrease of the signal attributed to hydroxymethyl groups (3.5-5 ppm) (Figure A20B). With the addition of FA, the signals in this range (3.5-5 ppm) did not disappear due to the combination of the lower condensation of hydroxymethyl groups and to the appearance of newly formed hydroxymethyl groups (Figure A20B). The GPC chromatograms and NMR spectra suggest that FA reacts with the lignin model compound to block the reactive sites of the aromatic ring, thereby reducing the condensation of lignin. Nevertheless, FA reduces rather than completely eliminates the condensation reaction of VA, as VA-VA condensation and VA-FA reactions proceed simultaneously. In addition, the newly formed hydroxymethyl groups can participate in the condensation reaction if the negatively charged positions on the aromatic ring are not completely occupied (or blocked) (Figure A20A).

A.2.3 Recovering the Principal Biomass Polysaccharides after Lignin Extraction

(1) Digestibilities of substrates after lignin extraction

In addition to catalyst recovery issues, a disadvantage of direct hydrogenolysis methods is that they generally irreversibly modify a significant portion of the polysaccharide fraction in biomass. The C_5 fraction is often lost or converted to xylitol and the C_6 fraction is often partially converted to sorbitol.[133,195] The treatment of beech wood particles with acidic 9:1 1,4-dioxane:water with or without FA removed 77-79 wt% of the lignin, completely removed xylan and left in place 87-88% of glucan. In the absence of FA, xylan depolymerized and dehydrated to furfural at a 95% yield with no xylose recovered. This high furfural yield is consistent with high furan yields in the presence of organic solvents and low water concentrations.[196] With FA present, 40% of xylan was dehydrated into furfural and 50% reacted with FA to form diformyl-xylose (structure given in Figure A6). This compound has a boiling point of 250-260 ºC, is aprotic, and is soluble in hexane, making it a potential fuel additive. Even though xylose can react with FA, this reaction is reversible as diformyl-xylose can be nearly quantitatively converted back to xylose under aqueous acidic conditions (Figure A21). The remaining biomass solids were subjected to enzymatic hydrolysis to explore the depolymerization of the remaining sugars. After treatment in the absence of FA, the remaining cellulose rapidly depolymerized to glucose in the presence of enzymes (95% in 64 h, Figure A22A). However, the substrate pretreated in the presence of FA had much lower enzymatic digestibility (~10% glucose in 96 h, Figure A22A). Similar to xylose and lignin, it is likely that FA reacted with the cellulose surface to form 1,3-dioxane structures. Changes in the surface chemistry of cellulose can severely affect enzyme binding and complexation to the surface resulting in decreased enzymatic digestibility, as reported for acetic acid pretreatment and GVL (γ-valerolactone) pretreatment.[197] However, as the formation of this structure can be reversed, we treated the substrate in a dilute acid aqueous environment at mild temperatures (120 °C) for 2 h. Following this treatment, enzymatic digestibility was partially recovered leading to an increase in glucose yield from 10% to 78% in 112 h. When the same lignin extraction conditions were used with F5H-poplar, molar yields of 75% and 14% were obtained for diformyl-xylose and furfural, respectively, leading to a combined yield of 89% from xylan. Upon enzymatic hydrolysis, the substrate obtained in the absence of FA had a much higher digestibility, reaching a glucose yield of 98% in 64 h, compared to a yield of less than 10% for the substrate obtained after treatment in the presence of FA (Figure A22B). An acid treatment at 120 °C for 2 h improved the final glucose yield to 48% in 112 h of hydrolysis and an acid treatment at 140 °C for 1 h could improve the final glucose yield to 77% after 112 h of hydrolysis (Figure A22B). This difference shows that the recovery of high enzymatic digestibilities likely depends on various reaction conditions, including temperature, residence time, and acid concentration. Further optimization is planned in future work.

(2) Hydrolysis of diformyl-xylose to xylose

Diformyl-xylose is an acetal structure that can be hydrolyzed in an acidic aqueous solution. This reaction allows for the flexibility of targeting either diformyl-xylose or xylose as the final product. Our initial results (Figure A21) show that diformyl-xylose can be nearly quantitatively converted into xylose and furfural. As xylose can easily dehydrate to furfural under acidic conditions, the ratio of xylose to furfural could vary with various reaction conditions including temperature and acid concentration. Further optimization will be carried out in the future.

(3) Mass balance of the principal polysaccharides (glucan and xylan) using 1,4-dioxane and GVL as extraction solvents

In addition to using 1,4-dioxane as an extraction solvent, we also tested the effect of using GVL because it has been proven to be an effective solvent for extracting lignin in previous reports.[195,197,198] Using lignin extracted with GVL and FA (Table A1, entry 13) led to a comparable lignin monomer yield (43% vs. 47%) to those extracted with 1,4-dioxane and FA. However, using GVL led to a higher diformyl-xylose yield and less furfural compared to 1,4-dioxane (Table A7). This difference indicates that organic solvent effects can help control the selectivity of these two products, and could help target one of the other products. These effects will be further investigated in the future.

A.2.4 Formaldehyde Mass Balance and Economic Analysis

(1) FA mass balance during pretreatment

FA before and after pretreatment was measured by HPLC. After pretreatment at 100 ºC for 2h, around 56-67% of FA remained in the pretreatment liquor, which means that 33-44% of FA was consumed by reacting with lignin and polysaccharides during pretreatment (Table A8).

(2) FA consumption and recovery

We have observed that during solvent evaporation FA is easy to remove from the reaction mixture. Therefore, any residual FA in the pretreatment liquor can be easily recovered through evaporation. To investigate the viability of recovering the portion of FA that had reacted with cellulose and hemicelluloses, we conducted studies with different model saccharides present in comparable amounts to polysacharrides during pretreatment. Monomers were reacted for shorter times to reflect their limited accessibility at the start of pretreatment. The reaction mixtures were diluted with water and hydrolyzed to regenerate FA. The mass balance results (Table A9) show that FA grafted on saccharides can be recovered quantitatively by acid hydrolysis.

(3) The fate of FA during hydrogenolysis

The FA grafted on the aromatic ring (hydroxymethyl group) will be hydrogenated into a methyl group. To further understand the fate of FA in the 1,3-dioxane ring during hydrogenolysis, we used ^{13}C labeled FA to track the products during the pretreatment and hydrogenolysis of VG (**1**, a lignin model compound). After pretreatment but before hydrogenolysis, unreacted FA was evaporated under reduced pressure so that any compounds containing ^{13}C could unambiguously attributed to the decomposition of the 1,3-dioxane ring. While the absence of ^{13}C labeled compounds such as methanol was observed by ^{13}C NMR of the liquid phase after hydrogenolysis, production of 0.95 mol of $^{13}CH_4$ per initial mol of VG was detected by GC-MS. Given that we observed nearly quantitative protection of **1** by FA to form **5** and **6**, we conclude that all FA reacting to form the 1,3-dioxane ring ends up as CH_4 after hydrogenolysis.

(4) H_2 consumption during hydrogenolysis

(1) **For beech wood, both processes (direct hydrogenolysis and our method) yield about 100 mg monomers per gram of field-dry biomass (all calculations are based on data of entry 1 and 5 in Table A1).**

Direct hydrogenolysis:

We assume that all lignin present during hydrogenolysis consumes H_2 regardless of monomer production. This is likely the case because lignin units linked by C–C bonds typically contain several oxygenated moieties similar or identical to those present on lignin monomers before hydrogenolysis. We calculate the molar amount of lignin aromatic units subjected to hydrogenolysis per g biomass as: 0.21 g / 211 g·mol^{-1}=1.0 mmol (0.21 g is the amount of Klason lignin per g of field-dry biomass (6% moisture))

211 mg.mmol^{-1} is the average molecular weight of β-O-4 bonded monomeric units in native lignin (guaiacyl units (196 mg·mmol^{-1}) or syringyl units (226 mg·mmol^{-1})).

H_2 consumption for producing monomers: 1.0 mmol * 3 * 2 g·mol= **6.0 mg**

Both types of polymerized monomeric units require 3 eq. amount of H_2 to be hydrogenated to lignin monomers (propyl syringol or propyl guaiacol).

For producing 1 ton of monomers, 60 kg H_2 are needed with direct hydrogenolysis.

With a price of H_2 at \$4/kg *(48)*, the hydrogen cost of producing 1 kg monomers is: 0.060 kg * \$4/kg= **24 cents**

Hydrogenolysis of FA-extracted lignin (method reported in this article):

Molar amount of lignin monomeric units subjected to hydrogenolysis per g biomass: 0.21 g * 72% / 211 g·mol^{-1}=0.72 mmol (*"72%" is the fraction of extracted lignin based on native lignin*)

(i) H_2 consumed by lignin structure (assuming no FA addition)

H_2 consumption for producing monomers: 0.72 mmol * 3 * 2 g·mol= **4.3 mg**

Both types of polymerized monomeric units require 3 mol/mol H_2 to be hydrogenated to a lignin monomer (propyl syringol or propyl guaiacol).

(ii) H_2 consumption resultant from FA addition

H_2 consumption for reducing FA (on the dioxane ring) into methane:

1 mmol*47%*2* 2 g·mol= 1.9 mg

"2" refers to the two equivalents of H_2 required to convert FA into methane and produce compound ***8*** *(Figure 18). According to entry 5 in Table A1, about 47% (monomer yield) of aromatic units react with FA to form compound 6, which resulted in monomers after hydrogenolysis. We assume other 53% of lignin units cannot form dioxane structure due to the lack of diol structure.*

H_2 consumption for reducing the hydroxymethyl group on the aromatic ring into a methyl group:

0.72 mmol*1/4*1* 2 g·mol= 0.4 mg

"1" refers to the one equivalent of H_2 required to convert the hydroxymethyl group on the aromatic ring into a methyl group. According to entry 5 in Table A1, about 1/4 of monomers are methylated. Therefore, we assume that all lignin units (including those with condensed structures) have a 1/4 chance of reacting with FA to form hydroxymethyl groups.

Total H_2 consumption per g biomass: 4.3+1.9+0.4=**6.6 mg**

For producing 1 ton monomers, 66 kg H_2 are needed for hydrogenolysis of FA-extracted lignin.

With a price of H_2 at \$4/kg *(48),* the hydrogen cost of producing 1 kg monomers is: 0.066 kg * \$4/kg= **26 cents**

(2) **For transgenic F5H poplar, both processes yield about 100 mg monomers per g of biomass (all calculations will be based on data of entry 14 and 16 in Table A1**)

Direct hydrogenolysis:

Lignin molar amount per g biomass: 0.13 g/ 211 g.mol^{-1}=0.62 mmol (0.13 g is the amount of Klason lignin per g of field dry biomass (7% moisture))

H_2 consumption: 0.62 mmol * 3 * 2 g·mol= 3.7 mg

For producing 1 ton monomers, 37 kg H_2 is needed with the direct hydrogenolysis process.

With a price of H_2 at $4/kg *(48)*, the hydrogen cost of producing 1 kg monomers is: 0.037 kg * $4/kg= **14.8 cents**

Hydrogenolysis of FA-extracted lignin:

Molar amount of lignin monomeric units subjected to hydrogenolysis per g biomass: 0.13 g * 73% / 211 $g \cdot mol^{-1}$=0.45 mmol (*"73%" is the fraction of extracted lignin based on native lignin*)

(i) H_2 consumed by lignin structure (assuming no FA addition)

0.45 mmol * 3 * 2 g·mol= **2.7 mg**

Both types of polymerized monomeric units require 3 mol/mol H_2 to be hydrogenated to lignin monomer (propyl syringol or propyl guaiacol).

(ii) H_2 consumption resultant from FA addition

H_2 consumption for reducing FA (on the dioxane ring) into methane:

0.62 mmol*79%*2* 2 g·mol= 2.0 mg (*"2" refers to the two equivalents of H_2 required to convert FA into methane and produce compound **8** (Figure 18). According to entry 5 in Table A1, about 79% (monomer yield) of aromatic units react with FA to form compound 6, which resulted in monomers after hydrogenolysis. We assume other 21% of lignin units cannot form dioxane structure due to the lack of diol structure*)

H_2 consumption for reducing the hydroxymethyl group on the aromatic ring into a methyl group:

0.45 mmol*1/3*1* 2 g·mol= 0.3 mg

"1" refers to the one equivalent of H_2 required to convert the hydroxymethyl group on the aromatic ring into a methyl group. According to entry 16 in Table A1, about 1/4 of monomers are methylated. Therefore, we assume that all lignin units have a 1/4 chance of reacting with FA to form a hydroxylmethyl group, which will consume H_2 during hydrogenation.

Total H_2 consumption per g biomass: 2.7+2.0+0.3=5.0 mg

For producing 1 ton monomers, 50 kg H_2 is needed for hydrogenolysis of FA-extracted lignin.

With a price of H_2 at $4/kg *(48)*, the hydrogen cost of producing 1 kg monomers is: 0.050 kg * $4/kg= **20.0 cents**

With a price of phenol of \$1.1/kg *(49)*, when using beech as the feedstock, the H_2 cost represents 24% of the final monomer price with our method, which is comparable to 22% resulting from direct hydrogenolysis. When using F5H-Poplar as a feedstock, the H_2 cost represents 18% of the final monomer price with our method, compared to 14% with direct hydrogenolysis. . In summary, the comparison of H_2 costs between the two methods shows that the difference is small. The higher H_2 consumption with our method is mainly caused by the reduction of FA into methane. However, upon partial reforming of methane, H_2 can be regenerated ($CH_4 \rightarrow HCHO + 2H_2$), which can cut the consumption of H_2 by 30-40% according to aforementioned cost analysis. If methane is completely reformed into H_2 ($CH_4 + 2H_2O \rightarrow CO_2 + 4H_2$), the consumption of can be reduced by 58-80% (not including the processing cost) but will require the process to buy replacement FA.

(5) FA consumption

We have shown that FA reacted with not only lignin, but also mono- and poly- saccharides. However, the FA grafted on these saccharides could be quantitatively regenerated by aqueous acid hydrolysis of these compounds (Table A9). In Table A9, it can be seen that about 31% of FA was consumed during the reaction with saccharides. Since 33-44% of FA was consumed in the pretreatment process, it can be calculated that 2-13% of FA was consumed for stabilizing lignin, which corresponds to 0.6-1.6 mol/mol of the aromatic units in lignin.

These numbers are consistent with the FA consumption that can be calculated based on the reactions discussed during the H_2 cost analysis:

(1) Beech wood

FA consumed by forming the dioxane ring:

1 mmol*47%=0.47 mmol

FA consumed by forming hydroxymethyl groups:

0.72 mmol*1/4=0.18 mmol

Total FA consumed: 19.5 kg FA/kg monomers.

(2) F5H-poplar

FA consumed by forming the dioxane ring:

0.62 mmol*79%=0.49 mmol

FA consumed by forming hydroxymethyl:

0.45 mmol*1/3=0.15 mmol

Total FA consumed: 19.2 kg FA/kg monomers.

Assuming that the high range of FA could be consumed, this corresponds to 24-64 kg FA per ton biomass (or 240-640 kg FA / 1000 kg monomers). Producing 1 kg monomers would require 0.24-0.64 kg FA. The price of FA is approximately 0.5 $/kg. Therefore, the cost of FA is 12-32 cents per kg of monomers.

However, at a minimum, the produced methane can be sold. With a price of 0.3 $/kg methane and a production of 0.53 kg methane per kg FA, the cost of external FA will be reduced to 8-22 cents per kg of monomers. With a price of phenol at $1.1/kg, FA costs represent 7-20% of the final monomer price.

Furthermore, methane could be converted into formaldehyde with established methane reforming processes. The regenerated FA ($CH_4 + CO_2 \rightarrow 2HCHO$, more FA is generated due to the reduction of CO_2 by methane) could be used to make up the FA consumption caused by the formation of hydroxymethyl groups (see calculation above).

Therefore, in theory, no FA make-up would be needed after the initial loading of FA. Given that fully regenerating the FA would cause a surplus, an autothermal reforming process could be used to generate less FA during methane reforming methane but save minimize heating costs. Furthermore, methoxy groups could be an additional source for methanol production, which can be dehydrogenated into formaldehyde.

In conclusion, although some FA will be converted into methane and hydroxymethyl groups during hydrogenolysis, its cost is at most 20% of the final monomer price. Regenerating FA from methane or hydroxymethyl groups could potentially further reduce this cost.

A.2.5 Figures

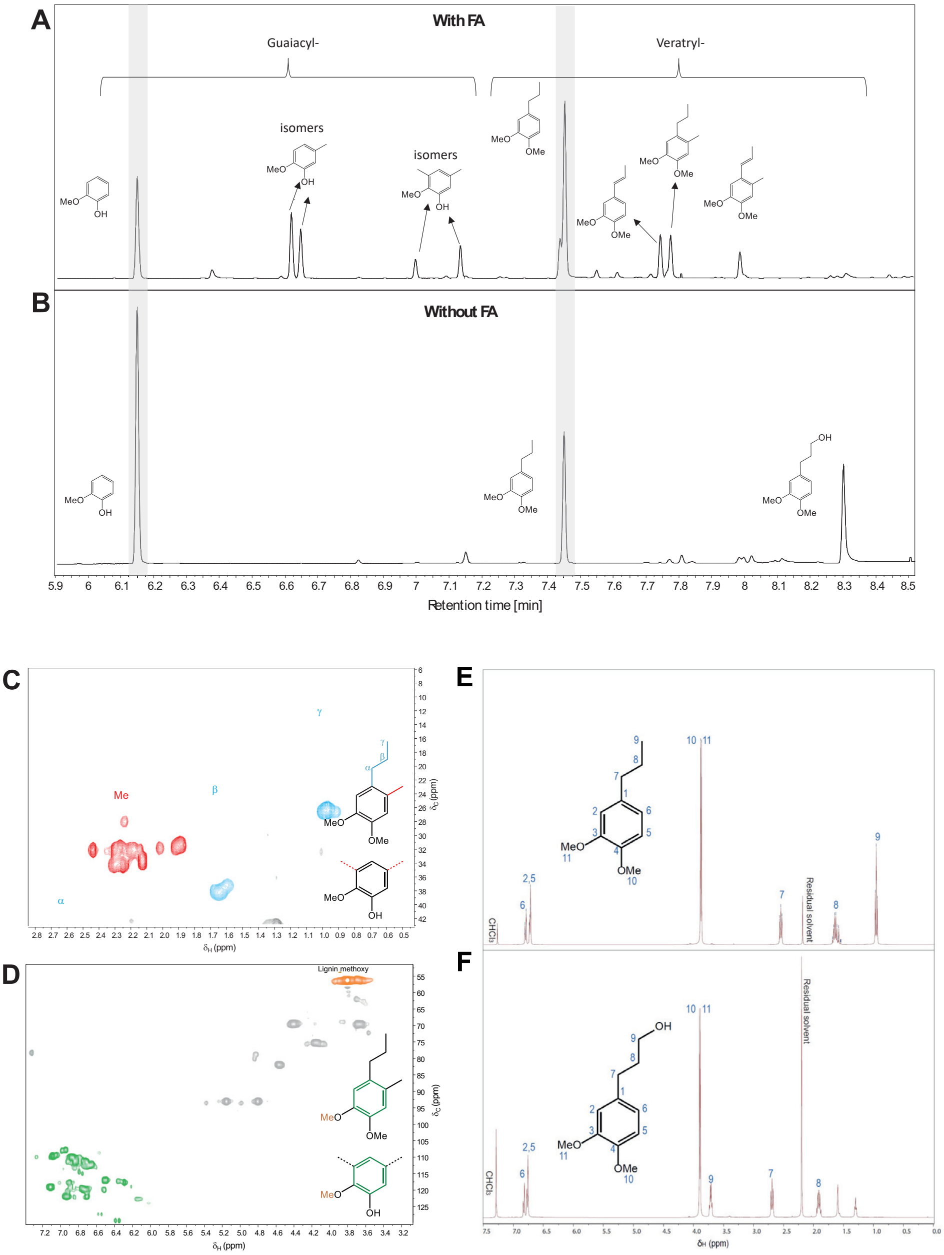


Figure A1. Acid-catalyzed depolymerization (at 80 °C for 7 h) of a lignin model compound (VG, compound 1, Figure 18) followed by hydrogenolysis (at 200 °C for 6 h, 40 bar H_2). (A) GC chromatogram of product mixture in the presence of formaldehyde (FA) (corresponding mass spectra of products are shown in Figure A10A), (B) GC chromatogram of the product mixture from treatment in the absence of FA (corresponding mass spectra of products are shown in

Figure A10B), (C) 2D HSQC NMR (Nuclear Magnetic Resonance) (side-chain region) of product mixture in the presence of FA (compounds **8** and **9** in Figure 18), (D) 2D HSQC NMR (aromatic region) of the product mixture in the presence of FA, (E) ^{1}H NMR of propylveratrol (compound **7** in Figure 18) isolated from the reaction mixture in chloroform-d and (F) ^{1}H NMR of propanolveratrol (compound **7** in Figure 18) isolated from the reaction mixture in chloroform-d. Propylveratrol and propanolveratrol were the major products from both reactions. Due to the complexity of the products in the presence of FA, 2D NMR was used to demonstrate the appearance of the methyl group on the aromatic ring, as is consistent with the results shown in Figures A5 and A15.

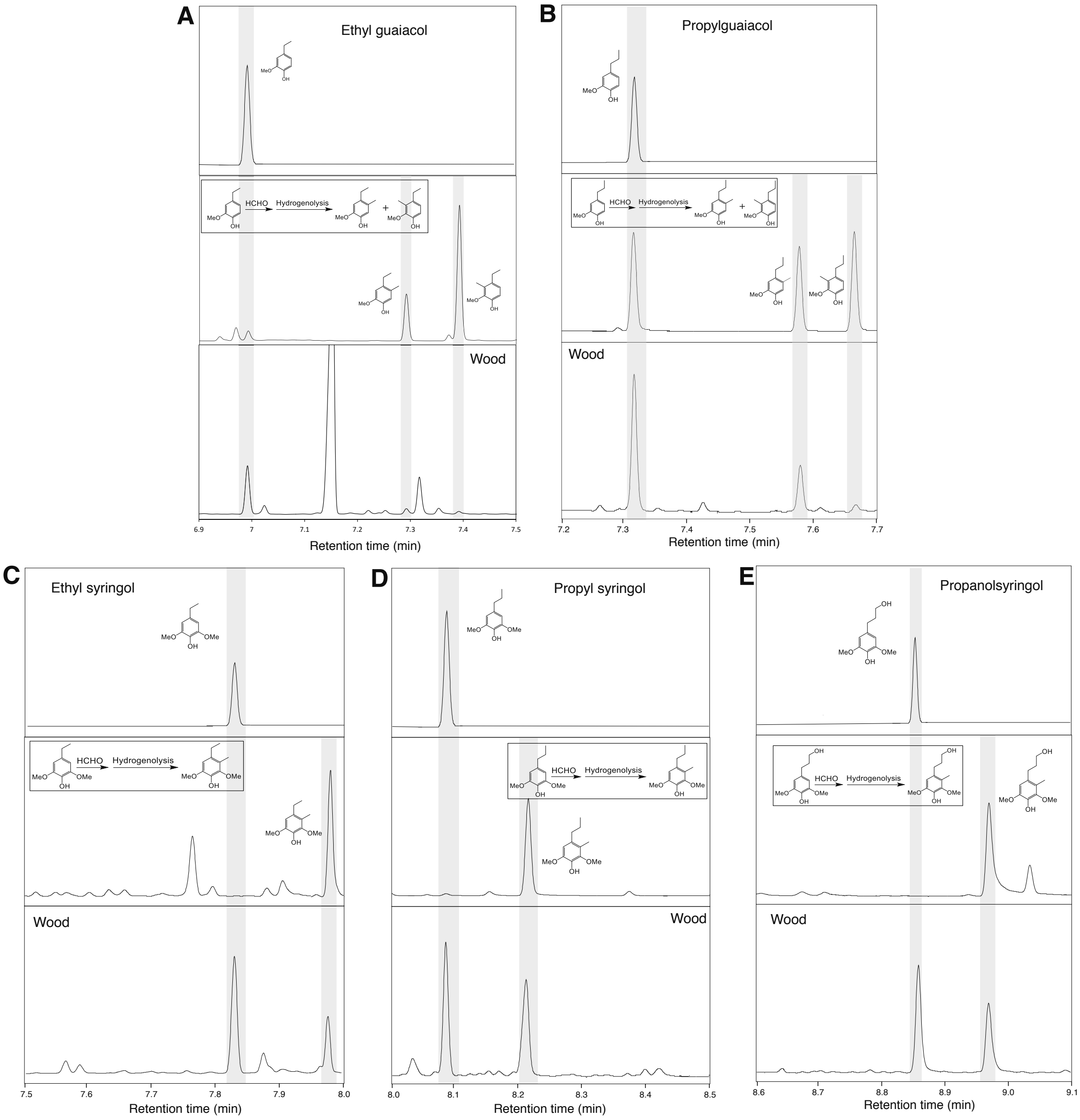


Figure A2. Chromatographic comparison of the products resulting from the reaction of formaldehyde (FA) with authentic standards (GC chromatograms of all authentic standards and associated lignin monomers are shown in Figures A12 and A13) and lignin monomers derived from wood (mixture from entry 3 in Table A1). This comparison confirms the appearance of methylated monomers in the presence of FA. (A) Ethylguaiacol (EG), (B) propylguaiacol (PG), (C) ethylsyringol (ES), (D) propylsyringol (PS), and (E) propanolsyringol (PSA). Detailed preparation procedure of methylated lignin monomers is described in Section A1.2.

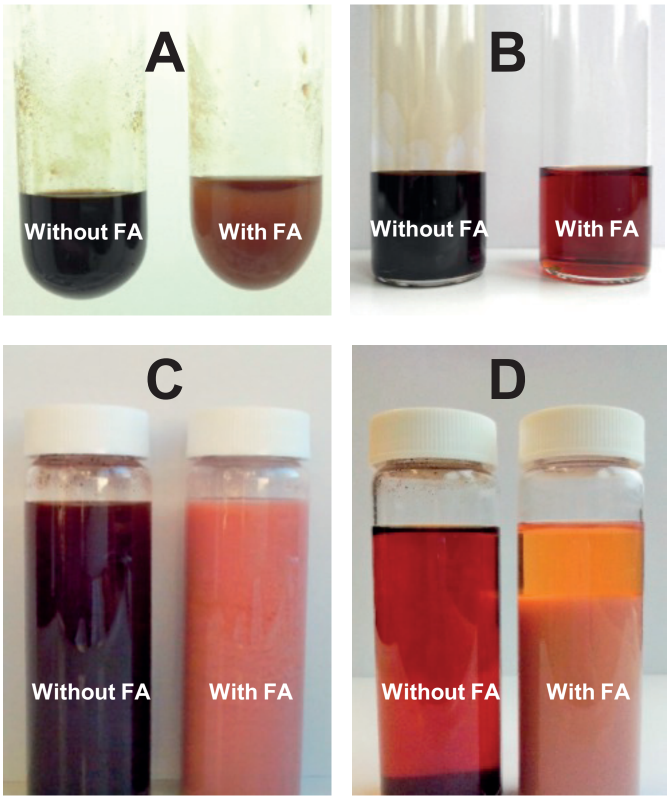

A
B
C
D
Without FA
With FA
Without FA
With FA
Without FA
With FA
Without FA
With FA

Figure A3. Extraction of lignin from beech wood. (A) Slurries after reaction (or pretreatment), (B) filtrates of the slurries, (C) 0 h after the addition of water into the extraction liquors (or filtrates) for lignin precipitation, and (D) 15 h after the addition of water into the extraction liquors (or filtrates) for lignin precipitation. (The dark color suggests a severe condensation of lignin; the light color suggests limited condensation of lignin.)

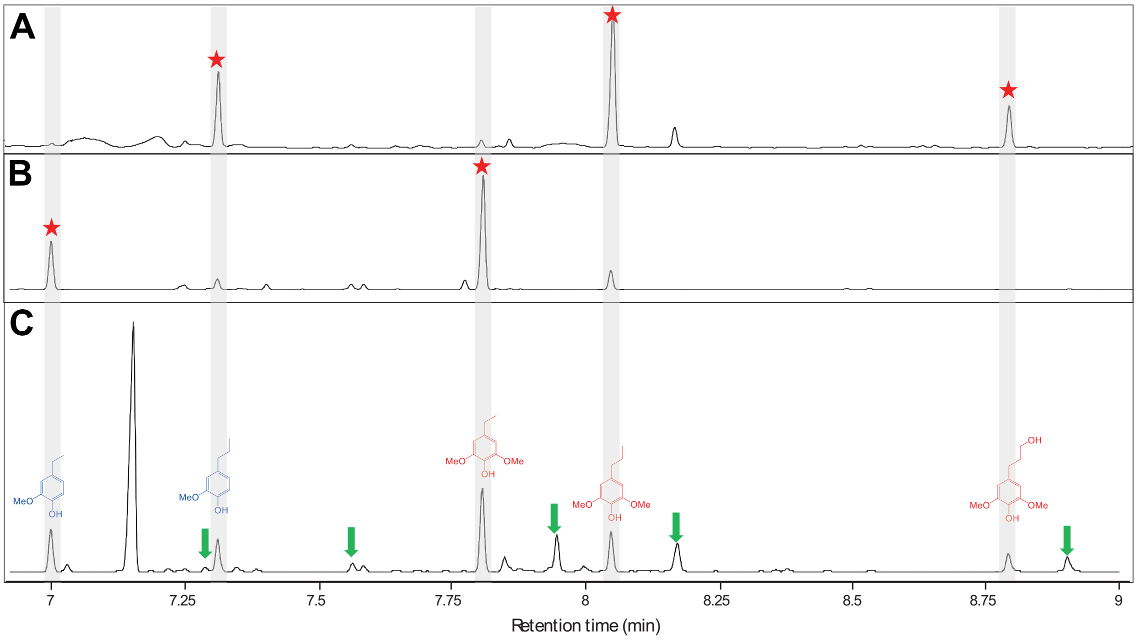


Figure A4. Chromatographic comparison of lignin monomers from hydrogenolysis of native lignin (A and B) and extracted lignin, with added FA (C). (A) Entry 1 in Table A1, (B) entry 12 in Table A1, and (C) entry 3 in Table A1. Direct hydrogenolysis of native lignin (A and B) generated five major lignin monomers (highlighted with red stars); with the addition of FA, hydrogenolysis of extracted lignin generated five additional methylated lignin monomers, the peaks of which are highlighted with green arrows. The corresponding structures are shown in Figure A11.

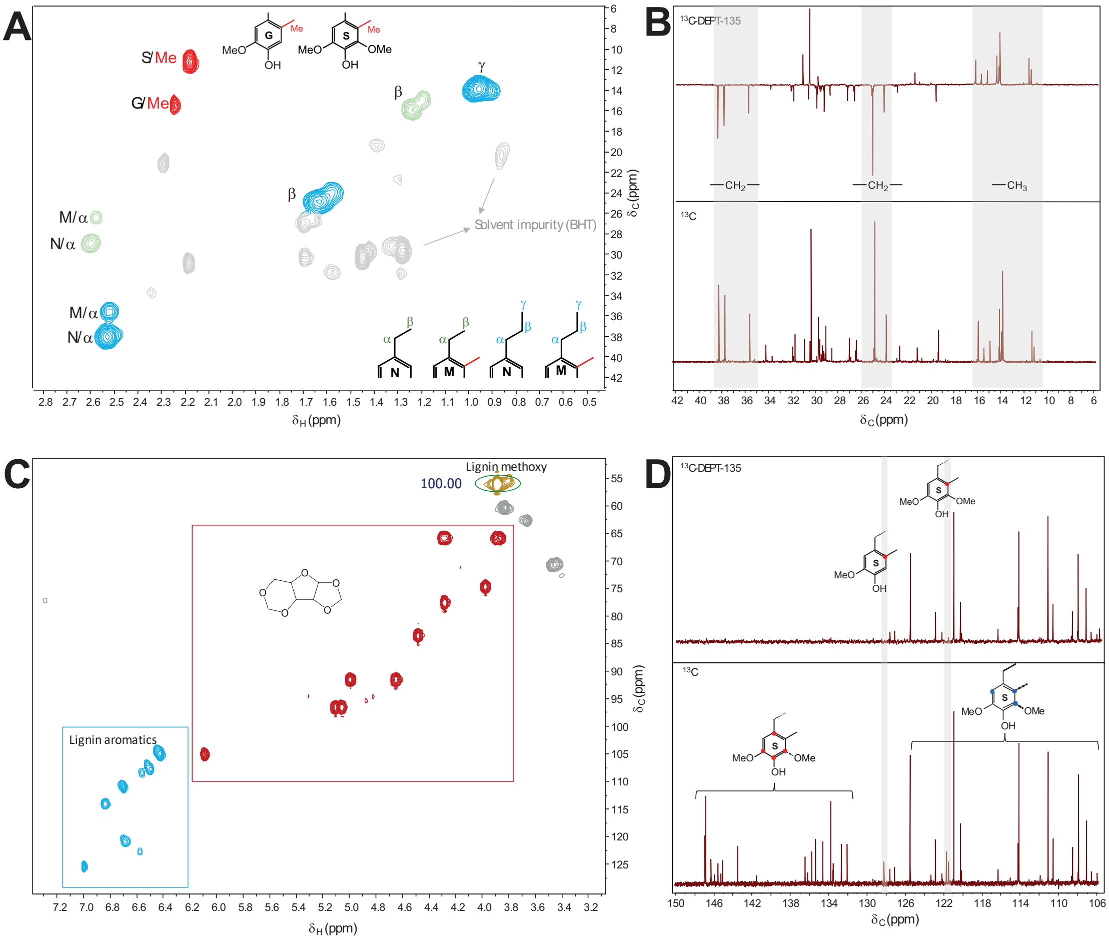


Figure A5. NMR characterization of the lignin monomer mixture (Entry 5 in Table A1). (A) 2D HSQC NMR spectra of the region corresponding to the side-chain signals of lignin monomers in chloroform-d ("N" represents unmethylated lignin monomers, "M" represents methylated lignin monomers, "S" represents syringyl monomers, "G" represents guaiacyl monomers, (B) ^{13}C and ^{13}C-Distortionless Enhancement by Polarization Transfer (DEPT-135) NMR spectra of the region corresponding to the side-chain signals of lignin monomers, (C) 2D HSQC NMR spectra of the region corresponding to the aromatic and methoxy signals of lignin monomers, and (D) ^{13}C and ^{13}C-DEPT-135 NMR spectra of the region corresponding to the aromatic and methoxy signals of lignin monomers in chloroform-d. The appearance of methyl groups in Figure A4A (red) and A4B and the signal from the tertiary carbon atoms (121 and 128 ppm) in Figure A4D are all indicative of the reaction of formaldehyde with lignin aromatic rings. In addition, the chemical shifts of methyl groups were consistent with those observed in pure methylated guaiacyl and syringyl monomers, shown in Figures A13, A14 and A15. Solvent impurity (BHT): butylated hydroxytoluene (the antioxidant in THF or 1,4-dioxane).

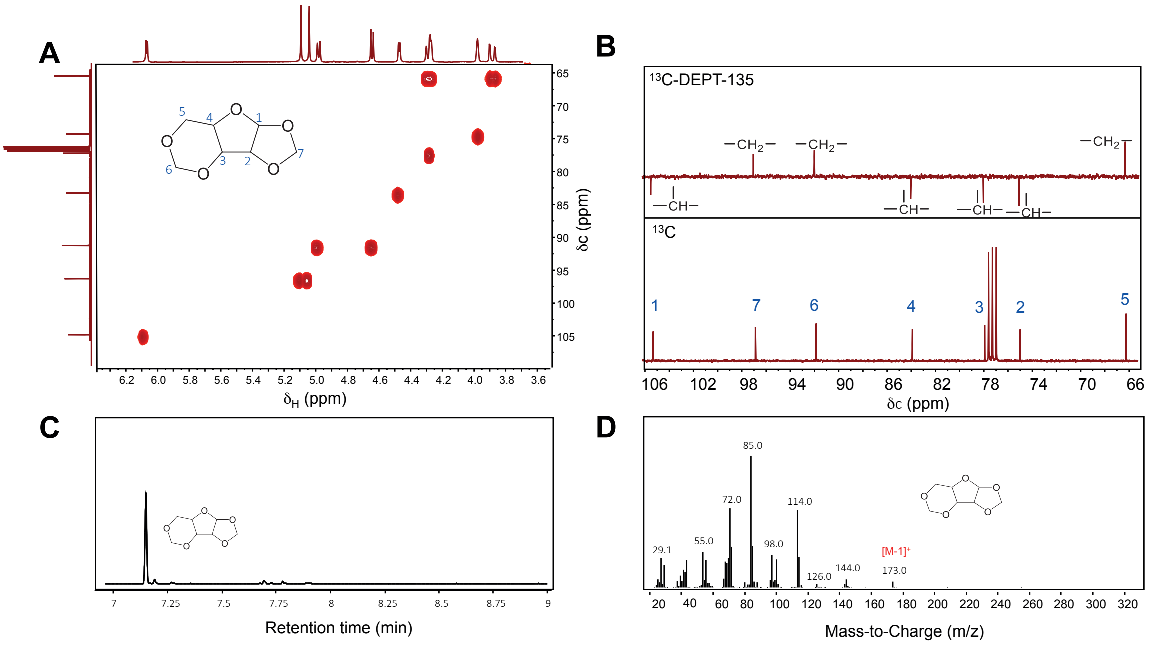


Figure A6. Identification of diformyl-xylose. (A) 2D HSQC NMR spectra of purified diformyl-xylose in chloroform-d, (B) ^{13}C and ^{13}C-DEPT-135 NMR spectra of diformyl-xylose in chloroform-d, (C) GC chromatogram of diformyl-xylose, and (D) mass spectra of diformyl-xylose. The 2D HSQC NMR peaks shown here (A) are consistent with those assigned in Figure 19A.

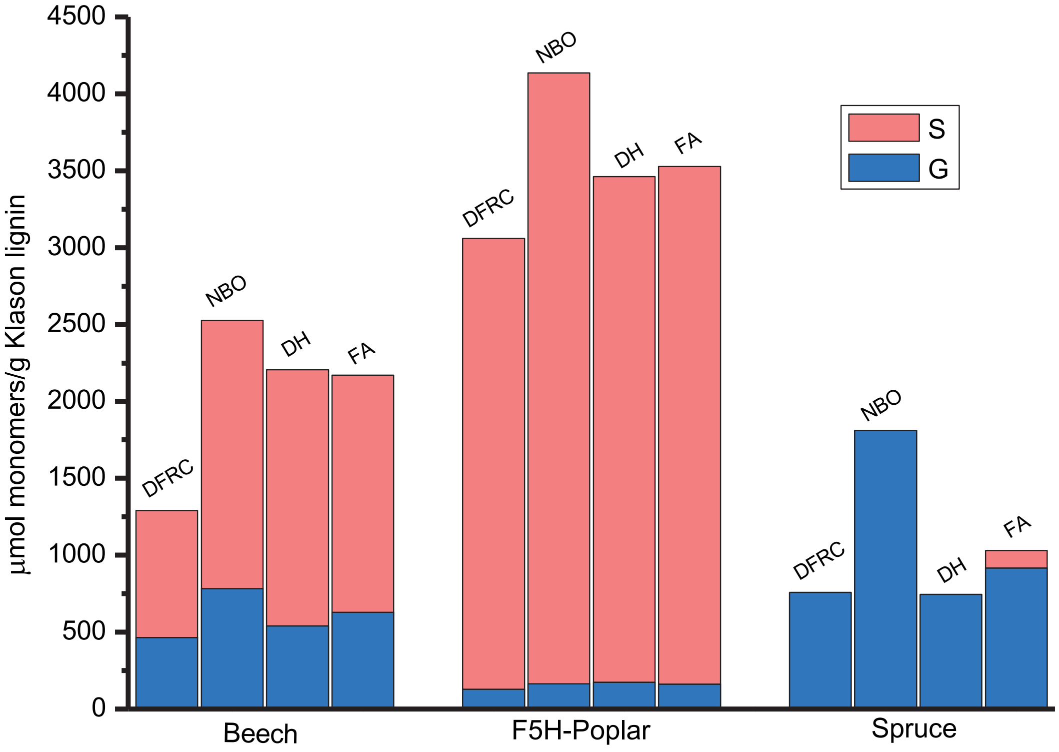


Figure A7. Lignin monomer analysis. "S" refers to syringyl units; "G" refers to guaiacyl units; "DFRC" refers to the derivatization followed by reductive cleavage method; "NBO" refers to the nitrobenzene oxidation method; "DH" refers to direct hydrogenolysis method; "FA" refers to FA extraction followed by hydrogenolysis (i.e. the method reported in this appendix);

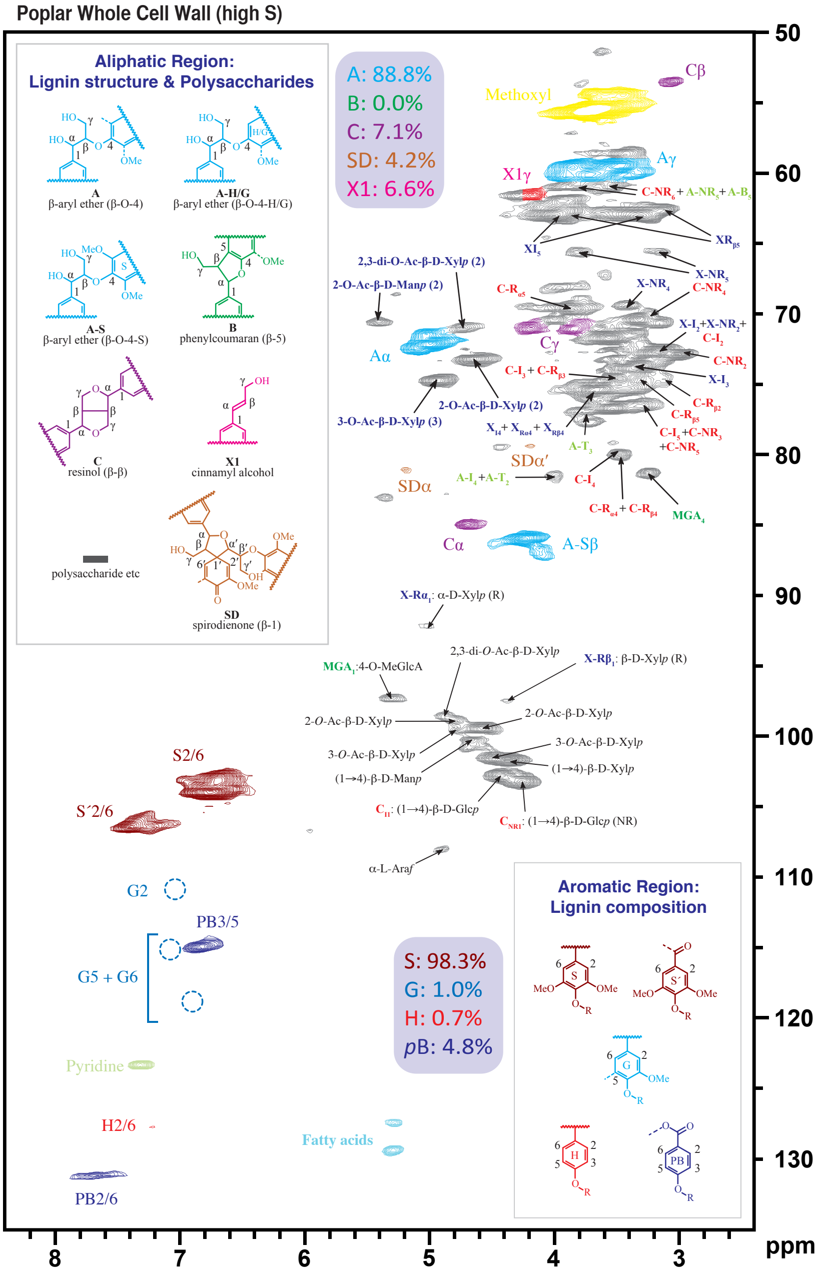


Figure A8. Partial 2D HSQC NMR spectrum of the high-syringyl (*F5H*-upregulated) transgenic poplar whole cell wall material in DMSO-d_6/pyridine-d_5 (4:1). The data shown include both aromatic (lignin aromatic units) and aliphatic regions (lignin structures and residual polysaccharides). Contour coloration matches that of the structures shown. The quantification values are from volume integration and are indicative of the relative levels of lignin aromatics or of the various aliphatic structures. This transgenic poplar has unusually high syringyl content (~98.3% S) compared to wild-type poplar (68%).[192]

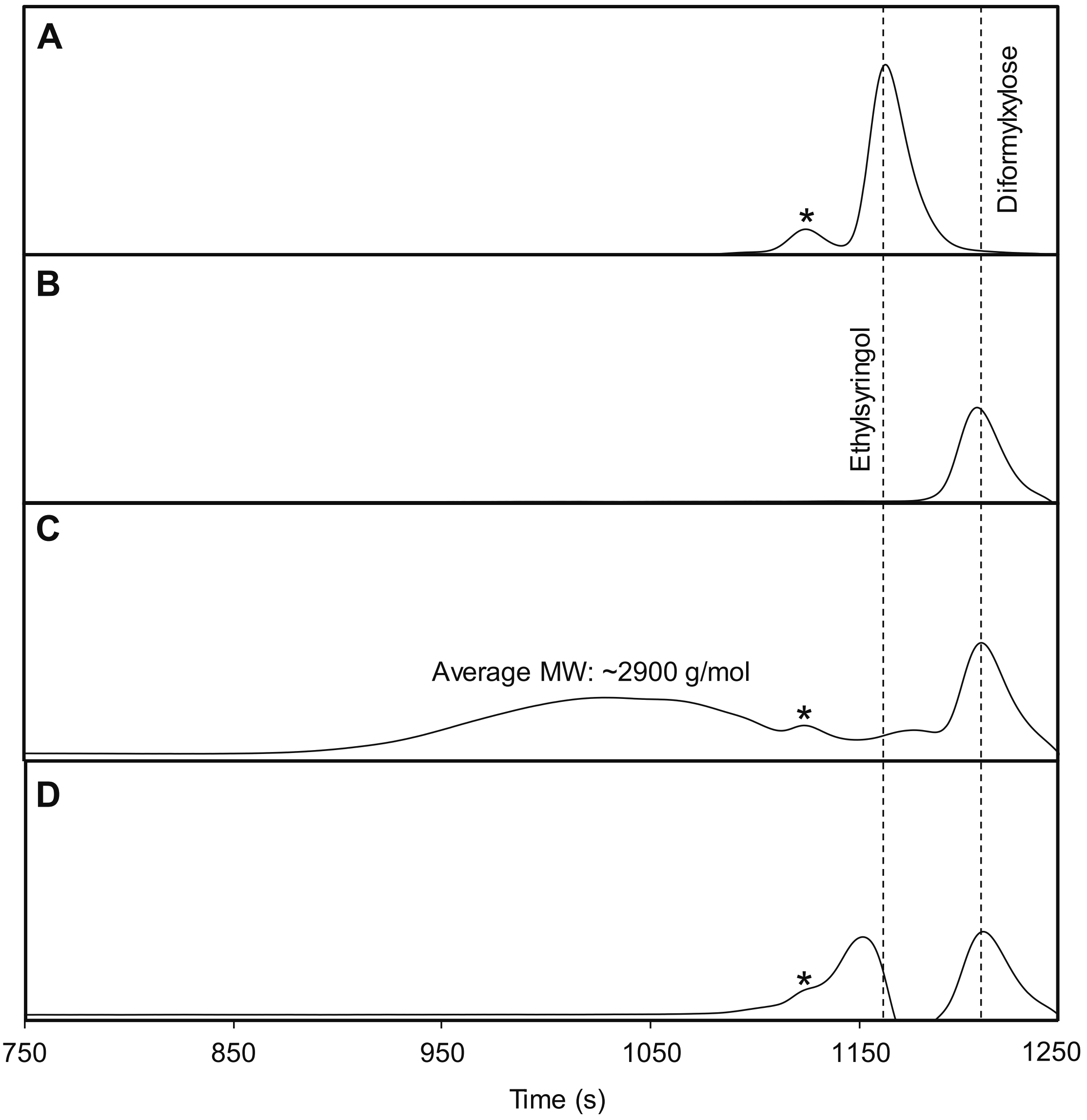


Figure A9. Gel permeation chromatograms of FA-extracted lignin before and after hydrogenolysis. (a) Ethyl syringol (182 g/mol), (b) diformylxylose (162 g/mol), (c) lignin liquor before hydrogenolysis (~2900 g/mol) and (d) lignin liquor after hydrogenolysis (152–212 g/mol).

“*”: the peak refers to the stabilizer in THF (butylated hydroxytoluene, 220 g/mol)). Unstabilized THF was purposefully used to solubilize diformylxylose leading to the absence of this peak in panel B.

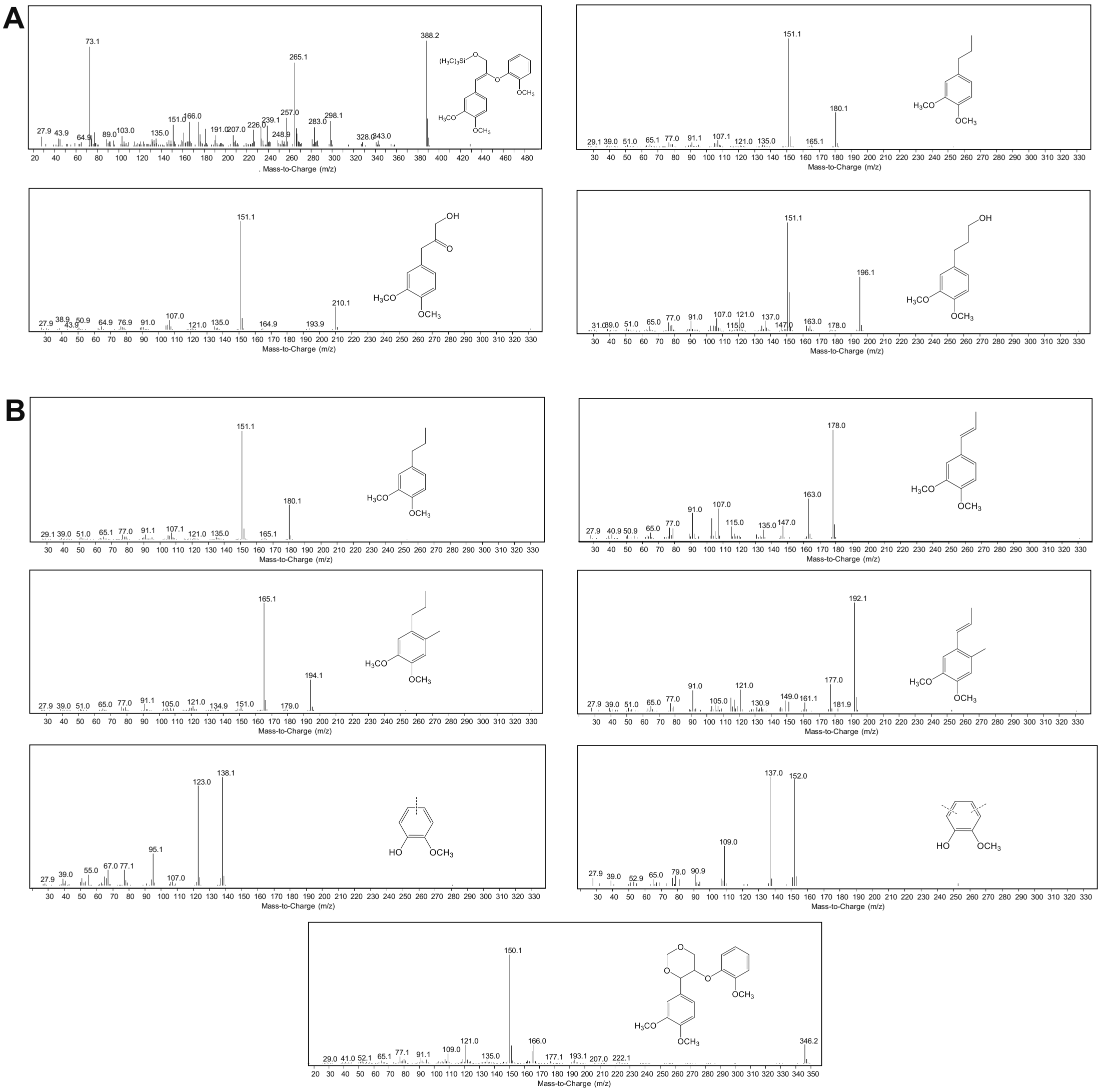


Figure A10. GC-MS spectra of products resulting from the hydrolysis of dimer model compound (VG) followed by hydrogenolysis. (A) Reaction without formaldehyde (FA) addition, and (B) reaction with FA addition.

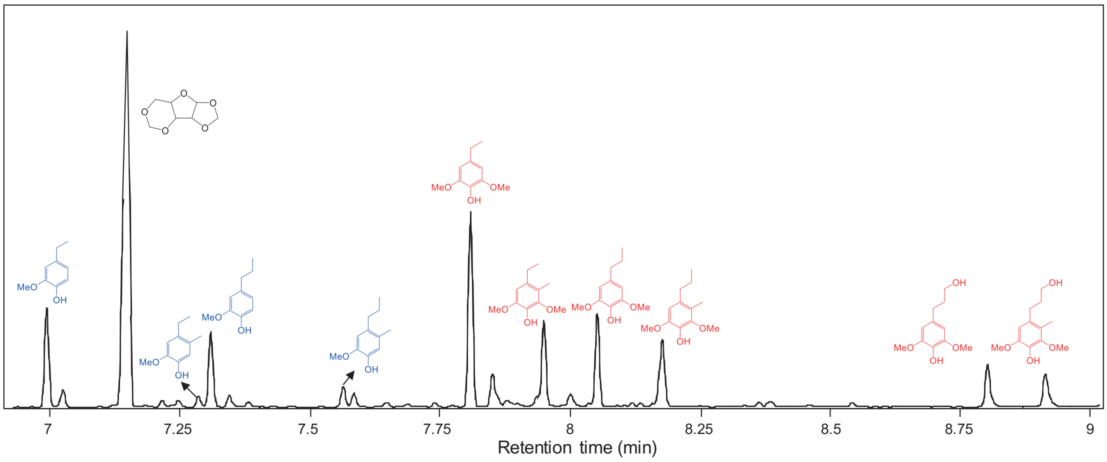


Figure A11. GC-FID chromatogram of the lignin monomer mixture showing the identified monomers.

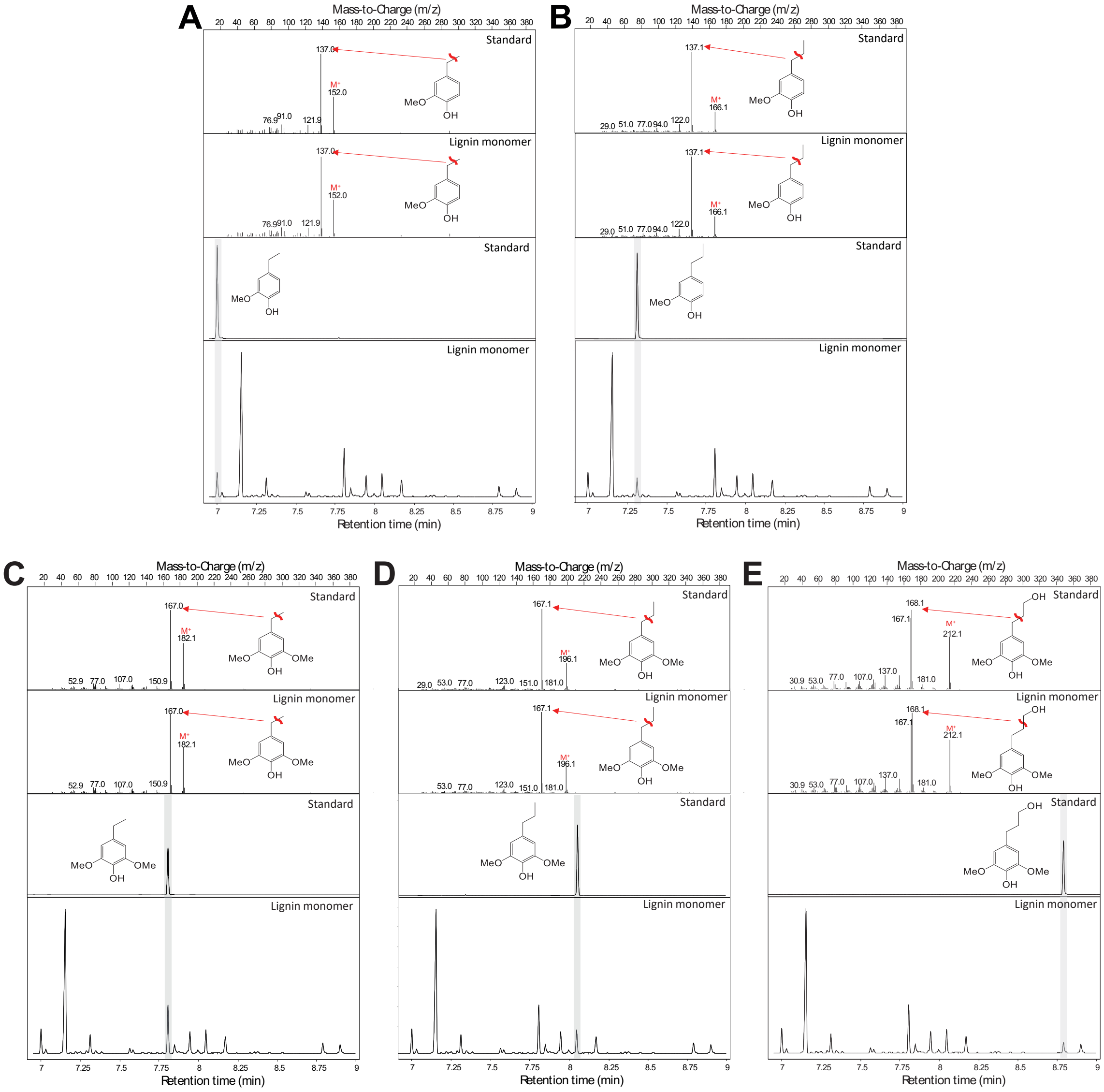


Figure A12. Comparison of standard monomers with identified lignin monomers. (A) Ethylguaiacol, (B) propylguaiacol, (C) ethylsyringol, (D) propylsyringol, and (E) propanolsyringol.

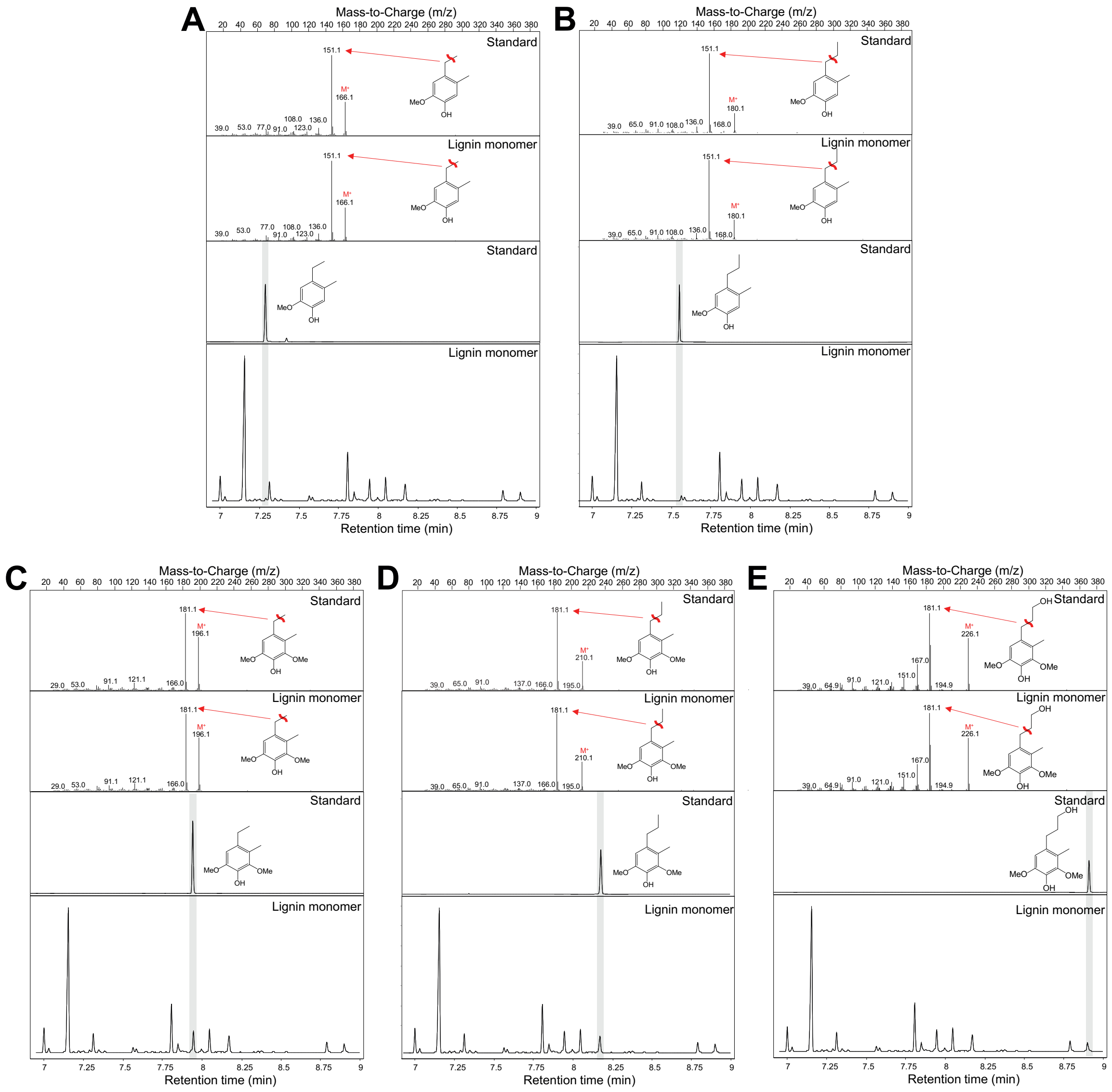


Figure A13. Comparison of prepared standard methylated monomers with identified lignin monomers. (A) Methylated ethylguaiacol, (B) methylated propylguaiacol, (C) methylated ethylsyringol, (D) methylated propylsyringol, and (E) methylated propanolsyringol.

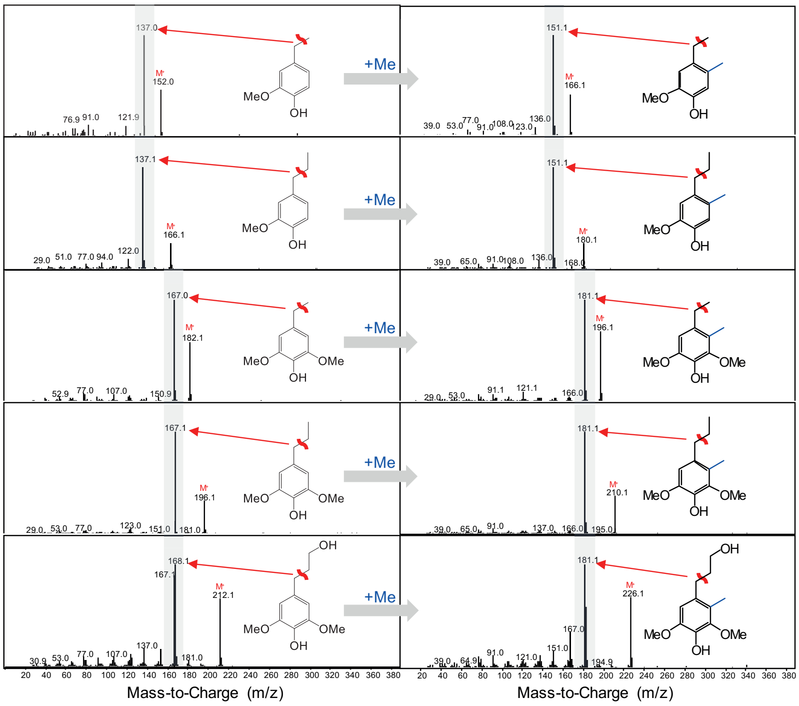


Figure A14. Comparison of the GC-MS spectra of the 5 unmethylated (left) and 5 methylated (right) monomers detected in the biomass-derived mixtures. M^+ is the molecular ion peak, representing the molecular weight of lignin monomers; the highest peak identified with a red arrow represents the molecular weight of the most stable ion fragment from each lignin monomer.

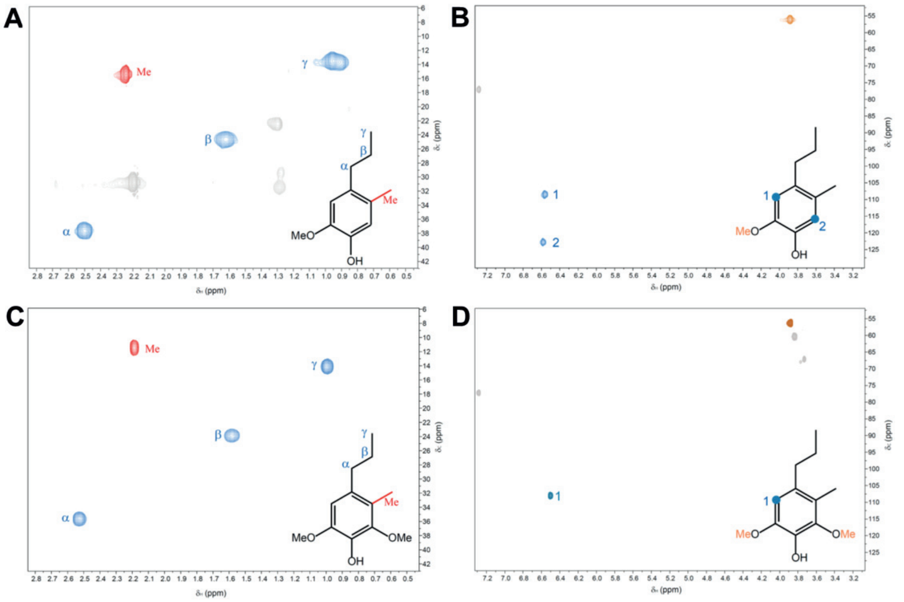


Figure A15. 2D HSQC NMR of methylated propylguaiacol and methylated propylsyringol in chloroform-d. The appearance of a methyl group on the aromatic ring was consistent with Figure A1C and Figure A4. In addition, comparison of the proton signals from the aromatic rings on spectra (B) and (D) confirmed the substitution of the proton at the position *para* to the methoxyl group

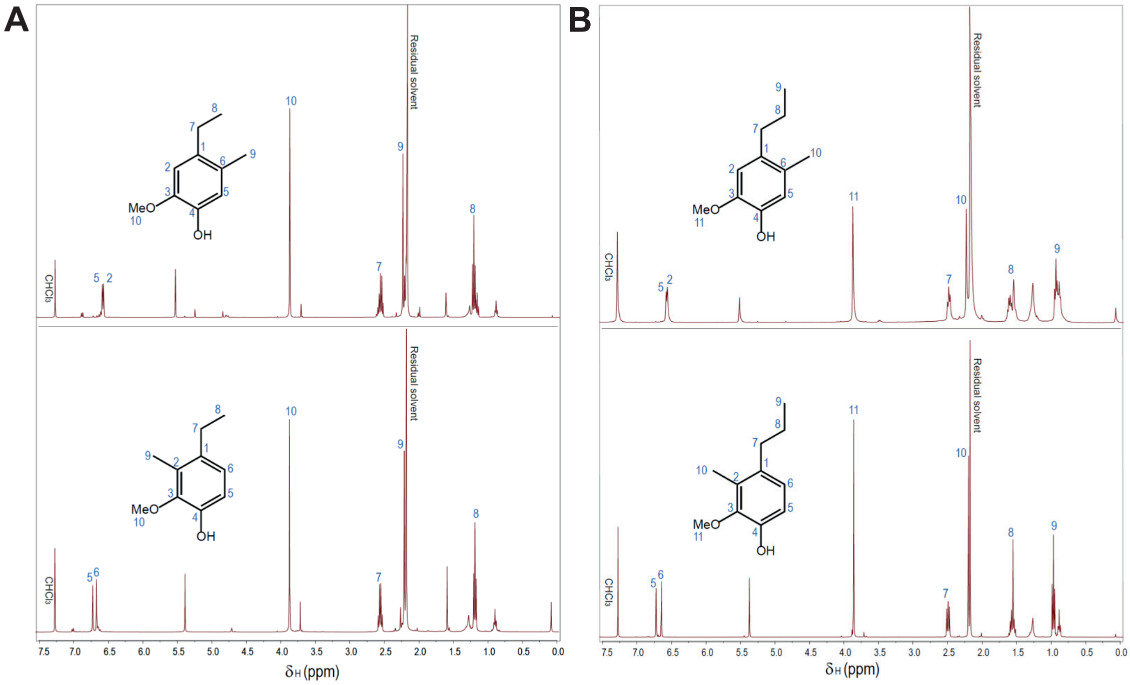


Figure A16. H^1 NMR of (A) Methylated ethylguaiacol isomers, and (B) methylated propylguaiacol isomers in chloroform-d. The NMR spectra confirmed that the methyl group detected in the biomass-derived solution (Figure A4) was substituted onto the position *para* to the methoxyl group.

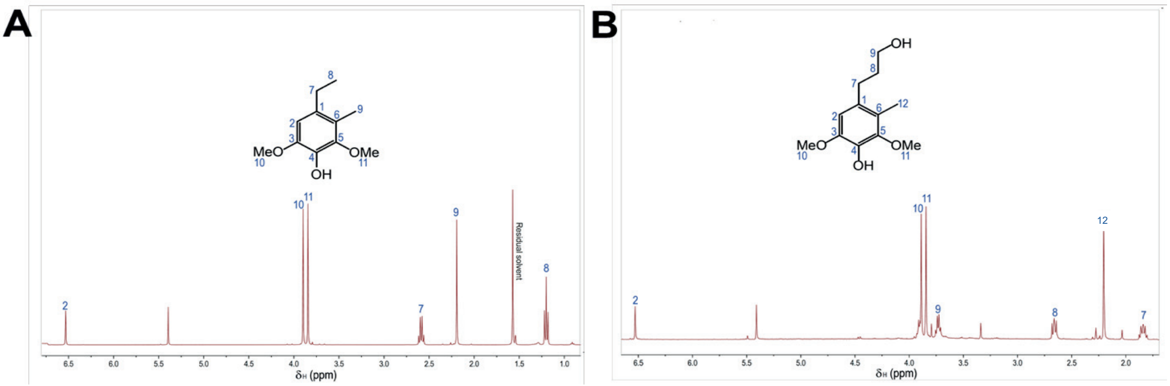


Figure A17. H^1 NMR of (A) Methylated ethylsyringol and (B) methylated propanolsyringol in chloroform-d.

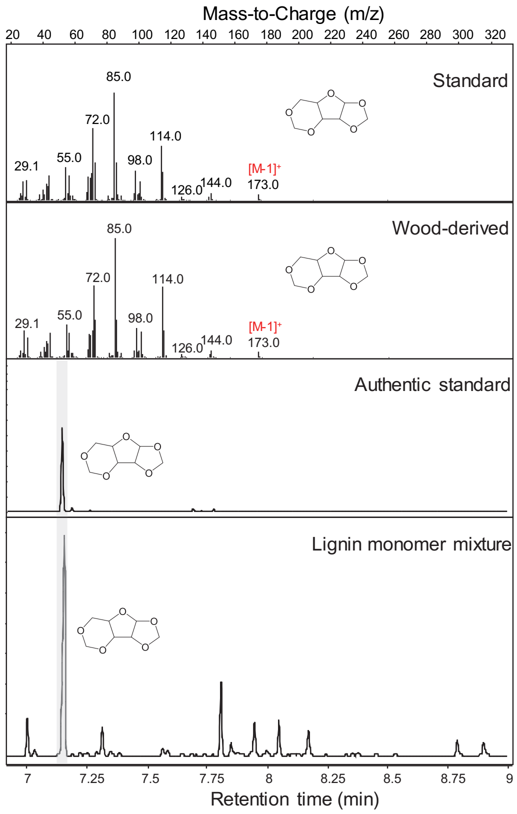


Figure A18. Comparison of prepared standard diformyl-xylose with wood-derived diformyl-xylose.

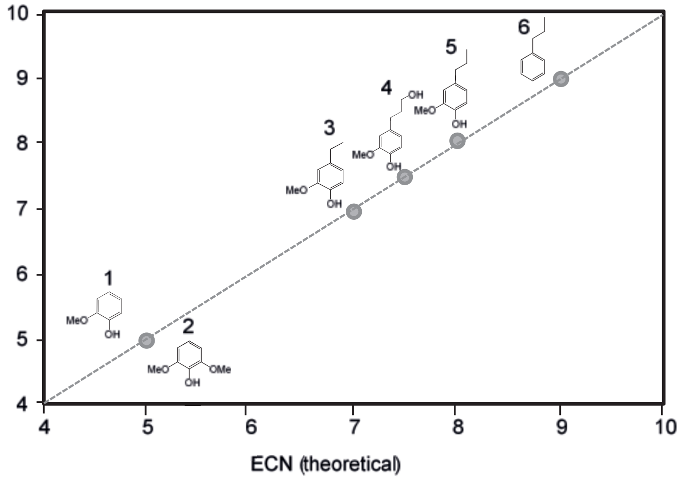


Figure A19. Validation of Effective Carbon Number (ECN) rule for lignin monomer quantification (X axis: ECN based on empirical rule (Table A4); Y axis: ECN calculated based on FID response using decane as an internal standard). Compound **1** (guaiacol), compound **2** (syringol), compound **3** (ethylguaiacol), compound **4** (propanolguaiacol), compound **5** (propylguaiacol) and compound **6** (propylbenzene).

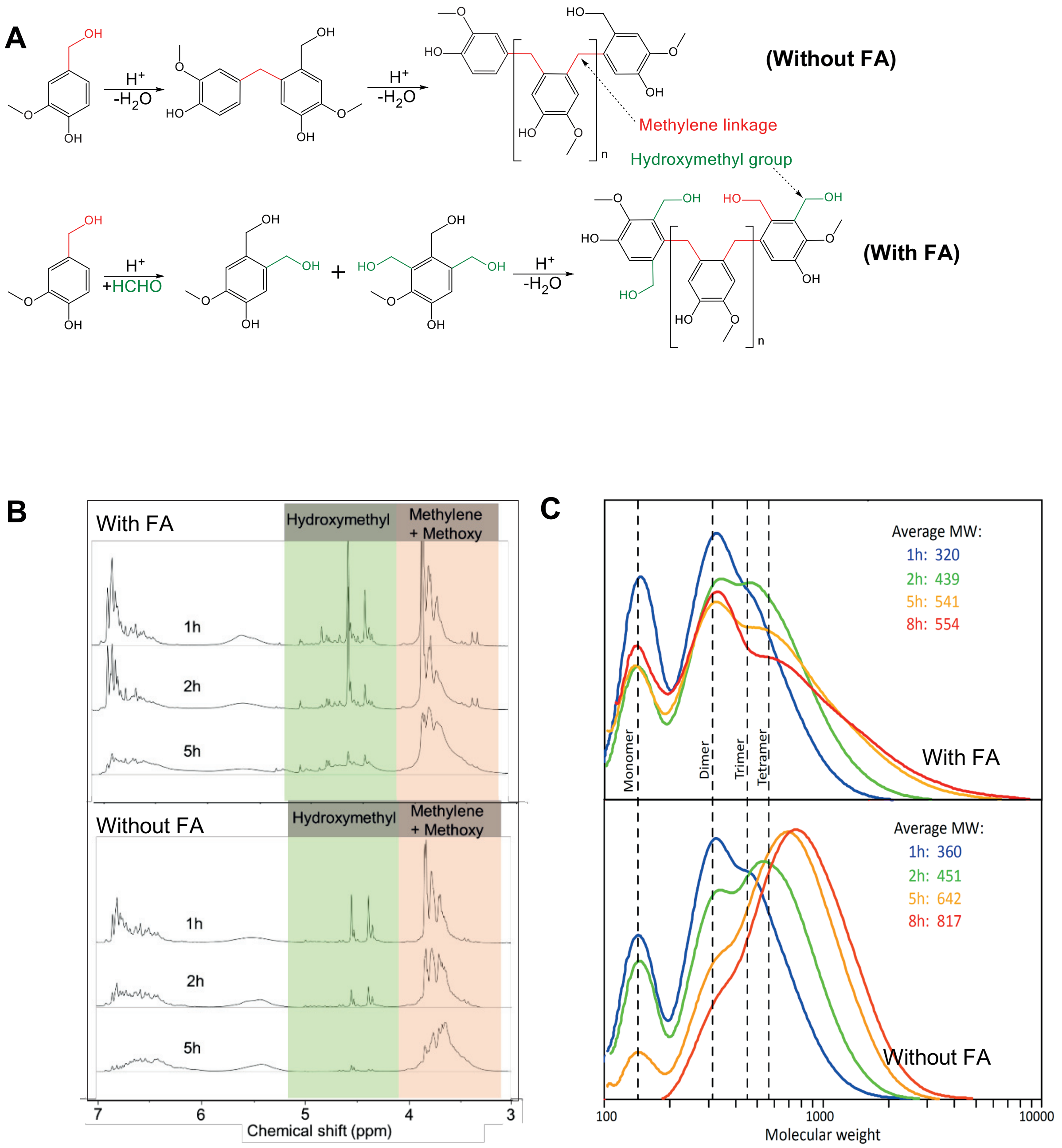


Figure A20. Acid-catalyzed condensation of vanillyl alcohol in a 9:1 1,4-dioxane/water mixture (with or without formaldehyde). (A) Proposed mechanism, (B) NMR spectra of resultant product mixtures in chloroform-d at different reaction times, and (C) GPC chromatograms of resultant product mixtures at different reaction times.

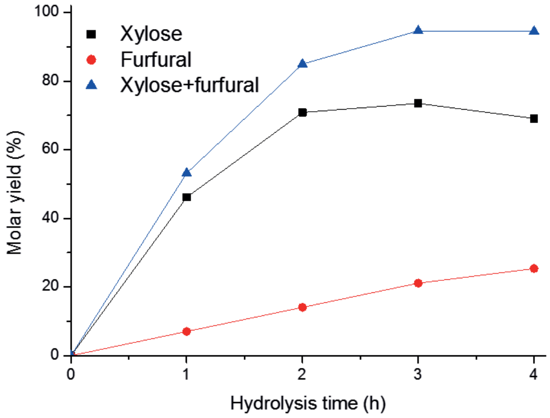


Figure A21. Hydrolysis of diformyl-xylose to xylose.

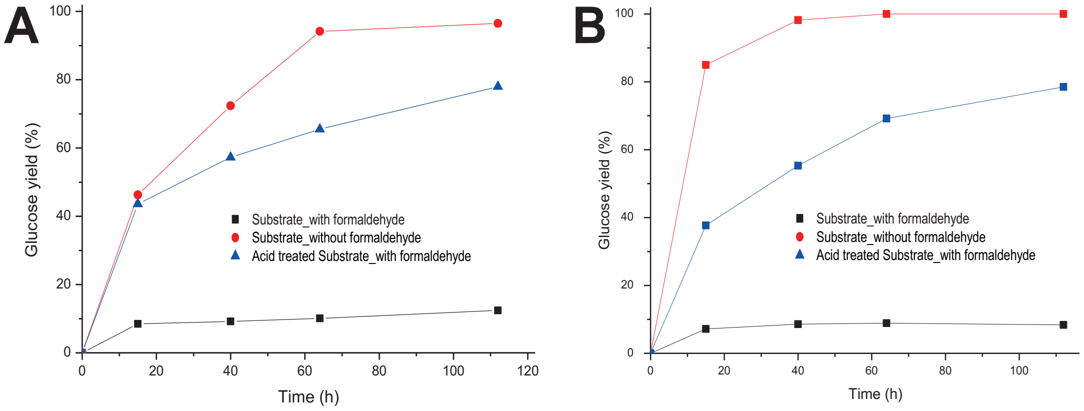


Figure A22. Enzymatic digestibility of substrates after lignin extraction. (A) Beech wood (additional acid treatment: 120 ºC, 2 h, 1 wt% H_2SO_4 solution, liquid/solid: 10:1), and (B) genetically modified wood (additional acid treatment: 140 ºC, 1 h, 1 wt% H_2SO_4 solution, liquid/solid: 10:1). (Lignin extraction conditions refer to entries 5 and 18 in Table A1; enzymatic hydrolysis conditions: 50 ºC; 250 rpm; and enzyme (Novozymes *Cellic*® ctec2) loading of 30 FPU/g glucan).

A.2.6 Tables

Table A1. Yields of lignin monomers for different reaction conditions and substrates.

Entry[a]	Feedstock	Extraction[b]	Hydrogenolysis [c]	Extracted lignin yield[d] (%)	M 1 (%)	M 2 (%)	M 3 (%)	M 4 (%)	M 5 (%)	M 6 (%)	M 7 (%)	M 8 (%)	M 9 (%)	M 10 (%)	Monomer yield based on native lignin (%)
1	Beech wood	/	250 °C, 15h[e]	/	0.53	0.00	10.06	0.00	1.30	0.00	24.15	0.00	12.22	0.00	48.26
2		80 °C, 5h [f]	200 °C, 6h	76.85	0.55	0.00	0.66	0.00	2.07	0.00	1.82	0.00	2.00	0.00	7.10
3		80 °C, 5h	200 °C, 6h	78.96	0.64	0.00	7.01	2.27	2.05	1.07	13.31	9.44	5.67	3.60	45.06
4		80 °C, 1.5h	250 °C, 15h	47.39	2.53	0.25	5.79	1.04	7.78	1.96	12.05	2.46	0.00	0.27	34.13
5		100 °C, 2h	250 °C, 15h[g]	71.59	5.31	0.60	5.01	1.39	15.52	5.78	9.49	3.83	0.00	0.26	47.19
6		80 °C, 5h	250 °C, 15h	78.96	0.30	0.65	6.26	2.28	9.71	5.65	12.13	7.02	0.00	0.40	44.39
7		80 °C, 5h	230 °C, 15h[h]	78.96	1.02	0.00	3.99	1.27	2.29	1.54	7.86	7.60	8.09	4.70	38.35
8		100 °C, 2h	250 °C, 15h	71.59	2.62	0.46	7.10	1.93	8.90	4.02	14.31	6.03	0.00	0.41	45.79
9		120 °C, 1h	250 °C, 15h	75.80	2.18	0.31	4.22	1.64	6.32	5.01	8.44	6.62	0.80	1.11	36.66
10		80 °C, 5h[i]	250 °C, 15h	76.85	4.26	0.00	2.87	0.83	15.08	0.15	7.06	0.74	0.26	0.32	31.56
11		100 °C, 3h	250 °C, 15h	77.90	2.62	0.57	4.14	1.99	7.35	6.26	7.87	6.74	0.47	0.76	38.76
12		/	250 °C, 15h[j]	/	8.16	0.00	1.86	0.00	22.64	0.00	3.37	0.00	0.00	0.00	36.04
13		100 °C, 2h[k]	200 °C, 15h	63.17	0.32	0.00	6.84	2.84	0.53	0.53	14.31	9.37	5.37	3.37	43.46
14	F5H-Poplar	/	250 °C, 15h[e]	/	0.00	0.00	3.41	0.00	2.53	0.00	49.57	0.00	22.19	0.00	77.70
15		80 °C, 5h[f]	250 °C, 15h	52.95	1.33	0.00	0.00	0.00	15.05	0.00	5.55	0.00	2.60	0.00	24.53
16		80 °C, 5h	250 °C, 15h	73.29	1.24	0.00	2.01	0.66	15.43	5.57	18.81	9.08	16.66	9.19	78.64
17		100 °C, 2h[f]	250 °C, 15h	54.22	0.00	0.00	0.00	0.00	16.21	0.00	5.65	0.00	4.56	0.00	26.42
18		100 °C, 2h	250 °C, 15h	74.57	0.81	0.00	3.28	0.74	2.33	1.75	25.44	13.51	13.65	9.77	71.29
19		100 °C, 2h	250 °C, 6h	74.57	0.96	0.00	1.64	0.57	7.26	3.54	17.84	22.82	12.73	11.90	79.25
21		/	150 °C, 20h[e]	/	0.00	0.00	0.00	0.00	0.17	0.00	0.04	0.00	0.00	0.00	0.21
22		/	150 °C, 20h[l]	/	0.00	0.00	0.90	0.00	0.73	0.00	43.24	0.00	2.17	0.92	47.95
23		80 °C, 5h	150 °C, 20h[m]	73.29	0.00	0.00	1.38	0.65	0.00	0.00	22.13	23.73	0.00	5.15	60.42[n]
24	Spruce	/	250 °C, 15h[e]	/	2.03	0.00	12.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	21.05[o]
25		100 °C, 2h	250 °C, 15h	Not measured	6.85	1.61	7.69	1.85	0.00	0.00	0.00	0.00	0.00	0.00	20.56[o]

Notes for Table S1: (a) Entries 1-5 are shown in Figure 19B and entries 13-16 are shown in Figure 19C; (b) other extraction procedures are detailed in Section A.1.2; (c) other hydrogenolysis procedures are detailed in Section A.1.2; solvent was tetrahydrofuran (THF) unless specified, H_2 pressure was 40 bar; (d) extracted lignin yield was calculated based on the amount of native lignin; (e) direct hydrogenolysis of native lignin in biomass (methanol as solvent, 1 g wood particles, 200 mg Ru/C); (f) without the addition of formaldehyde (FA); (g) 1,4-dioxane instead of THF as hydrogenolysis solvent; (h) methanol instead of THF as solvent; (i) 1/10 FA loading. M1-M10 correspond to the monomers in the order from left to right in Figure A11; (j) direct hydrogenolysis of native lignin in biomass (THF as solvent, 1 g wood particles, 200 mg Ru/C); (k) γ-valerolactone (GVL) instead of 1,4-dioxane as extraction solvent; (l) THF was used as the hydrogenolysis solvent with 200 µl of 36% HCl added to the mixture prior to hydrogenolysis.; (m) 200 µl of 36% HCl were added to the pretreatment liquor prior to hydrogenolysis. (n) 7.4% of the final yield was contributed by a newly identified monomer (methylated propylene syringol, 4-(2-hydroxyethyl)-2,6-dimethoxyphenol), which was only identified by GC-MS and listed as M11. (o) 6.45% (entry 24) and 2.56% (entry 25) of the final yield was provided by a newly identified monomer (propanol guaiacol, (Z)-4-(2-hydroxyvinyl)-2,6-dimethoxy-3-methylphenol), which was only identified by GC-MS and listed as M12. M1: ethylguaiacol, M2: methylated ethylguaiacol, M3: propylguaiacol, M4: methylated propylguaiacol, M5: ethylsyringol, M6: methylated ethylsyringol, M7: propylsyringol, M8: methylated propylsyringol, M9: propanolsyringol, and M10: methylated propanolsyringol. The reported yields are given on a molar basis and are based on Klason lignin.

Table A2. High-resolution ESI-MS of purified methylated lignin monomers. The monomers detected in the wood-derived mixtures are shown in bold.

Lignin monomers	Molecular Formula	$[M+H]^+$ (Theoretical)	$[M+H]^+$ (High resolution ESI-MS, positive mode)
Methylated ethylguaiacol (M2)[a]	$C_{10}H_{14}O_2$	167.1072	167.1063
Methylated propylguaiacol (M4) [a]	$C_{11}H_{16}O_2$	181.1228	181.1215
Methylated ethylsyringol (M6) [a]	$C_{11}H_{16}O_3$	197.1178	197.1168
Methylated propylsyringol (M8) [a]	$C_{12}H_{18}O_3$	211.1334	210.1325
Methylated propanolsyringol (M10) [a]	$C_{11}H_{16}O_4$	213.1129	213.1127
Methylated ethylguaiacol isomer [b]	$C_{10}H_{14}O_2$	167.1072	167.1047
Methylated propylguaiacol isomer[b]	$C_{11}H_{16}O_2$	181.1228	181.1211
Compound 7[c]	$C_{11}H_{16}O_2$	181.1228	181.1211
Compound 7[c]	$C_{11}H_{16}O_3$	197.1178	197.1177
Compound 3[b]	$C_{11}H_{14}O_4$	233.0790	233.0764

Notes: [a] monomers are shown in Figure A11, [b] isomers are shown in Figure A14, and [c] monomers are shown in Figure 18A

Table A3. Compositional analysis of biomass

Substrate	Moisture (wt %) [a]	Extractives (wt %) [a]	Weight fraction (wt %)[b]						
			Glu	Xyl	Ara	Gal	Man	KL.	ASL
Beech	5.7	2.6	37.1	15.3	2.1	2.2	1.9	22.8	3.7
F5H-poplar	6.7	4.7	41.4	17.4	ND	ND	ND	14.2	7.5
Spruce[c]	6.6	5.4	/	/	/	/	/	27.8	0.7

Notes: [a] On the basis of wet biomass, [b] on the basis of dried extracted biomass. “glu” refers to glucose; “xyl” refers to xylose; “ara” refers to arabinose; “gal” refers to galactose; “man” refers to mannose; “KL” refers to Klason lignin; and “ASL” refers to acid soluble lignin, [c] Structural polysaccharides were not determined for spruce.

Table A4. Increments to calculate the ECN

Atom/group	ECN contribution
Carbon-aliphatic	1
Carbon-aromatic	1
Oxygen-primary alcohol	-0.6
Oxygen-phenol	-1
Oxygen-ether	-1

Table A5. Effective carbon number (ECN) for lignin monomers used in this study

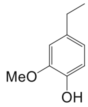
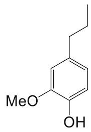
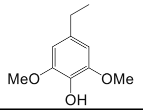
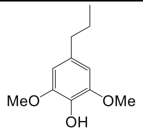
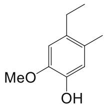
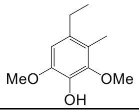
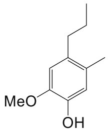
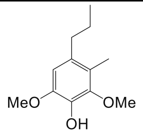
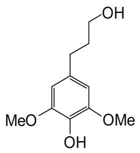
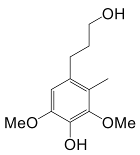
Lignin monomer structure	Effective carbon number calculated based on adjusted ECN rule (ECN$_{\mathrm{monomer}}$)
	7
	8
	7
	8
	8
	8
	9
	9
	7.4
	8.4

Table A6. Comparison of lignin monomer quantification with NMR integration and GC-FID.

Hydrogen type[a]	NMR integration	GC-FID	Error (%)
Methoxy	100[b]	100[c]	/
N/α	19.77	19.21	2.88
M/α	7.64	7.95	3.96
N/α	8.77	9.08	4.43
M/α	3.87	3.93	1.64
S/Me	13.79	13.76	0.22
G/Me	4.15	4.07	1.85
β (ethyl)	18.76	19.52	3.91
β (propyl)	28.19	27.17	3.75
γ	44.15	43.25	2.08

Notes: [a] The NMR spectrum used for hydrogen integration is shown in Figure A4A, [b] The area of integration of methoxy protons was set as 100 (each proton accounts for 33.33% of the area), and [c] The molar amount of methoxy group was set as 33.33; therefore, the total molar amount of hydrogen in methoxy groups was 100.

Table A7. Mass balance of lignin extraction of beech wood particles.

Solvent	Solid residue	Extraction liquor
1,4-dioxane+FA	Glucose: 88% Xylose: ND	Glucose: 11% Xylose: ND Diformyl-xylose: 50% Furfural: 40%
GVL+FA	Glucose: 88% Xylose: 9%	Glucose: 14% Xylose: ND Diformyl-xylose: 82% Furfural: 10%

Table A8. FA mass balance during pretreatment.

Biomass	Residual FA (%)
Beech	55.74
Poplar	67.17

Table A9. FA consumption and recovery with polysaccharides.

Feedstock	Amount (mg)	Reaction condition	FA consumption after reaction (%)	FA recovery after hydrolysis of reaction mixtures (%)
Cellulose	150	80 ºC, 5h	18	99
Cellobiose	170	80 ºC, 1h	19	99
Glucose	170	80 ºC, 0.5 h	17	97
Xylan	63	80 ºC, 3h	13	99
Xylose	72	80 ºC, 0.5h	7	103

Appendix B Appendix for chapter 3

B.1 Experimental design

As discussed previously, here we detail two optimized protocols for isolating and purifying both the formaldehyde and propionaldehyde stabilized lignins in solid form. Beyond extracting isolated lignins that retain their full upgrading potential, the protocols also allow for the full fractionation of the lignocellulosic biomass producing highly digestible cellulose (in the case of the propionaldehyde fractionation) and stabilized xylose. Before embarking on these procedures, there are a few considerations and adjustments that should be made.

B.1.1 Prior to the fractionation

As the biomass source largely dictates the results obtained for the fractionation protocols, it is important to fully characterize the feedstock composition beforehand. With that in mind, we have also detailed methods to quantify the ash, hydration, extractives, structural sugar, and lignin of the biomass. These procedures are based on available protocols and are re-described here in Boxes 1-5 for convenience. Should certain characteristics of the extracted materials be desired, we recommend that the reader first optimize their choice of biomass using one or more of the available characterization protocols. Of particular concern is the quality of the lignin in the raw biomass (e.g. fraction of native interunit ether linkages vs. native interunit C—C linkages).[182,199] If the biomass has been processed in any way (e.g. heated, dried, etc.), the lignin that is extracted may already be cross-linked and will therefore provide low yields of monomers upon hydrogenolysis. We highly recommend performing a direct hydrogenolysis on a sample of the biomass that you intend to extract. If it gives you poor yields, then while the extraction will afford you a stabilized lignin, it will give similar poor yields when upgraded.

B.1.2 Fractionation

Here we present the procedure that is most optimal for hardwood biomass sources. However, depending on the biomass source, it can be modified for dramatically improved results. Variables that should be considered include acid loading, temperature, and time. We have found that the reaction can tolerate acid ranges of 0.3 M to 1.4 M, temperatures between 75 °C and 100 °C, and times of 3-5 hours. Modifications outside of those parameters may be necessary, but we have found them to consistently provide optimal results.

In cases of unusual lignin structures, including those that have a lower degree of polymerization and/or a high acid soluble lignin content, modifications will need to be made to obtain the most optimal results. Given the nature of the formaldehyde extraction procedure, no modifications will need to be made; however, the propionaldehyde-based procedure may need to be modified to obtain optimal yields. For some biomass sources, the solubility of the extracted propionaldehyde stabilized-lignin can be substantially altered. Normally during the procedure, we perform a final ether trituration of the lignin to extract residual stabilized-sugars, but unusual lignins may be partially soluble in ether and performing this step may remove a portion of the lignin and drastically reduce the yield of monomers after subsequent hydrogenolysis. The impact of this phenomenon can be seen during the extraction of lignin from High Syringyl Poplar where the low Klason Lignin content gives comparatively high hydrogenolysis yields versus other biomass sources when compared on the basis of that component (Figure 28). This comparative advantage is reduced when the yields are instead compared on the basis of the total biomass. Though this is partially due to the reduced total lignin content in the plant, the higher fraction of acid soluble lignin content in HSP relative to other biomass sources contributes significantly to this advantage (see Table 4). Because of the modified solubility of this unusual lignin, the ether trituration step had to be eliminated for HSP poplar to achieve the high yields presented in Figure 28 for the extracted propionaldehyde lignin. Depending on your needs, you may wish to avoid this step as well.

B.1.3 Depolymerization of the extracted biopolymers

Once the biomass is fractionated, the cellulose and lignin can be depolymerized using enzymatic hydrolysis and hydrogenolysis respectively. The cellulose produced from the propionaldehyde-based fractionation and washed with a saturated sodium bicarbonate solution can be used directly for enzymatic hydrolysis, leading to near quantitative yields of glucose. For the formaldehyde-based fractionation, the formaldehyde grafting can dramatically impact the enzymatic hydrolysis. Dilute sulphuric acid can cleave the acetals on the cellulose back-bone, improving digestibility. Depending on the source biomass and any additional modifications to the procedure, the concentration of the sulphuric acid may need to be varied along with the temperature and time of the reaction to obtain the best enzymatic hydrolysis results. As for the hydrogenolysis, the solvent, temperature, time, and catalyst loading can dramatically impact the yield and distribution of the monomers from the reaction. In contrast, the reaction seem to be insensitive to hydrogen pressure as we have observed nearly identical monomer yields with pressures as low as 3 bar. However, given that catalyst reducibility can be highly dependent on various factors (storage conditions, identity of the metal precursors used for the catalyst preparation, time elapsed since preparation, etc.), we recommend operating the hydrogenolysis at 40 bar of hydrogen to avoid any issues associated with this variable. Here we present one set of conditions that should provide a good gauge of the quantity of monomers that can be produced from a given stabilized lignin sample and use this to determine the quality of the stabilized lignin.

B.2 Procedure

B.2.1 Formaldehyde biomass fractionation protocol

Timing for Complete Procedure: ~ 10 hours 5 minutes
Timing for Isolation of Formaldehyde-Stabilized Lignin: ~ 6 hours 5 minutes

B.2.1.1 Pretreatment of the biomass

Timing: ~ 4 hours 35 minutes

1. Mass the extracted and dried biomass (4.5 g) into a 60 mL, thick-walled, glass reactor with an oval, PTFE-coated stir-bar (20 mm length x 10 mm diameter).
2. Add to the reactor sequentially formaldehyde (37 wt%, 5.2 mL, 66 mmol, 2.6 equiv.), 1,4-dioxane (25 mL), and hydrochloric acid (37 wt%, 2.1 mL, 25 mmol, 1.0 equiv.).

CAUTION! Formaldehyde is extremely toxic, use proper protective equipment and a fume hood while handling it.

CAUTION! 1,4-dioxane is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical

CAUTION! Hydrochloric acid is extremely corrosive. Use proper protective equipment and a fume hood while handling it.

3. Heat the reaction to 95 °C with stirring for 3.5 hours. Swirling the reaction solution every 30 minutes to ensure homogeneity.

CRITICAL STEP! Incomplete extraction of lignin from the biomass will result should the reaction not be properly stirred.

4. Cool the reaction to room temperature.

B.2.1.2 Cellulose collection
Timing: ~ 2 hours 50 minutes

5. Assemble a filtration apparatus consisting of a 250 mL filter flask, a neoprene adapter, and a ground-glass-frit Büchner funnel (porosity grade 3).
6. Filter the reaction to collect the cellulose and wash it with dioxane (2 x 10 mL) followed by methanol (2 x 10 mL) to ensure full extraction of the cellulose, which will have a pink hue. Set the filtrate aside for further processing under the *Formaldehyde-Stabilized Lignin Collection* procedure.

CAUTION! Methanol is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical.

7. Transfer the cellulose to a 60 ml, thick-walled, glass reactor with an oval, PTFE-coated stir-bar (20 mm length x 10 mm diameter).
8. Add 25 ml of 1 wt% H_2SO_4 aqueous solution to the reactor.
9. Heat the reaction to 140 ^{o}C with stirring for 1 hour.
10. Assemble a filtration apparatus consisting of a 250 mL filter flask, a neoprene adapter, and a ground-glass-frit Büchner funnel (porosity grade 3).
11. Filter the reaction to collect the cellulose and wash it with 50 mL of deionized water followed by 20 mL of acetone.

CAUTION! Acetone is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical.

12. Transfer the cellulose to a tared, 29/32, 100 mL, round-bottom flask washing with dichloromethane.
13. Add dichloromethane (~10 mL) to the flask then remove the organic solvent *in vacuo* on a rotavap (40 °C bath temperature, 300 mbar to 10 mbar).

CAUTION! Dichloromethane is highly toxic. Use proper protective equipment and a fume hood while handling it.

14. Re-tare the flask to obtain the mass of the isolated cellulose.

B.2.1.3 Formaldehyde-stabilized lignin collection

Timing: ~ 1 hour 40 minutes

15. Transfer the filtrate from the previous section to a tared, 29/32, 250 mL, round-bottom flask washing with 1,4-dioxane (10 mL).
16. Re-mass the round-bottom flask and remove a 1 mL aliquot. Place it into an HPLC vial and cap the vial.
17. Inject the aliquot on the C18 reverse-phase HPLC to determine the quantity of 2-furfural and 5-hydroxymethylfurfural produced in the pre-treatment reaction. This data will be relevant for the cellulose and hemicellulose quantifications.
18. Re-mass the round-bottom flask to determine the amount of solution removed.
19. Gradually add saturated $NaHCO_3$ solution (35 mL) to the filtrate from the previous section and swirl the flask until the acid is neutralized.

CAUTION! This neutralization will result in vigorous bubbling due to the formation of CO_2. Proceed with care.

20. Concentrate the solution using a rotavap (35 °C bath temperature, 60 mbar final pressure). The dioxane will evaporate causing the formaldehyde-stabilized lignin to precipitate.
21. Assemble a filtration apparatus consisting of a 250 mL filter flask, a neoprene adapter, and a membrane filtration apparatus with a nylon membrane filter.
22. Filter the solution washing with de-ionized water (~50 mL). Let the brown filter cake air dry for 10 min. Set aside the filtrate for further processing as described under the *Formylated C_5 Sugar Collection* section.
23. Transfer the filter cake into a tared, 29/32, 100 mL, round-bottom flask and dry the filter cake in a vacuum desiccator (~15 mbar) overnight to afford the formaldehyde-stabilized lignin as a dark-brown powder.

PAUSEPOINT. Once the lignin is transferred to the desiccator, the fractionation procedure can be paused overnight. Or, if the lignin is the only desired product, discard the filtrate that was set aside.

B.2.1.4 Formylated C_5 sugar collection

Timing: ~ 2 hours 5 minutes

24. Transfer the filtrate to 29/32, 250 mL, round-bottom flask and concentrate the solution using a rotavap (40 °C bath temperature, 40 mbar final pressure) to approximately 50 mL.
25. Transfer the concentrated solution to a 250 mL separatory funnel washing with ethyl acetate (5 mL) and de-ionized water (5 mL) and dilute it with ethyl acetate (50 mL).

CAUTION! Ethyl acetate is highly flammable. Ensure that there are no open flames or spark generating devices nearby while handling this chemical.

26. Shake the separatory funnel and separate the layers.

27. Collect the organic layer and return the aqueous fraction to the separatory funnel. Repeat the extraction of the aqueous layer twice more with ethyl acetate (50 mL)
28. Transfer the organic fractions into a 29/32, 500 mL, round-bottom flask washing with ethyl acetate (10 mL).
29. Concentrate the ethyl acetate solution using a rotavap (40 °C bath temperature, 25 mbar final pressure).
30. Add this concentrated ethyl acetate solution dropwise into a 250 mL Erlenmeyer flask containing 100 mL of hexanes being stirred at 700 RPM with a bar-type, PTFE-coated stir-bar (30 mm length, 10 mm diameter).

CAUTION! Hexanes is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical.

31. Assemble a filtration apparatus consisting of a 250 mL filter flask, a neoprene adapter, and a ground-glass-frit Büchner funnel (porosity grade 4).
32. Filter the reaction to remove the insoluble impurities washing with hexanes (10 mL).
33. Transfer the filtrate to a tared, 29/32, 500 mL, round-bottom flask washing with hexanes (10 mL)
34. Concentrate the solution using a rotavap (40 °C bath temperature, 10 mbar final pressure) to afford the diformylxylose as a yellow oil ≥95% pure by ^{1}H-NMR.

PAUSEPOINT. Having completed the procedure to this point, the fractionated materials can be stored on the benchtop in sealed vials for at least three months prior to proceeding with the *Enzymatic Cellulose Hydrolysis* or the *Lignin Hydrogenolysis* discussed in subsequent sections.

B.2.2 Propionaldehyde biomass fractionation protocol

Timing for Complete Procedure: ~ 10 hours 20 minutes
Timing for Isolation of Propionaldehyde-Stabilized Lignin: ~ 6 hours 40 minutes

B.2.2.1 Pretreatment of the biomass

Timing: ~ 4 hours 20 minutes

35. Mass the extracted and dried biomass (4.5 g) into a 29/32, 100 mL, round-bottom flask containing an oval, PTFE-coated stir-bar (20 mm length x 10 mm diameter).
36. Add to the flask sequentially propionaldehyde (4.8 mL, 67 mmol, 6.6 equiv.), 1,4-dioxane (25 mL), and hydrochloric acid (37 wt%, 0.85 mL, 10 mmol, 1.0 equiv.).

CAUTION! Propionaldehyde is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical

CAUTION! 1,4-dioxane is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical

CAUTION! Hydrochloric acid is extremely corrosive. Use proper protective equipment and a fume hood while handling it.

37. Fit a 29/32, Dimroth condenser onto the flask and connect it to a source of cooling water.
38. Fit a gas bubbler onto the top of the reflux condenser to create an air-lock.

CRITICAL STEP! This air-lock is essential for the complete extraction of the biomass.

39. Heat the reaction to 85 °C with stirring for 3 hours.

CRITICAL STEP! Incomplete extraction of lignin from the biomass will result should the reaction not be properly stirred.

40. Cool the reaction to room temperature.

B.2.2.2 Cellulose collection

Timing: ~ 1 hour 25 minutes

41. Assemble a filtration apparatus consisting of a 250 mL filter flask, a neoprene adapter, and a ground-glass- frit Büchner funnel (porosity grade 3).
42. Filter the reaction to collect the cellulose and wash it with dioxane (2 x 10 mL) followed by methanol (2 x 10 mL) to ensure full extraction of the cellulose, which will have a pink hue.

CAUTION! Methanol is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical.

43. Set the filtrate aside for further processing under the *Propionaldehyde-Stabilized Lignin Collection* procedure and place the Büchner funnel containing the filter cake on another 250 mL filter flask.
44. Without pulling a vacuum, add 20 mL of saturated sodium bicarbonate solution to the cellulose and stir it with a spatula. The solution will bubble, and the cellulose will turn from a pinkish hue to a light grey.
45. Let the cellulose solution rest for 30 min then pull a vacuum on the filtration apparatus.
46. Wash the cellulose with 50 mL of deionized water followed by 20 mL of acetone.

CAUTION! Acetone is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical.

47. Transfer the cellulose to a tared, 29/32, 100 mL, round-bottom flask washing with dichloromethane.

48. Add dichloromethane (~10 mL) to the flask then remove the organic solvent *in vacuo* on a rotavap (40 °C bath temperature, 300 mbar to 10 mbar) to afford the cellulose as light-grey, fibrous material.
49. Re-tare the flask to obtain the mass of the isolated cellulose.

B.2.2.3 Propionaldehyde-stabilized lignin collection

Timing: ~ 2 hours 30 minutes

50. Add $NaHCO_3$ (1.680 g, 20 mmol, 2.0 equiv.) and a bar-type, PTFE-coated stir-bar (30 mm length x 10 mm diameter) to the filtrate from the previous section.
51. Stir the solution for 30 min or until the acid is neutralized. If it doesn't neutralize, add more methanol (20 mL).
52. Assemble a filtration apparatus consisting of a 250 mL filter flask, a neoprene adapter, and a ground-glass-frit Büchner funnel (porosity grade 3).
53. Filter the reaction to remove the $NaHCO_3$ and NaCl then transfer the filtrate into a tared, 29/32, 250 mL, round-bottom flask washing with 1,4-dioxane (10 mL)
54. Re-mass the round-bottom flask, remove a 1 mL aliquot, place it into an HPLC vial, and cap the vial.
55. Re-mass the round-bottom flask to determine the amount of solution removed.
56. Inject the aliquot set aside on the C18, reverse-phase HPLC to determine the quantity of 2-furfural and 5-hydroxymethylfurfural produced in the pre-treatment reaction. This data will be relevant for the cellulose and hemicellulose quantifications.
57. Concentrate the solution using a rotavap (40 °C bath temperature, 25 mbar final pressure)
58. To the resulting dark brown oil, add ethyl acetate (10 mL). The solution should not be viscous and should be easily pipette-able.
59. Add the solution rinsing with ethyl acetate (5 mL) drop-wise with a pipette to a 500 mL Erlenmeyer flask containing 250 mL of hexanes being stirred at 700 RPM with a bar-type,

PTFE-coated stir-bar (30 mm length, 10 mm diameter). A reddish-brown precipitate will form.

CAUTION! Ethyl acetate is highly flammable. Ensure that there are no open flames or spark generating devices nearby while handling this chemical.

CAUTION! Hexanes is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical.

60. Assemble a filtration apparatus consisting of a 500 mL filter flask, a neoprene adapter, and a membrane filtration apparatus with a nylon membrane filter.
61. Filter the hexanes solution through the filtration apparatus washing with more Hexanes.
62. Collect the filter cake in a tared, 29/32, 250 mL, round-bottom flask.
63. Transfer the filtrate to a 29/32, 500 mL, round-bottom flask washing with diethyl ether.
64. Add diethyl ether (50 mL) to the filter cake and sonicate the solution for 5 min.

CAUTION! Diethyl Ether is highly flammable. Ensure that there are no open flames or spark generating devices nearby while handling this chemical.

65. Assemble a filtration apparatus consisting of a 500 mL filter flask, a neoprene adapter, and a membrane filtration apparatus with a nylon membrane filter.
66. Decant the diethyl ether into the filtration apparatus.
67. Add more diethyl ether (50 mL) to the filter cake then sonicate it for 5 minutes and decant it through the filtration apparatus.
68. Collect any solids that accumulated on the nylon membrane filter and transfer them into the flask containing the residual prior filter cake.
69. Transfer the diethyl ether solution into the flask containing the hexanes and ethyl acetate solution from the earlier precipitation.

70. Dry the filter cake *in vacuo* using a rotavap (40 °C bath temperature, 25 mbar final pressure) to afford the propionaldehyde-stabilized lignin as a reddish-brown powder.

PAUSEPOINT Once the lignin is transferred to the desiccator, the fractionation procedure can be paused overnight. Or, if the lignin is the only desired product, discard the filtrate that was set aside.

B.2.2.4 Propylated C_5 sugar collection

Timing: ~ 3 hours 15 minutes

71. Concentrate the hexanes, ethyl acetate, and diethyl ether solution from the previous section (*Propionaldehyde-Stabilized Lignin Collection*) *in vacuo* on a rotavap (40 °C bath temperature, 25 mbar final pressure).
72. To the resulting dark-brown oil, add diethyl ether (10 mL). The solution should not be viscous and should be easily pipette-able.
73. Add this solution rinsing with diethyl ether (5 mL) drop-wise using a pipette into a 500 mL Erlenmeyer Flask containing 250 mL of hexanes being stirred at 700 RPM with a bar-type, PTFE-coated stir-bar (30 mm length, 10 mm diameter).
74. Add activated carbon (1 g) to the filter flask. Stir for 10 minutes.
75. Assemble a filtration apparatus consisting of a 500 mL filter flask, a neoprene adapter, and a ground-glass-frit Büchner funnel (porosity grade 3). Create a 1 cm pad of vacuum-compressed Celite® in the ground-glass-frit Büchner funnel.
76. Filter the hexanes solution through the filtration apparatus washing with more hexanes (25 mL).
77. Transfer the filtrate to a 29/32, 500 mL, round-bottom flask washing with hexanes and concentrate *in vacuo* on a rotavap (40 °C, 25 mbar final pressure). The resulting yellow oil should be ≥70 wt% dipropylxylose by NMR assuming that the impurities are largely alkyl

in nature (R–CH_2–R). To purify the dipropylxylose further, follow the remaining steps.

78. Prepare a 100 g, silica-gel column on an Automated Column Machine using a hexanes:ethyl acetate gradient with an initial solvent ratio of 94:6 and that increases over 10 column volumes to 50:50. For a detailed description of how to run such a column, please see the *Equipment Setup: Automated Column Machine* section.

CAUTION! Silica gel is known to cause silicosis. Use proper protective equipment and a fume hood while handling it.

79. Once the column has equilibrated, load the yellow oil from step 10 onto the column washing with hexanes and then run the programmed sequence collecting all the fractions.
80. Transfer the appropriate fractions as determined by comparison of the R_f of their contents with that provided for dipropylxylose in the anticipated results section into a 29/32, 500 mL, round-bottom flask. There are two diastereomers of the dipropylxylose that are produced.
81. Concentrate the collected fractions using a rotavap (40 °C, 25 mbar final pressure).
82. Transfer the resulting yellow oil washing with ethyl acetate into a tared, 29/32, 100 mL, round-bottom flask washing with ethyl acetate.
83. Concentrate the resulting solution using a rotavap (40 °C, 10 mbar final pressure) to afford the dipropylxylose as a mixture of diastereomers. The diastereomers of dipropylxylose can be separated, but it takes multiple silica gel columns along with the cutting and pooling of fractions to achieve the separation.

PAUSEPOINT. Having completed the procedure to this point, the fractionated materials can be stored on the benchtop in sealed vials for at least three months prior to proceeding with the *Enzymatic Cellulose Hydrolysis* or the *Lignin Hydrogenolysis* discussed in subsequent sections.

B.2.3 Enzymatic cellulose hydrolysis

Complete Section Timing: ~ 78 hours 40 minutes

B.2.3.1 Enzymatic hydrolysis of cellulose

Timing: ~ 78 hours 40 minutes

84. Prepare 50 mL of a 0.1 M pH 5 citrate buffer by diluting trisodium citrate dihydrate (956 mg, 3.25 mmol) and citric acid monohydrate (368 mg, 1.75 mmol) to 50 mL with Milli-Q water in a 50 mL volumetric flask.
85. Prepare a tetracycline stock solution by first dissolving the tetracycline (20 mg, 0.045 mmol) with 1.4 mL of absolute ethanol and then diluting the resulting solution with 0.6 mL of Milli-Q water in a 5 mL vial with screw cap.

CAUTION! Tetracycline is toxic. Use proper protective equipment and a fume hood while handling it.

CAUTION! Ethanol is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical.

86. Prepare a cycloheximide stock solution by combining cycloheximide (20 mg, 0.071 mmol) with 2 mL of Milli-Q water in a 5 mL vial with screw cap.

CAUTION! Cycloheximide is toxic. Use proper protective equipment and a fume hood while handling it.

87. Mass 300 mg of cellulose into three separate 20 mL glass vials with PTFE coated stir-bars.
88. To each vial, add via pipette 11.3 mL of the citrate buffer, 0.4 mL of the tetracycline solution, and 0.3 mL of the cycloheximide solution.
89. Cap the vials and place them in a shaking incubator at 250 RPM and 50 °C for 1 hour
90. Remove the vials and add 10 Filter Paper Units (FPU) of cellulases.

91. Return the vials to the incubator at 250 RPM and 50 °C and continue to heat and shake them for 72 hours
92. Let the vials cool to room temperature then transfer their contents to 50 mL volumetric flasks and dilute to 50 mL with Milli-Q water
93. Shake the flasks and let any solids settle then remove a 1 mL aliquot, filter it though a syringe filter into an HPLC vial.
94. Cap the vial and inject the solution on the pH 2 aqueous phase HPLC to determine the glucose concentration.
95. Calculate the yield of glucose from the cellulose taking into account the hydration of the cellulose (see below)

B.2.3.2 Determine the cellulose hydration

Timing: ~ 17 hours 30 minutes

96. Mass 300 mg of cellulose into three, separate, tared, 50 mL, self-standing centrifuge tubes.
97. Lightly cap the centrifuge tubes and place them into a vacuum oven at 60 °C and dry them for at least 16 hours *in vacuo* (~50 mbar final pressure).
98. Remove the centrifuge tubes from the vacuum oven and cool them in a vacuum desiccator at room temperature (~25 mbar) for an hour.
99. Re-mass the centrifuge tubes and calculate the mass loss. This is the hydration of the cellulose.

B.2.3.3 Determine the glucose content of the cellulose

Timing: ~ 38 hours 20 minutes

100. Prepare 25 mL of 72 wt% H_2SO_4 (Specific Gravity = 1.634 $g\cdot mL^{-1}$) by adding 30.97 g of concentrated sulphuric acid to 8 g of deionized water in a 25 mL volumetric flask and then diluting with de-ionized water to a final solution volume of 25 mL.

CAUTION! This dilution is extremely exothermic. Always add acid to water and not *vice versa*. Let the solution cool to room temperature before diluting to 25 mL.

CAUTION! Sulphuric acid is extremely corrosive. Use proper protective equipment and a fume hood while handling it.

101. Into three, separate, 50 mL, self-standing centrifuge tubes, add a 0.2 µm nylon membrane filter.
102. Place the centrifuge tubes from step 101, lightly-capped, into the vacuum oven at 60 °C.
103. Into the three centrifuge tubes from the previous section, *Determine the Cellulose Hydration*, add oval stir-bars (20 mm long x 10 mm diameter).
104. Into each centrifuge tube, add 7.5 mL of 72 wt% (12 M) H_2SO_4.
105. Cap the centrifuge tubes, shake and vortex them to distribute the solid, and sonicate them for 2 hours at 30 °C.
106. Transfer the contents of the centrifuge tubes to 500 mL reagent bottles with GL 45 polypropylene caps and dilute the solutions to ~250 mL with Milli-Q water.
107. Autoclave the bottles for 1 hour at 120 °C.
108. Transfer the hot solutions (ca. 85 °C) to a refrigerator and let them cool.

CAUTION! These solutions will be extremely hot.

PAUSEPOINT. The procedure can be paused here for the day as is recommended, or once the solutions have cooled one can proceed.

109. The next day, remove the centrifuge tubes containing the nylon membrane filters from the vacuum oven and cool them in a vacuum desiccator for 1 hour at room temperature (~25 mbar).
110. Mass the centrifuge tubes and record the mass.
111. Remove the reagent bottles from the refrigerator and filter the solutions through the dried,

tared, 0.2 µm nylon membrane filters washing with Milli-Q water.

112. Place the nylon membrane filters and filter cakes into their corresponding centrifuge tubes and lightly cap those centrifuge tubes. Place them in a vacuum oven at 60 °C and dry them for 24 hours *in vacuo* (~50 mbar final pressure). If there is residual precipitate adhered to the walls of the filtration apparatus after the filtration, wash it into the centrifuge tubes using ethanol.
113. Transfer the filtrates to separate 500 mL volumetric flasks diluting with Milli-Q water then return the filtrates to the 500 mL reagent bottles.
114. Remove 1 mL from each of the 500 mL reagent bottles and filter it through a syringe filter into an HPLC autosampler vial. Label and cap the HPLC vial and then inject it on the pH 2 HPLC column to determine the concentration of glucose, xylose, 5-hydroxymethylfurfural, and 2-furfural in the sample. When presenting the data, add the HPLC responses ($g \cdot L^{-1}$) of 5-hydroxymethylfurfural and 2-furfural reconstituted as glucose (multiply the 5-hydroxymethylfurfural response by 1.43) and xylose (multiply the 2-furfural response by 1.56) to the observed yields for those of glucose and xylose. Use **Equation (4)** from **Box 4** to calculate the contribution of each sugar to the overall mass of the material.
115. Remove the filters and filter cakes from the vacuum oven along with their centrifuge tubes from step 112 and cool them in a vacuum desiccator for 1 hour at room temperature (~25 mbar).
116. Mass the filters and filter cakes and subtract the mass of the filters to determine the Klason lignin content.

B.2.4 Lignin hydrogenolysis

Timing: 6 hours 40 minutes (Formaldehyde-Stabilized Lignin) or 5 hours 40 minutes (Propionaldehyde-Stabilized Lignin)

117. Add the stabilized lignin (200 mg), ruthenium on carbon (5 wt%, 100 mg), and tetrahydrofuran (20 mL) into a 50 mL Parr reactor with bar-type, PTFE-coated stir-bar (20 mm length x 10 mm diameter).

CAUTION! Tetrahydrofuran is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical.

CAUTION! Ruthenium on carbon is toxic. Use proper protective equipment and a fume hood while handling it.

118. Seal the Parr reactor and then back-fill it with H_2 gas by filling it with 40 bar of H_2 and slowly releasing the pressure.

CAUTION! Hydrogen gas is highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical.

CAUTION! High pressure gas is in use. Use proper protective equipment and appropriate apparatuses for filling and running the reaction.

119. Repeat the back-fill for a total of 3 times.
120. Fill the Parr reactor with 40 bar of H_2 gas.
121. Heat the Parr reactor to 250 °C with stirring for 4 hours for formaldehyde stabilized lignin and 3 hours for propionaldehyde stabilized lignin. Start the timer as soon as the reactor begins heating.

CRITICAL STEP! At 250 °C, lower monomer yields may result if the reaction is left longer than the prescribed amount of time through degradation or overreduction. Reaction temperatures as low as 175 °C can be used; however, more time will be required to convert the lignin (12+ hours). Here we present optimal conditions for the determination of the monomer yield from the stabilized lignin for the biomass sources used in this paper.

CAUTION! The reactor is extremely hot. Handle with care.

122. Let the Parr reactor cool to room temperature (ca. 25 °C).
123. Release the hydrogen gas and open the Parr reactor.
124. Add 200 µL of the *n*-decane stock solution to the reaction solution and stir it with a spatula.

CAUTION! The *n*-decane stock solution is toxic and highly flammable. Use proper protective equipment and a fume hood while handling it. Also, ensure that there are no open flames or spark generating devices nearby while handling this chemical.

125. Using a 20 mL syringe, withdraw the reaction solution from the Parr reactor.
126. Filter the reaction solution through a syringe filter to remove the catalyst.
127. Take a sample of the filtrate and inject it on the gas chromatography instrument using the method described under *Equipment Setup: Gas Chromatography*.
128. Integrate the appropriate peaks and, using the effective carbon number, calculate the yield of the reaction as described under *Equipment Setup: Gas Chromatography*.

B.3 Equipment setup

Autoclave

1. Seal all the reagent bottles to be autoclaved with their GL 45 Polypropylene Caps.
2. Fill one empty reagent bottle with de-ionized water to roughly the same level as those in the reagent bottles. Leave un-capped.
3. Place all the reagent bottles into the autoclave with the un-capped reagent bottle roughly in the centre.
4. Insert the thermocouple(s) into the uncapped reagent bottle.
5. Seal the autoclave.
6. Heat the autoclave to a temperature of at least 121 °C and a pressure of 220 kPa. (ca. 1 hour)
7. Hold at the autoclave at that temperature and pressure for 1 hour.
8. Let the autoclave slowly cool and exhaust to 90 °C and ~100 kPa (ca. 1.5 hours).

Automated Column Machine

Starting and Running a Silica Gel Column

1. Select an appropriately sized silica gel cartridge. Typically, a mass ratio of 1:50 is an advisable starting point with more difficult separations requiring more silica gel and *vice versa*. As the automated column machine cartridges typically come in pre-set sizes, always choose the slightly larger version for initial separation attempts. Here, we use 100 g cartridges due to the size of the sample and the difficulty of the separation.
2. Using the display, program the sequence to include the following phases with the respective column volumes of solvent (solvent required to fill the voids in the packed column being used). For the 100 g cartridge, 1 Column Volume is approximately 125 mL and the flow rate should be 50 $mL \cdot min^{-1}$.
    a. Equilibration at the initial solvent conditions (3 Column Volumes)
    b. After Loading Resting Phase at the initial solvent conditions (1 Column Volumes)
    c. Solvent Ramp (10 Column Volumes)
    d. Column Wash (2 Column Volumes)
3. Indicate in the method the type of test tube rack being used and fill the test tube rack with test tubes.
4. Set the sample collection to collect all fractions and enable the UV-Vis to observe the 254 and 280 nm bands.
5. Load the silica gel cartridge and start the run sequence. After equilibration, the system will stop and prompt you to load your sample. Press “load sample” and using a 10 mL syringe and Hexanes, quantitatively transfer the sample to the top of the silica gel cartridge.
6. Re-start the run and wait until it is completed.
7. Empty the waste container used to collect any run-off and the equilibration solvent.
8. Collect the test tube racks.

Test Tube Fraction Analysis

1. To determine which fractions to collect from the racks of test-tubes, thin layer chromatography (TLC) will be employed. As the products are not chromophores, $KMnO_4$ TLC stain will be used to visualize the TLC plates after they have been developed.
2. Obtain a 20 cm x 20 cm TLC plate and cut it into 5 cm x 5 cm strips. Draw a line in pencil about 0.75 cm from the bottom of the 5 cm x 5 cm TLC strip. Along this line, draw a perpendicular score every 2-3 mm then write an even number below that line from 2 to the number of scores on the strip. Repeat the same procedure on another TLC strip until the highest even number matches the number of test tube fractions.
3. Spot the numbered TLC plates using the TLC spotter and the correspondingly numbered test tube fractions.
4. Place the TLC plate into a TLC chamber containing the solvent mixture specified in the anticipated results section for the desired isolated compound.
5. The TLC plate will draw solvent up the silica gel through capillary action. Once the solvent line is ~0.75 cm from the top, withdraw the plate and mark where the solvent advanced to using a pencil.
6. Air dry the TLC plate then stick it completely in the $KMnO_4$ TLC stain. This stain requires heat to develop, so place it in an oven for a minute. The $KMnO_4$ is purple and it will oxidize oxidizable chemical species present on the TLC plate leaving a yellowish-brown spot.
7. Based on the R_f (fractional distance of elution versus the distance of solvent advance) of the reported compound, narrow down the range of possible fractions and repeat steps 2-6, but this time running every fraction of interest.
8. Collect the appropriate fractions in a 1 L, 29/32, round-bottom flask.
9. Concentrate the fractions on a rotavap (40 °C, 25 mbar final pressure) to afford the product.

Gas Chromatography

1. Column and Injection Conditions
2. Injection temperature: 250 °C.
3. Column Temperature: 50 °C for 1 min, ramp at 15 °C·min^{-1} to 300 °C (16 minutes, 40 seconds), and hold at 300 °C for 7 min.
4. Carrier Gas: N_2 at 25 mL·min^{-1}
5. FID detection temperature: 290 °C
6. FID gases: H_2 at 30 mL·min^{-1} and Synthesis Air at 400 mL·min^{-1}

Monomer yield quantification

1. Integrate the area of monomers and decane in the GC-FID chromatogram.
2. Use the following **Equations (B1-B4)** to calculate the yield of each monomer based on its integrated area and Effective Carbon Number (ECN). Abbreviations: mass (m), moles (n), molecular weight (MW), area of peak (A)
3. For the yield as a percent of the Klason Lignin content, use **Equation (B5)**. Abbreviations: MKL% (Monomer as a weight percent of Klason Lignin), wt% (weight percent).

$$n_{decane} = \frac{m_{decane\ in\ sample}}{MW_{decane}} \tag{B1}$$

$$n_{monomer} = \frac{\mathrm{A}_{monomer\ in\ sample}}{A_{decane\ in\ sample}} \cdot n_{decane} \cdot \frac{ECN_{decane}}{ECN_{monomer}} \tag{B2}$$

$$m_{monomer} = n_{monomer} \cdot MW_{monomer} \tag{B3}$$

$$Yield_{monomer} = \frac{m_{monomer}}{m_{biomass}} \cdot 100\% \tag{B4}$$

$$Yield_{MKL\%} = \frac{Yield_{monomer}}{wt\%\ Klason\ Lignin} \cdot 100\% \tag{B5}$$

High Pressure Liquid Chromatography

C18 Reverse Phase Chromatography

1. Flow Rate: 0.7 mL·min^{-1}
2. Injection Size: 1 µL
3. Column Temperature: 25 °C
4. Column Conditions
    a. Solvents: pH 7 Water (MilliQ, Ammonium Formate 1 mg·mL^{-1}) and Acetonitrile
    b. 5 min at 10% Water: 90% Acetonitrile
    c. 15 min ramp from 10% Water: 90% Acetonitrile to 90% Water: 10% Acetonitrile
    d. 5 min at 90% Water: 10% Acetonitrile
    e. 10 min at 10% Water: 90% Acetonitrile

pH 2 Aqueous Phase Chromatography

1. Flow Rate: 0.6 mL·min^{-1}
2. Injection Size: 20 µL
3. Column Temperature: 60 °C
4. Column Conditions
    a. Solvents: pH 2 Water (5 mL of 1M H_2SO_4 diluted to 1 L with MilliQ water)
    b. 60 min at 100% Water

pH 7 Aqueous Phase Chromatography

1. Flow Rate: 0.6 mL·min^{-1}
2. Injection Size: 20 µL
3. Column Temperature: 80 °C
4. Column Conditions
    a. Solvents: pH 7 Water (MilliQ, Ammonium Formate 1 mg·mL^{-1})
    b. 60 min at 100% Water

Machine Sieve

1. Stack the appropriate test sieve (0.45 mm) onto the bottom sieve pan.
2. Place the biomass to be sieved into the test sieve then cover it with the clamping lid.
3. Place the assembled sieve stack onto the machine sieve and strap it down.
4. Use the following settings to sieve out any and all fines ($\Phi < 0.45$ mm) from the biomass
    1. Pulse: 5 second period, 50% duty cycle (2.5 seconds on; 2.5 seconds off)
    2. Intensity: 8 out of 10
    3. Time: 30 minutes
5. After running the above sequence, remove the biomass and transfer the sieved materials to separate appropriate storage vessels.

NMR Parameters

^{1}H-NMR Parameters

1. NS (Number of Scans): 16
2. D_1 (Delays)= 30s
3. O1P (Transmitter frequency offset): 6.000 ppm
4. SW (Spectral Width): 14.701 ppm
5. DS (Dummy Scans) = 0

^{13}C-NMR (^{1}H decoupled) Parameters

1. NS (Number of Scans): 1024
2. D_1 (Delays)= 2s
3. O1P (Transmitter frequency offset): 100.000 ppm
4. SW (Spectral Width): 236.621 ppm
5. DS (Dummy Scans) = 4

HSQC NMR (^{1}H – ^{13}C multiplicity edited HSQC with gradient selection) Parameters

1. NS (Number of Scans): 32
2. D_1 (Delays)= 1.5s
3. O1P (Transmitter frequency offset): 4.700 ppm
4. SW (Spectral Width): 13.1536 ppm
5. DS (Dummy Scans) = 32

B.4 Boxes

Box 1 | Biomass Ash Quantification

Additional Equipment

- Aluminium Foil (Weita – Fresh)
- Muffle Furnace (100-1200°C, 220-240 V, 1560 Watts, 50/60 Hz; Thermo Scientific – F48020-33-80)
- Porcelain Crucibles (Haldenwanger – 79 MF/7)

Procedure

1. Heat three clean, dry crucibles to 120 °C for 16 hours in an oven.
2. Let the crucibles cool to room temperature in a vacuum desiccator for 1 hour at room temperature (~25 mbar).
3. Tare the crucibles then mass 1 g of the biomass into each of the crucibles.
4. Record the new mass of the crucibles and cover them with a small square of aluminium foil with holes punched in it.
5. Place the crucibles into a ventilated muffle furnace under air and heat it to 600 °C
6. Leave the samples for 24 hours. Note: any pencil or pen markings will burn off, so note the position of the crucibles on a sheet of paper so that they can be identified upon removal from the muffle furnace.
7. Remove the crucibles and cool them to room temperature in a vacuum desiccator for 1 hour at room temperature (~25 mbar).
8. Re-mass the crucibles with ashes. Use the following **Equation (B6)** to calculate the wt% of ashes in the sample.

$$Ashes\ (wt\%) = \left(\frac{m_{Crucible\ \&\ Ashes} - m_{Crucible}}{m_{Crucible\ \&\ Biomass} - m_{Crucible}}\right)\cdot 100\% \qquad \textbf{(B6)}$$

Box 2 | Biomass Hydration Quantification

Procedure

1. Into three separate, tared, 50 mL centrifuge tubes, mass 2g of biomass.
2. Record the new mass of the centrifuge tubes.
3. Lightly cap and place the tubes into a vacuum oven at 60 °C and dry them for at least 16 hours *in vacuo* (~50 mbar).
4. Remove the biomass from the vacuum oven and cool it in a vacuum desiccator for 1 hour at room temperature (~25 mbar).
5. Re-mass the biomass and calculate the mass loss using **Equation (B7)** below. This quantity is the hydration of the biomass. Abbreviations: mass (m), centrifuge tube (CFT)

$$Hydration\ (wt\%) = \left(\frac{m_{CFT\ \&\ Dry\ Biomass} - m_{CFT}}{m_{CFT\ \&\ Biomass} - m_{CFT}}\right)\cdot 100\% \qquad \textbf{(B7)}$$

Box 3 | Biomass Extractives Quantification

Additional Equipment

- Centrifuge (VWR Ventilated Benchtop Centrifuge Mega Star 1.6 – 521-1749)

Procedure

1. Mass 2 g of biomass into three tared 50 mL centrifuge tubes.
2. Record the new mass of the centrifuge tubes.
3. Prepare 400 mL of 80% ethanol by mixing 320 mL of absolute (100%) ethanol with 80 mL of Milli-Q water.
4. Add 40 mL of the 80% Ethanol to each centrifuge tube.
5. Cap the centrifuge tubes and sonic them at room temperature for 30 min
6. Centrifuge the tubes for 5 min @ 4500 RPM to separate the solids from the solution.
7. Decant the solution.
8. Repeat steps 4-7 twice more with 80% ethanol, thrice with Milli-Q water, and once with absolute ethanol.
9. Lightly cap the centrifuge tubes and place the biomass into a vacuum oven at 60 °C and dry it for at least 16 hours *in vacuo* (~50 mbar final pressure).
10. Remove the biomass from the vacuum oven and cool it in a vacuum desiccator for 1 hour at room temperature (~25 mbar).
11. Re-mass the biomass and calculate the mass loss. This quantity includes both the hydration and extractives of the biomass. Calculating the difference between the two mass losses yields the extractives. See **Equation (B8)** below to calculate this value. Abbreviations: mass (m), centrifuge tube (CFT)

$$Extractives\ (wt\%) = \left(\frac{m_{CFT\ \&\ Extracted\ Biomass} - m_{CFT}}{m_{CFT\ \&\ Biomass} - m_{CFT}}\right) \cdot 100\% - Hydration\ (wt\%) \qquad \textbf{(B8)}$$

Box 4 | Structural Sugar (Cellulose and Hemi-Cellulose), Acid Soluble Lignin, and Klason Lignin Quantification

Additional Equipment

- Agilent Technologies 1260 Infinity System with 1260 High Performance Degasser (G4225A), 1260 Binary Pump (G1312B), 1260 Automated Liquid Sampler (ALS) (G1329B), 1260 Thermostatted Column Compartment (TCC) (G1316A), 1260 Diode Array Detector (DAD) (G4212B), and 1260 Refractive Index Detector (RID) (G1362A) equipped with a BioRad Aminex HPX-87P Column (300 mm x 7.8 mm; 125-0098) and Micro-Guard De-Ashing Guard Column (Part Number: 125-0118)
- Bulb Pipette, 100 mL (Poulten & Graf – 1 2305104)
- Centrifuge tubes, 15 mL (Sarsedt – 62.554.502)
- Planetary Ball Mill (Retsch – PM 100 – 205400001)
- Reagent Bottle with GL 45 Polypropylene Cap, 250 mL (Simax – 1632414321250)

Procedure

1. Prepare 50 mL of 72 wt% H_2SO_4 (Specific Gravity = 1.634 $g \cdot mL^{-1}$) by adding 61.94 g of concentrated sulphuric acid to 16 g of deionized water in a 50 mL volumetric flask and then diluting with de-ionized water to a final solution volume of 50 mL.

CAUTION! This dilution is extremely exothermic. Always add acid to water and not *vice versa*. Let the solution cool to room temperature before diluting to 50 mL.

2. Extract and dry 5 g of biomass (see *Reagent Setup: Bulk Biomass Extractives Removal and Drying*)
3. Ball mill the biomass for 2 hours at 450 RPM using a 50% duty cycle (5 min on, 5 min off) until the biomass is a fine powder.
4. Mass 0.3 g of the ball-milled biomass and record the mass into three separate, tared, 50 mL centrifuge tubes. These will be used to determine the hydration of the ball-milled biomass.
5. Add a 0.2 µm nylon membrane filter into three separate 50 mL self-standing centrifuge tubes.
6. Place the centrifuge tubes from step 4 and 5, lightly-capped, into the vacuum oven at 60 °C.
7. Into another three separate, tared, 15 mL centrifuge tubes with oval stir-bars (20 mm long x 10 mm diameter) mass 0.5 g of the ball-milled biomass and record the mass.

8. Into each centrifuge tube, add 7.5 mL of 72 wt% (12 M) H_2SO_4 using a 1-10 mL Variable Volume Single-Channel Pipette.
9. Cap the centrifuge tubes, shake and vortex them to distribute the solid, and sonicate them for 2 hours at 30 °C.
10. Transfer the contents of the centrifuge tubes to 500 mL reagent bottles with GL 45 polypropylene caps and dilute the solutions to ~300 mL with Milli-Q water.
11. Autoclave the bottles for 1 hour at 120 °C.
12. Transfer the hot solutions (ca. 85 °C) to a refrigerator and let them cool overnight.
13. The next day, remove the centrifuge tubes from steps 4 and 5 from the vacuum oven and cool them in a vacuum desiccator for 1 hour at room temperature (~25 mbar).
14. Mass the centrifuge tubes and record the mass. Calculate the hydration of the biomass using the data from step 4 and the **Equation (B7)** from **Box 2: Biomass Hydration Quantification**.
15. Remove the reagent bottles from the refrigerator and filter the solutions through the dried, tared, 0.2 µm nylon membrane filters from step 5 washing with Milli-Q water.
16. Place the nylon membrane filters and filter cakes into their corresponding centrifuge tubes and lightly cap those centrifuge tubes. Place them in a vacuum oven at 60 °C and dry them for 24 hours *in vacuo* (~50 mbar final pressure).
17. Transfer the filtrates to separate 500 mL volumetric flasks diluting with Milli-Q water then return the filtrates to the 500 mL reagent bottles.
18. Mass $NaHCO_3$ (3 g, 35.7 mmol) into three separate 250 mL reagent bottles.
19. Using a 100 mL Mohr Pipette, transfer 100 mL of each of the diluted acidic filtrate into the reagent bottles with $NaHCO_3$.
20. Once neutralized, remove an aliquot from each of the neutralized filtrate and filter it through a syringe filter into an HPLC autosampler vial then cap it. Analyse the sample by HPLC using the pH 7 method described under *Equipment Setup* to determine the concentration of D-(+) glucose, D-(+) xylose, D-(+) galactose, L-(+) arabinose, and D-(+) mannose, 2-furfural, and 5-hydroxymethylfurfural in the filtrate. When presenting the data, add the HPLC responses ($g\cdot L^{-1}$) of 5-hydroxymethylfurfural and 2-furfural reconstituted as glucose (multiply the 5-hydroxymethylfurfural signal by 1.43) and xylose (multiply the 2-furfural signal by 1.56) to the observed yields for those of glucose and xylose. Use the following generalized **Equation (B8)** to calculate the contribution of each sugar to the overall mass of the material. Abbreviations: mass (m), High Pressure Liquid Chromatography (HPLC), Molecular Weight (MW), ball milled biomass (BMB), raw biomass (RBM), hydration (H), extractives (E).

$$Sugar\ Polymer\ (wt\%) = \left(\frac{[HPLC\ response\ (g\cdot L^{-1})] * (1\ L) * \left(\frac{MW_{sugar} - MW_{water}}{MW_{sugar}}\right)}{m_{BMB} * \left(1 - H_{BMB}(wt\%)\right)}\right) \cdot \left(1 - H_{RBM}(wt\%) - E_{RBM}(wt\%)\right) \cdot 100\% \qquad \textbf{(B8)}$$

21. Obtain a UV absorbance trace of the acidic diluted filtrate from 190 nm to 300 nm using a quartz cuvette. Record the absorbance at 205 nm. If the absorbance exceeds 2 A, dilute the solution with 0.18 M sulphuric acid (1 mL 97% sulphuric acid diluted to 100 mL with MilliQ Water) until it falls under that threshold and then record the dilution and the value of the absorbance. Typically, a dilution factor (d) of 3 is required.
22. Using the data collected from steps 20 and 21 along with the absorptivities for furfural, 5-hydroxymethylfurfural, and Acid Soluble Lignin (9.7 ± 0.3 $L\cdot g^{-1}\cdot cm^{-1}$, 20.3 ± 0.4 $L\cdot g^{-1}\cdot cm^{-1}$, 110 $L\cdot g^{-1}\cdot cm^{-1}$ respectively at 205 nm) to determine the Acid Soluble Lignin in the biomass according to the following E**quation (B9)**. Abbreviations: acid soluble lignin (ASL), ball milled biomass (BMB), raw biomass (RBM), mass (m), path length (b), absorptivity (ε), dilution factor (d), 5-hydroxymethylfurfural (HMF), hydration (H), extractives (E).

$$ASL\ (wt\%) = \left(\frac{\left(A_{ASL} - \varepsilon_{furfural} \cdot b \cdot [furfural] - \varepsilon_{HMF} \cdot b \cdot [HMF]\right)}{\varepsilon_{ASL} \cdot b \cdot m_{BMB} \cdot \left(1 - H_{BMB}(wt\%)\right)}\right) \cdot (1\ L) \cdot d \cdot \left(1 - H_{RBM}(wt\%) - E_{RBM}(wt\%)\right) \cdot 100\% \qquad \textbf{(B9)}$$

23. Remove the filters and filter cakes from the vacuum oven along with the centrifuge tubes from step 4 and cool them in a vacuum desiccator for 1 hour at room temperature (~25 mbar).
24. Mass the filters and filter cakes and subtract the mass of the filters to determine the Klason lignin

content using the following **Equation (B10)**. Abbreviations: mass (m), centrifuge tube (CFT), ball milled biomass (BMB), raw biomass (RBM), hydration (H), extractives (E).

$$Klason\ Lignin\ (wt\%) = \left(\frac{m_{CFT\ \&\ Klason\ Lignin} - m_{CFT}}{m_{Biomass} * \left(1 - H_{BMB}(wt\%)\right)}\right) \cdot \left(1 - H_{RBM}(wt\%) - E_{RBM}(wt\%)\right) \cdot 100\% \qquad \textbf{(B10)}$$

Box 5 | Determination of the Theoretical Monomer Yields from Biomass

Procedure

1. Into a 50 mL Parr reactor with bar-type, PTFE-coated stir-bar (20 mm length x 10 mm diameter) add the raw biomass (1 g), ruthenium on carbon (5 wt%, 200 mg), and methanol (20 mL).
2. Seal the Parr reactor and then back-fill it with H_2 gas by filling it with 40 bar of H_2 and slowly releasing the pressure.
3. Repeat the back-fill for a total of 3 times.
4. Fill the Parr reactor with 40 bar of H_2 gas.
5. Heat the Parr reactor to 250 °C with stirring for 15 hours. Start the timer as soon as the reactor begins heating.
6. Let the Parr reactor cool to room temperature (ca. 25 °C).
7. Release the hydrogen gas and open the Parr reactor.
8. Add 200 µL of the *n*-decane stock solution to the reaction solution and stir it with a spatula.
9. Using a 20 mL syringe, withdraw the reaction solution from the Parr reactor.
10. Filter the reaction solution through a syringe filter to remove the catalyst and other insoluble material.
11. Take a sample of the filtrate and inject it on the gas chromatography instrument using the method described under *Equipment Setup: Gas Chromatography*.
12. Integrate the appropriate peaks and using the effective carbon number, calculate the yield of the reaction as described under *Equipment Setup: Gas Chromatography* using the Effective Carbon Numbers (ECN).

B.5 Hemicellulose derivatives characterizations

Diformylxylose Characterization

(3aR,3bS,7aR,8aR)-tetrahydro-7H-[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine.

Appearance: white crystalline solid

TLC (3:1 v/v Hexanes: Ethyl Acetate, visualized with $KMnO_4$) $R_f = 0.2$

^{1}H NMR (400 MHz, Chloroform-*d*): δ 6.01 (d, *J* = 4.0 Hz, 1H), 5.01 (d, *J* = 20.0 Hz, 2H), 4.91 (d, *J* = 4.0 Hz, 1H), 4.58 (d, *J* = 4.0 Hz, 1H), 4.40 (d, *J* = 4.0 Hz, 1H), 4.22 (s, 1H), 4.21 (d, *J* = 12.0 Hz, 1H), 3.91 (s, 1H), 3.82 (dd, *J* = 2.0, 12.0 Hz, 1H)

^{13}C NMR (101 MHz, Chloroform-*d*): 105.1, 96.7, 91.6, 83.7, 77.7, 74.8, 65.9

Mass Spectrometry (GC-MS-EI): Calculated for $C_7H_{11}O_5$ (M-H$^+$) = 173.0; Found = 173.0

Dipropylxylose Characterization

(2R,3aR,3bS,5R,7aR,8aR)-2,5-diethyltetrahydro-7H-[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine.

Appearance: white crystalline solid

TLC (3:1 v/v Hexanes: Ethyl Acetate, visualized with $KMnO_4$) $R_f = 0.54$

^{1}H NMR (400 MHz, Chloroform-*d*): δ 5.91 (d, *J* = 4.0 Hz, 1H), 4.84 (t, *J* = 4.0 Hz, 1H), 4.36 (d, *J* = 4.0 Hz, 1H), 4.33 (t, *J* = 5.3 Hz, 1H), 4.20 (d, *J* = 13.2 Hz, 1H), 4.15 (d, *J* = 4.0 Hz, 1H), 3.97 – 3.92 (m, 1H), 3.86 (dd, *J* = 13.2, 2.0 Hz, 1H), 1.63 (qd, *J* = 7.5, 4.6 Hz, 2H), 1.55 (qdd, *J* = 7.5, 5.3, 1.2 Hz, 2H), 0.89 (t, *J* = 7.5 Hz, 3H), 0.84 (t, *J* = 7.5 Hz, 3H).

^{13}C NMR (101 MHz, Chloroform-*d*): δ 105.90, 105.18, 101.17, 84.26, 78.25, 72.39, 65.99, 27.72, 26.86, 8.10, 7.57.

HSQC (Chloroform-d): See B.6 Figures.

Mass Spectrometry (APPI): Calculated for C11H19O5 (M+H+) = 231.1227; Found = 231.1226

(2S,3aR,3bS,5R,7aR,8aR)-2,5-diethyltetrahydro-7H-[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine.

Appearance: white crystalline solid

TLC (3:1 v/v Hexanes: Ethyl Acetate, visualized with $KMnO_4$) $R_f = 0.51$

^{1}H-NMR (400 MHz, Chloroform-*d*): δ 5.98 (d, *J* = 3.6 Hz, 1H), 5.10 (t, *J* = 4.5 Hz, 1H), 4.44 (d, *J* = 3.6 Hz, 1H), 4.35 (t, *J* = 5.2 Hz, 1H), 4.25 – 4.17 (m, 2H), 3.88 – 3.78 (m, 2H), 1.58 (qd, *J* = 7.5, 4.5 Hz, 2H), 1.56 (qd, *J* = 7.5, 5.2 Hz, 2H), 0.87 (t, *J* = 7.5 Hz, 3H), 0.86 (t, *J* = 7.5 Hz, 3H).

^{13}C NMR (101 MHz, Chloroform-*d*): δ 107.75, 105.35, 101.12, 84.33, 78.33, 74.92, 66.23, 27.82, 27.64, 8.11, 7.44.

HSQC (Chloroform-d): See B.6 Figures.

Mass Spectrometry (APPI): Calculated for $C_{11}H_{19}O_5$ $(M+H^+)$ = 231.1227; Found = 231.1229

B.6 Figures

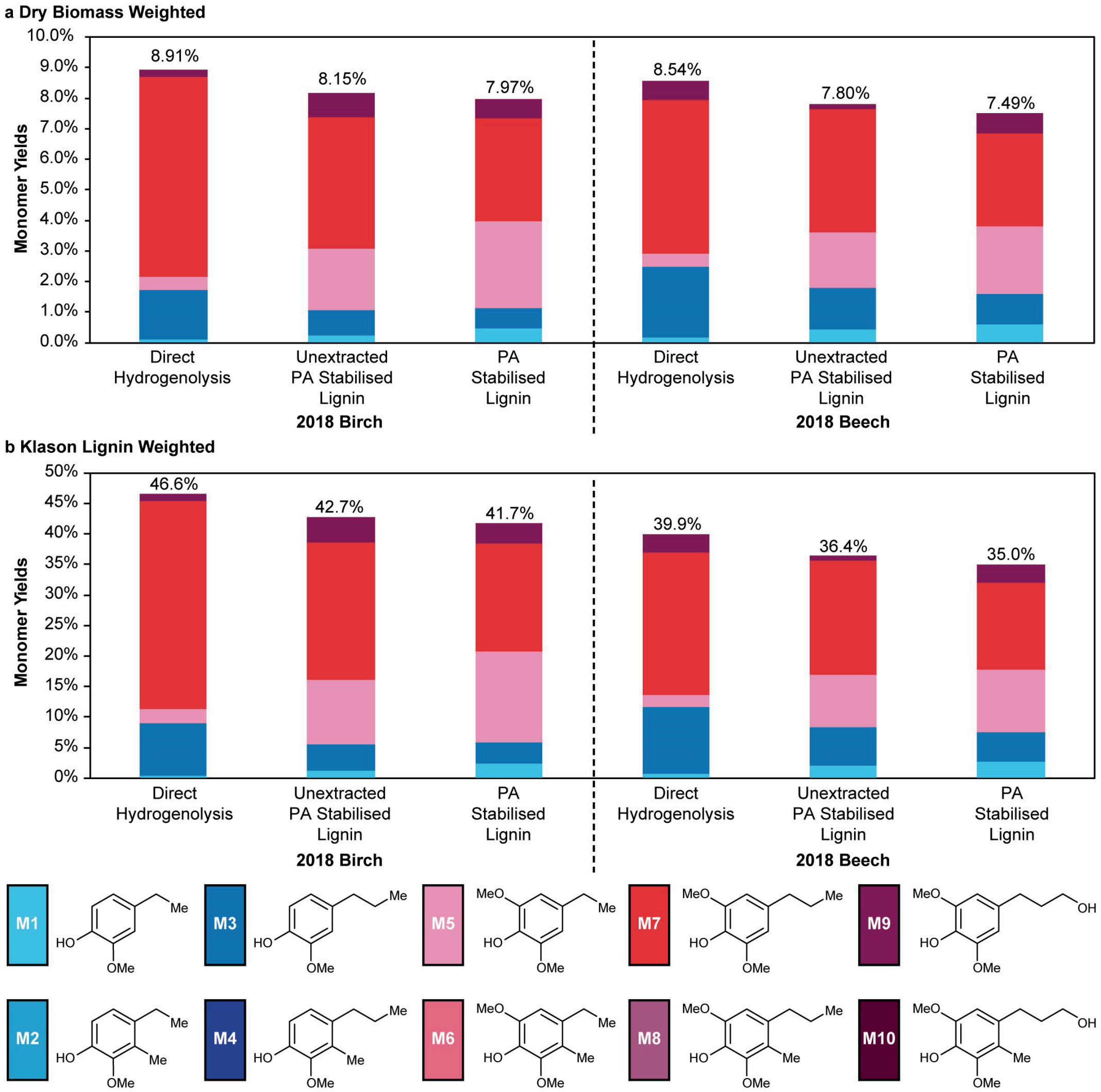


Figure B1. Hydrogenolysis Data for the Unextracted and Extracted Propionaldehyde Stabilised Lignins Compared with the Direct Hydrogenolyses of the Feedstock Biomass. These two charts compare the monomer yields from the hydrogenolysis of the raw biomass (Direct Hydrogenolysis), Propionaldehyde stabilised lignin derived from unextracted wood, and propionaldehyde stabilised lignin derived from extracted wood for two biomass sources: 2018 Birch, and 2018 Beech. The direct hydrogenolysis represents the highest possible yield of monomers for these biomass sources and was performed on biomass that had not been extracted or dried. The difference in yields between the extracted and unextracted propionaldehyde protected lignins is approximately 1% on a Klason Weighted Basis.

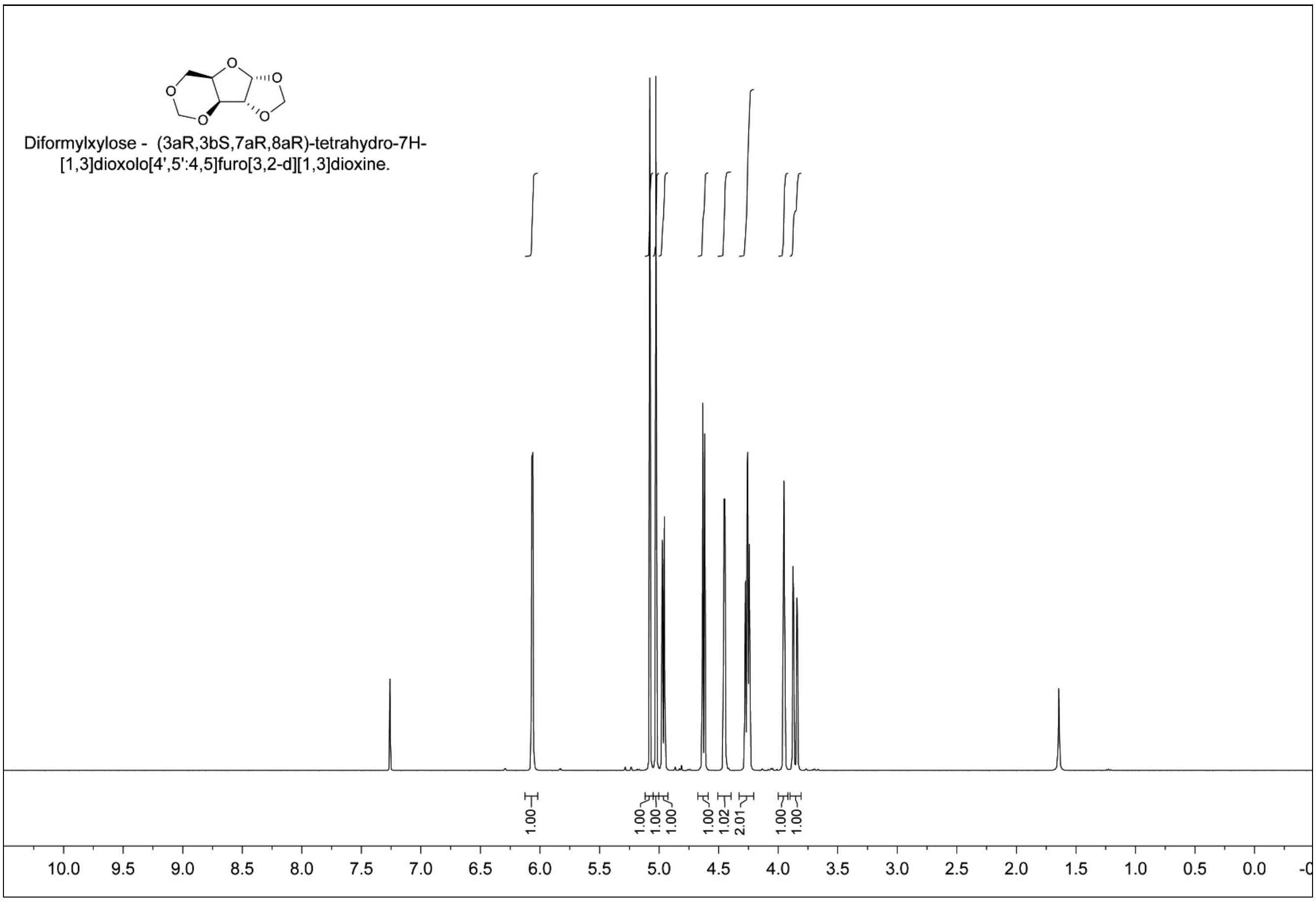

Diformylxylose - (3aR,3bS,7aR,8aR)-tetrahydro-7H-
[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine.
O
1.00
1.00
1.00
1.00
1.00
1.02
2.01
1.00
1.00
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

Figure B2

1H-NMR of Diformylxylose in CDCl3.

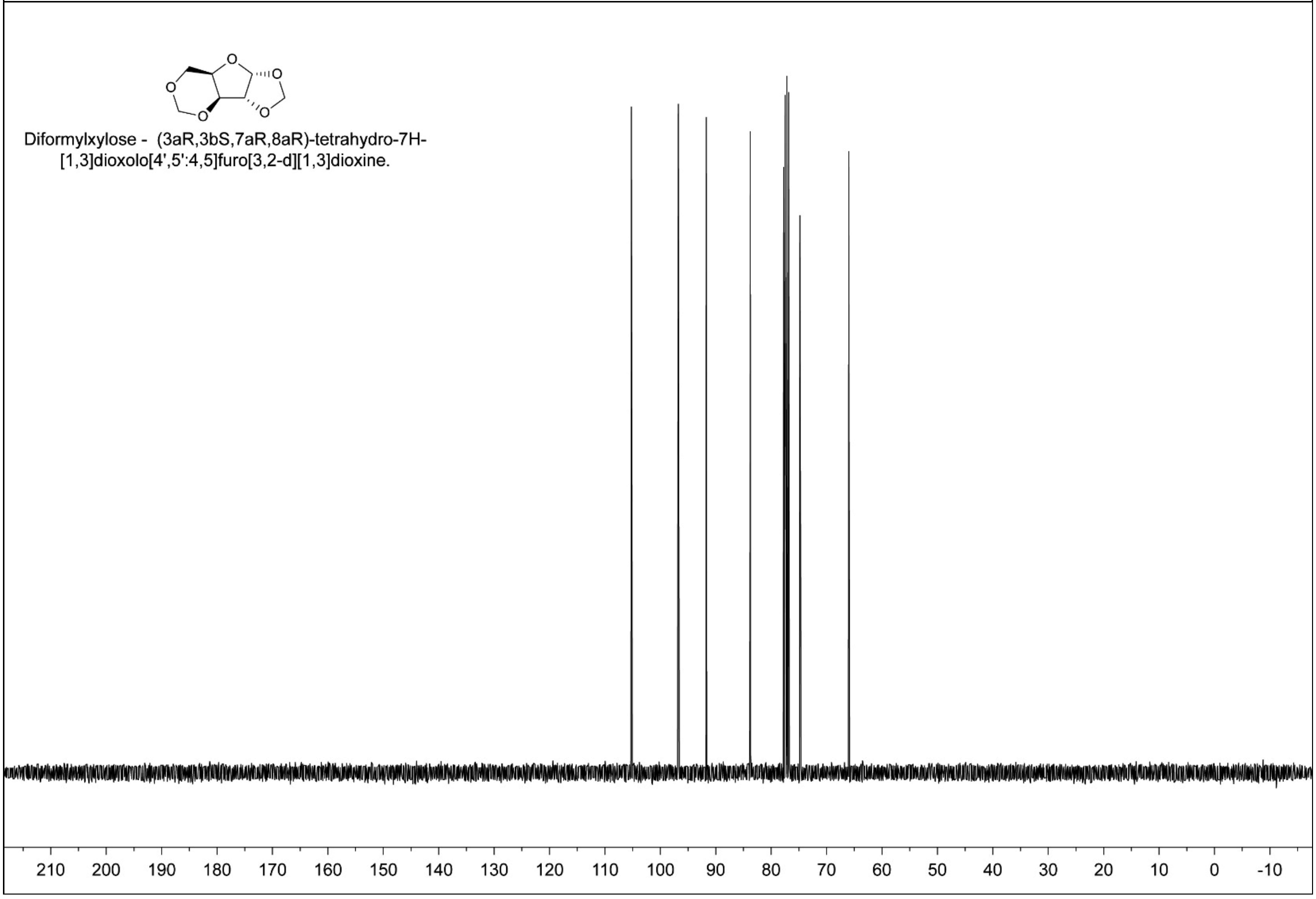


Figure B3

13C-NMR of Diformylxylose in CDCl3.

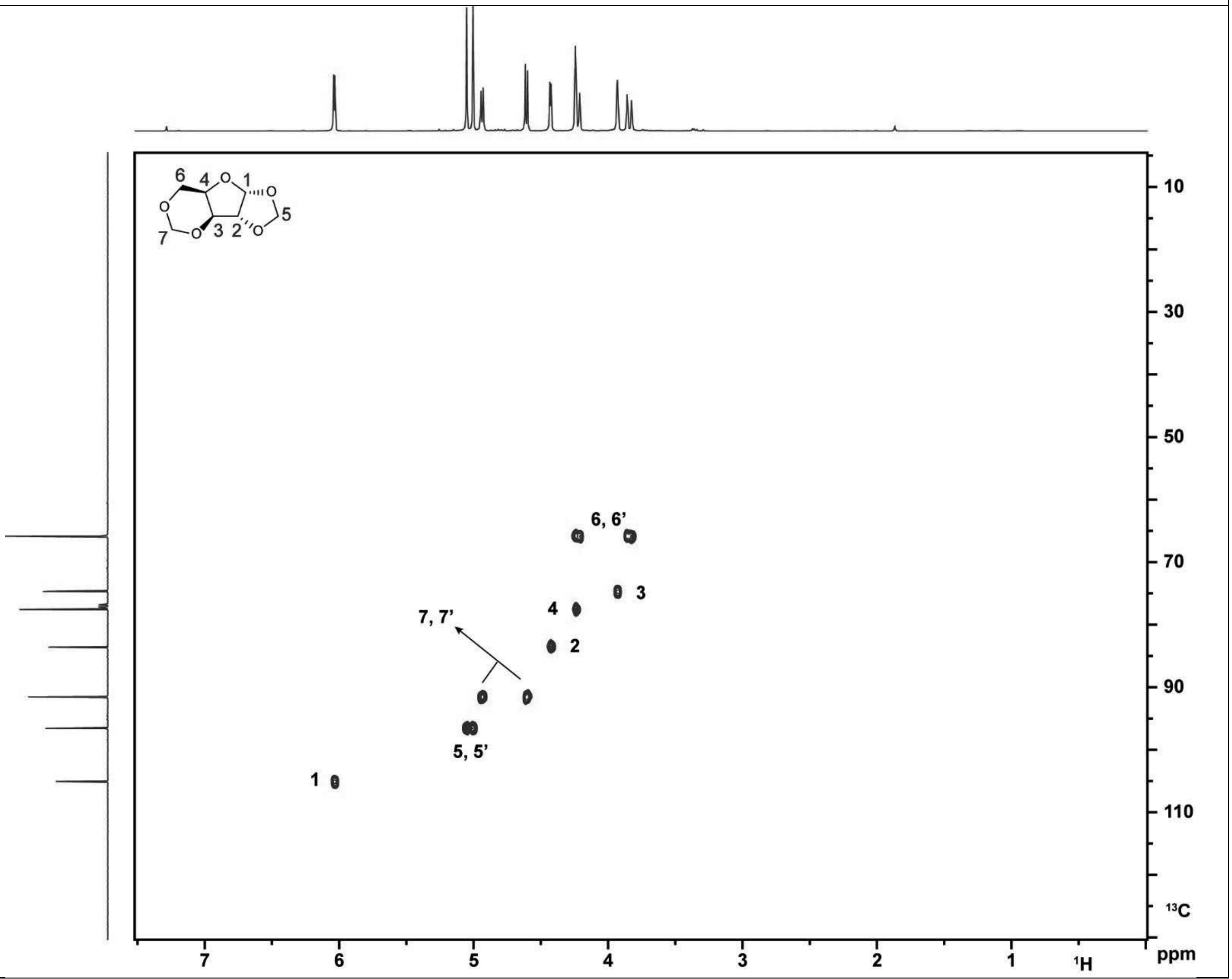


Figure B4

HSQC of Diformylxylose in CDCl3.

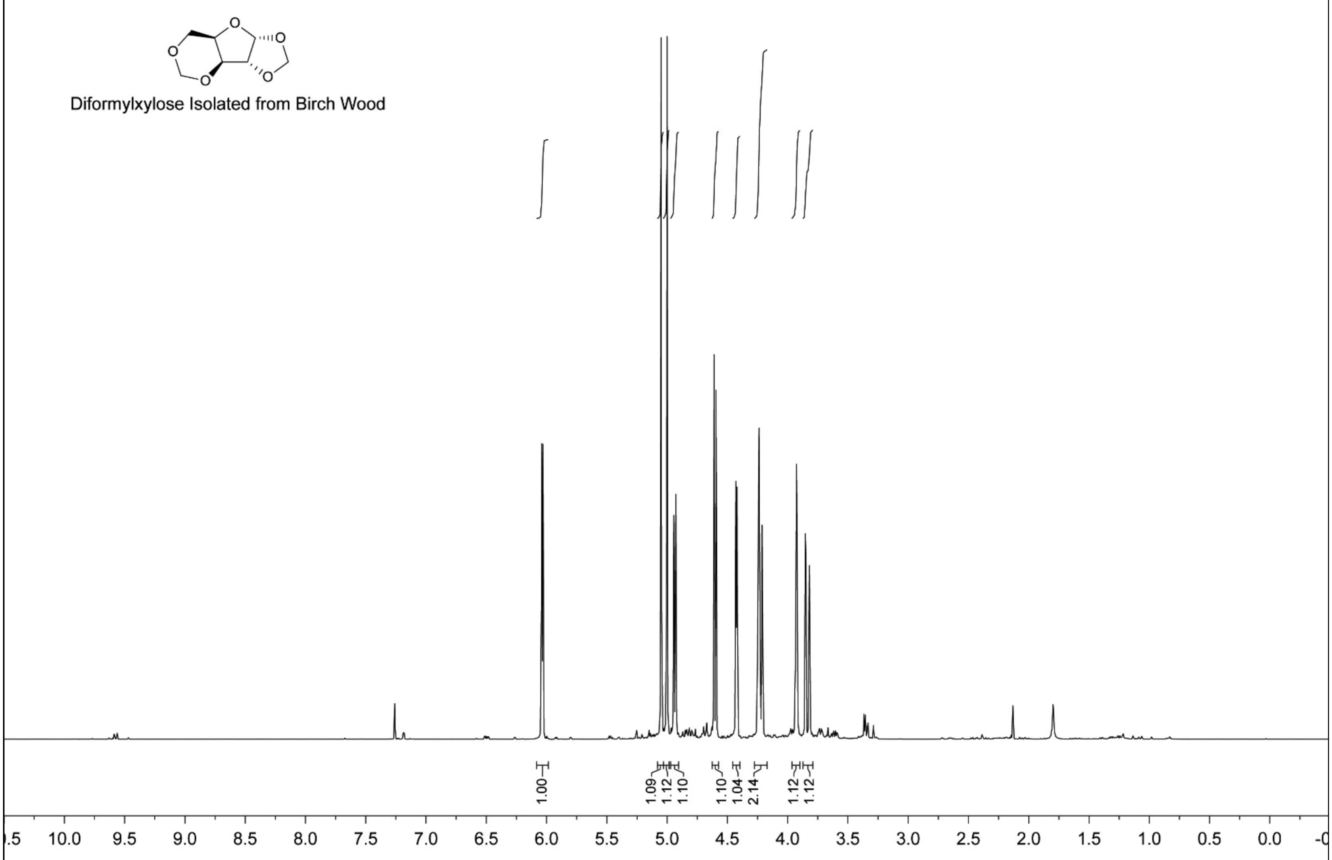


Figure B5

1H-NMR in CDCl3 of Diformylxylose Isolated from the 2018 Birch Wood Using the Formaldehyde Fractionation Protocol.

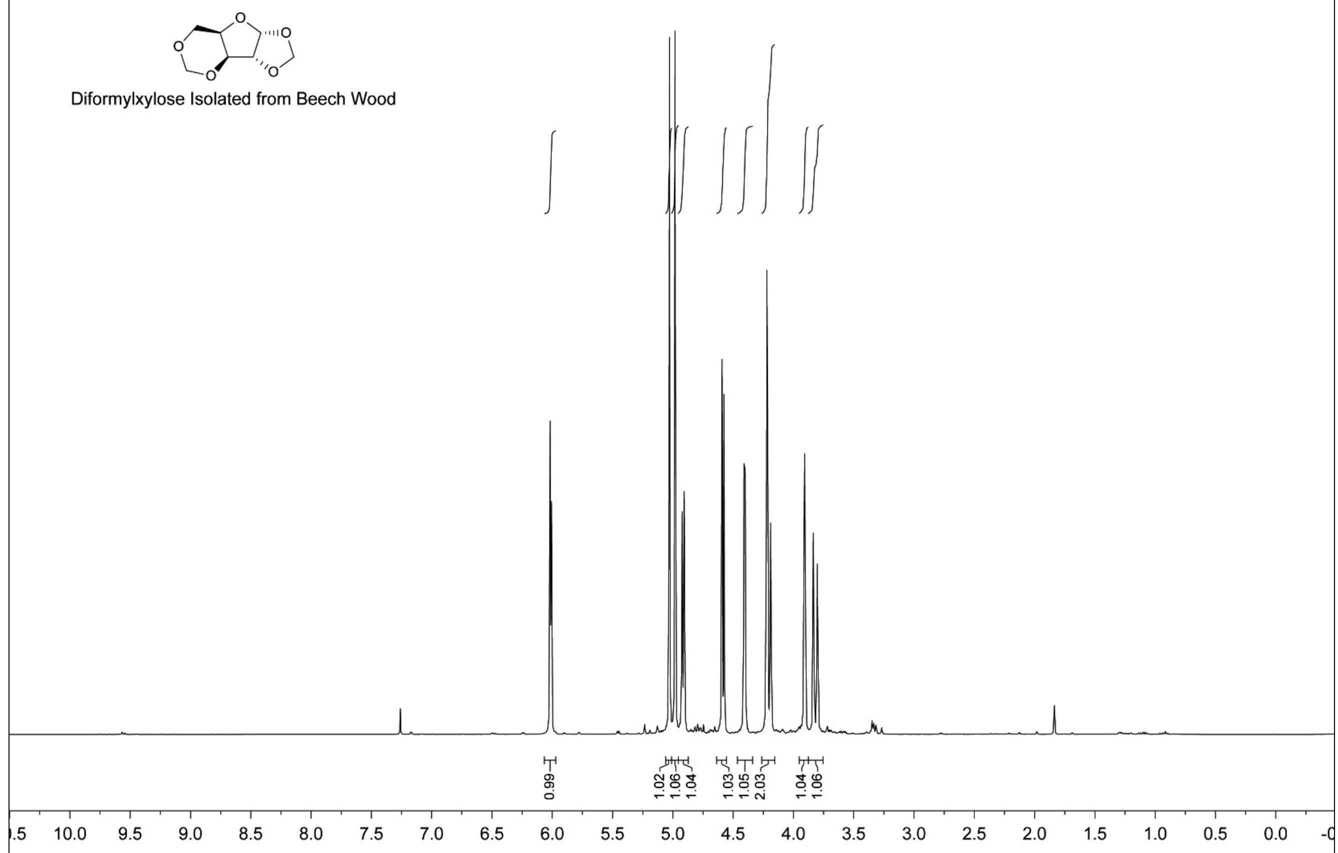
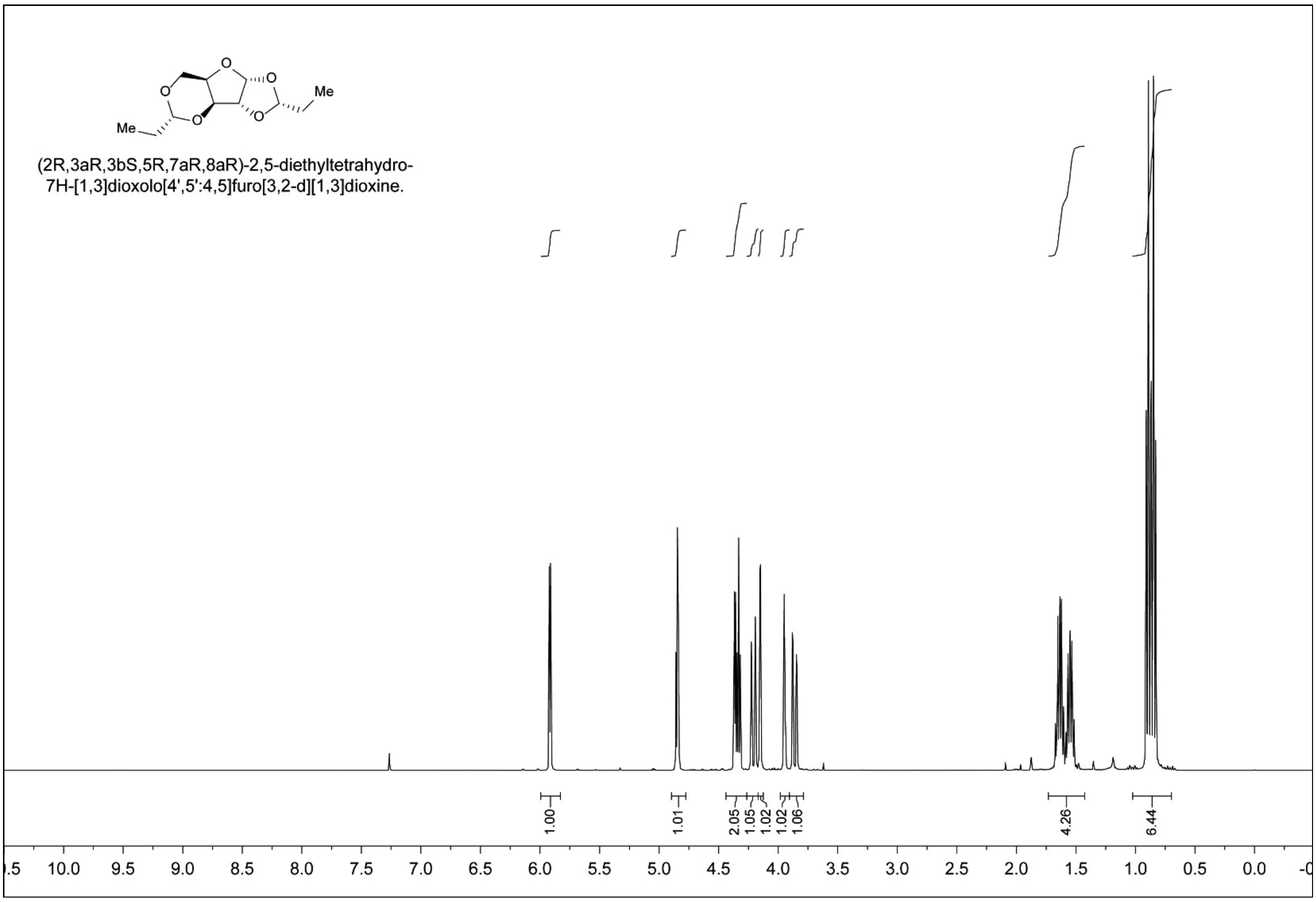


Figure B6

1H-NMR in CDCl3 of Diformylxylose Isolated from the 2018 Beech Wood Using the Formaldehyde Fractionation Protocol.



(2R,3aR,3bS,5R,7aR,8aR)-2,5-diethyltetrahydro-7H-[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine.

Figure B7

1H-NMR of (2R,3aR,3bS,5R,7aR,8aR)-2,5-diethyltetrahydro-7H-[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine (Dipropylxylose) in CDCl3.

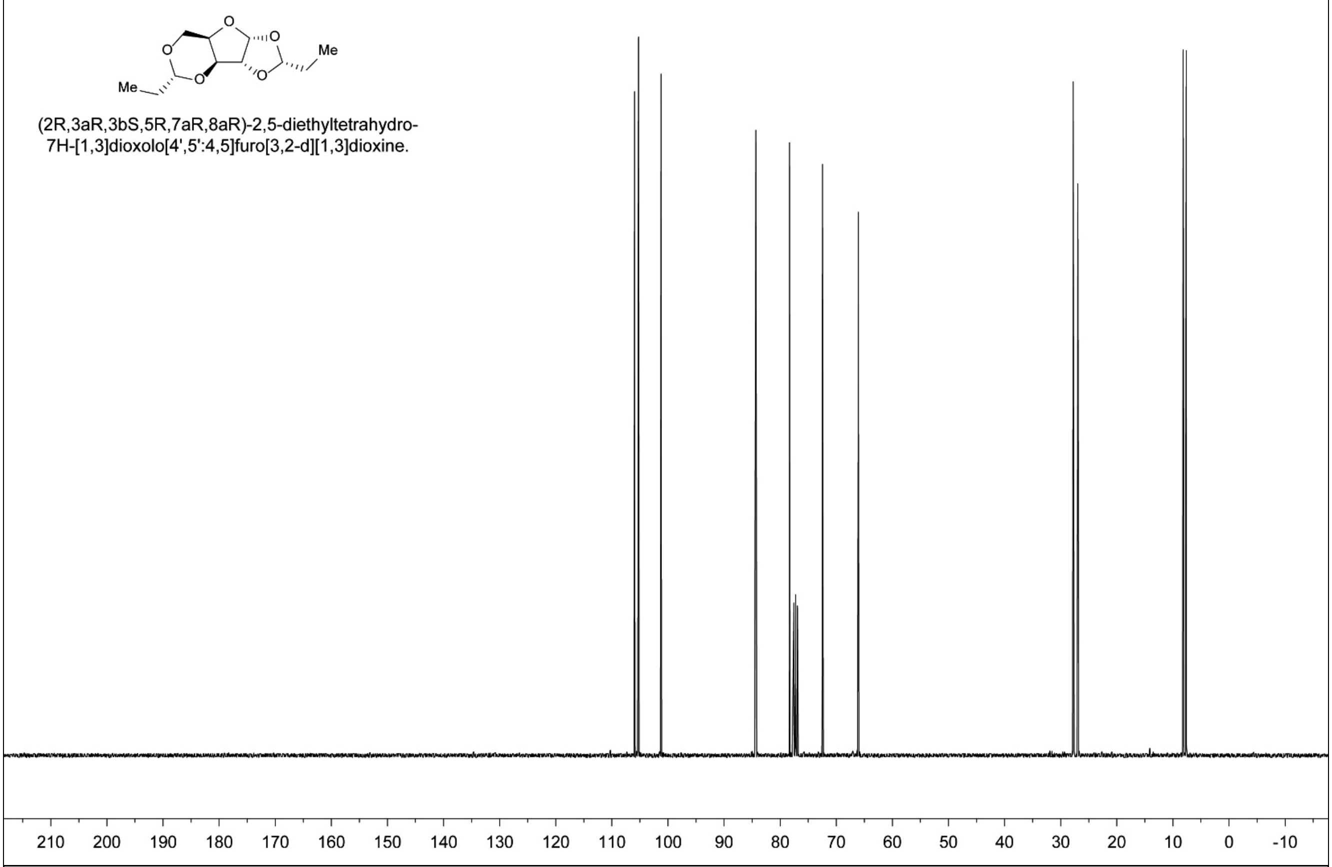


Figure B8

13C-NMR of (2R,3aR,3bS,5R,7aR,8aR)-2,5-diethyltetrahydro-7H-[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine (Dipropylxylose) in CDCl3.

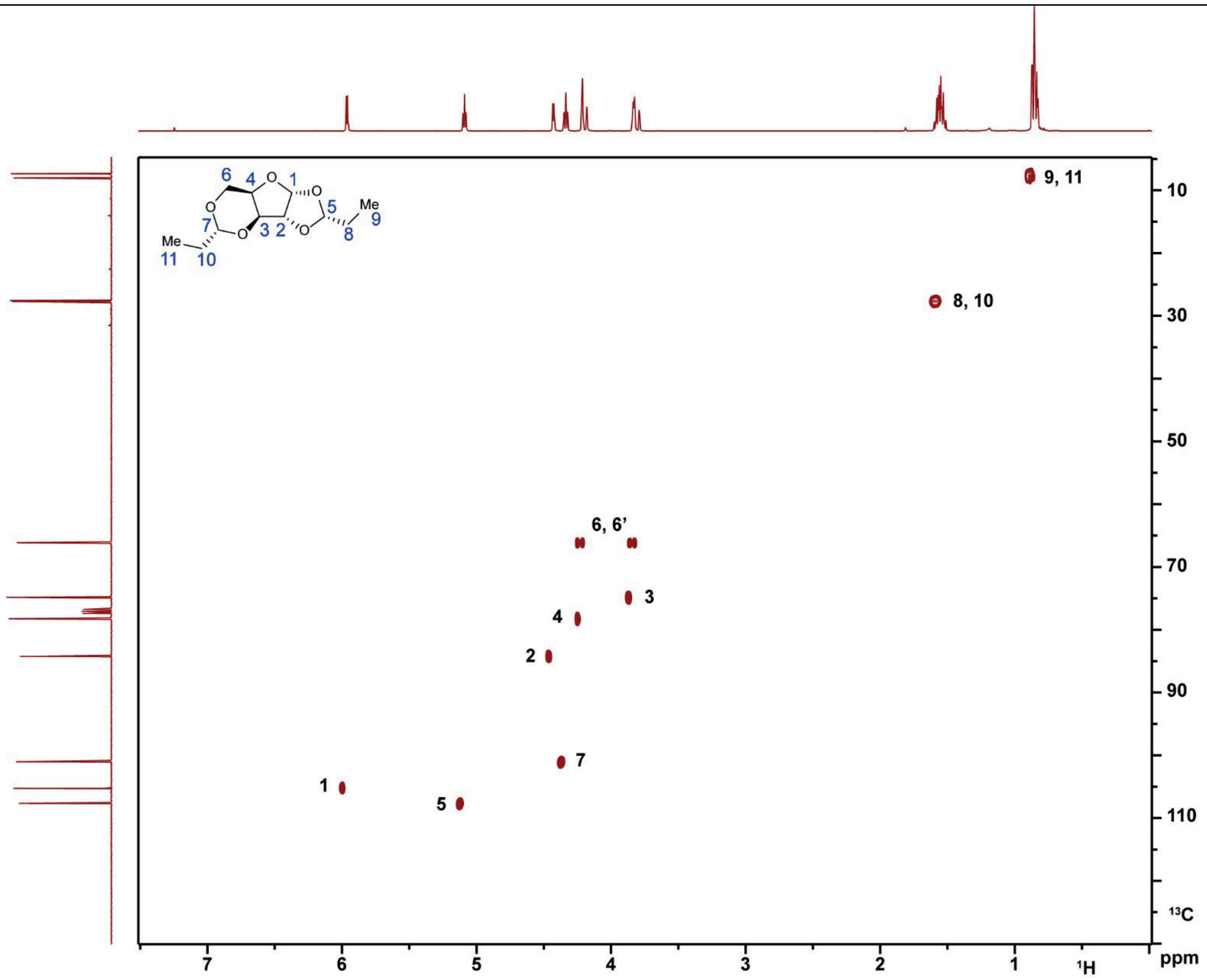


Figure B9

HSQC of (2R,3aR,3bS,5R,7aR,8aR)-2,5-diethyltetrahydro-7H-[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine (Dipropylxylose) in CDCl3.

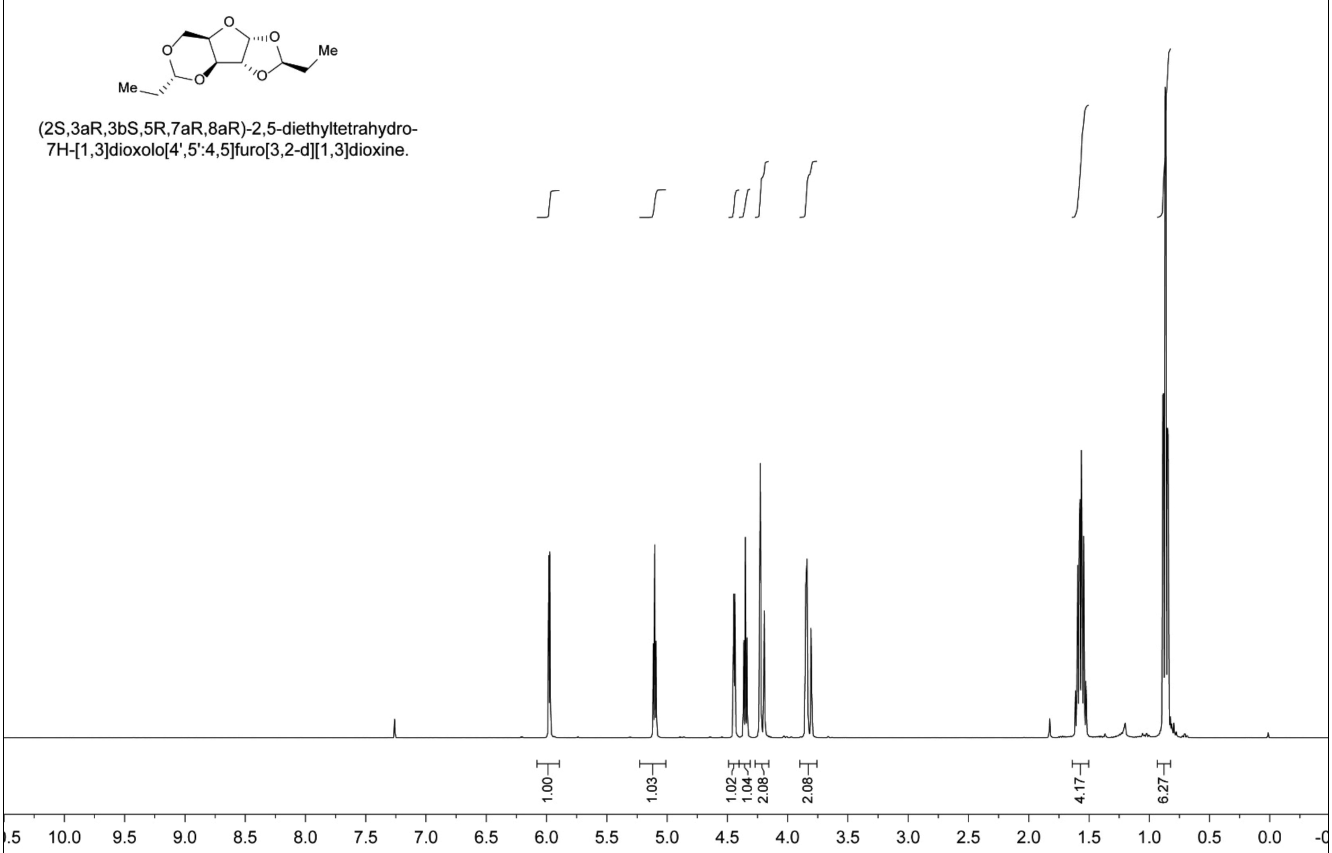


Figure B10

1H-NMR of (2S,3aR,3bS,5R,7aR,8aR)-2,5-diethyltetrahydro-7H-[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine.dioxine (Di-propylxylose) in CDCl3.

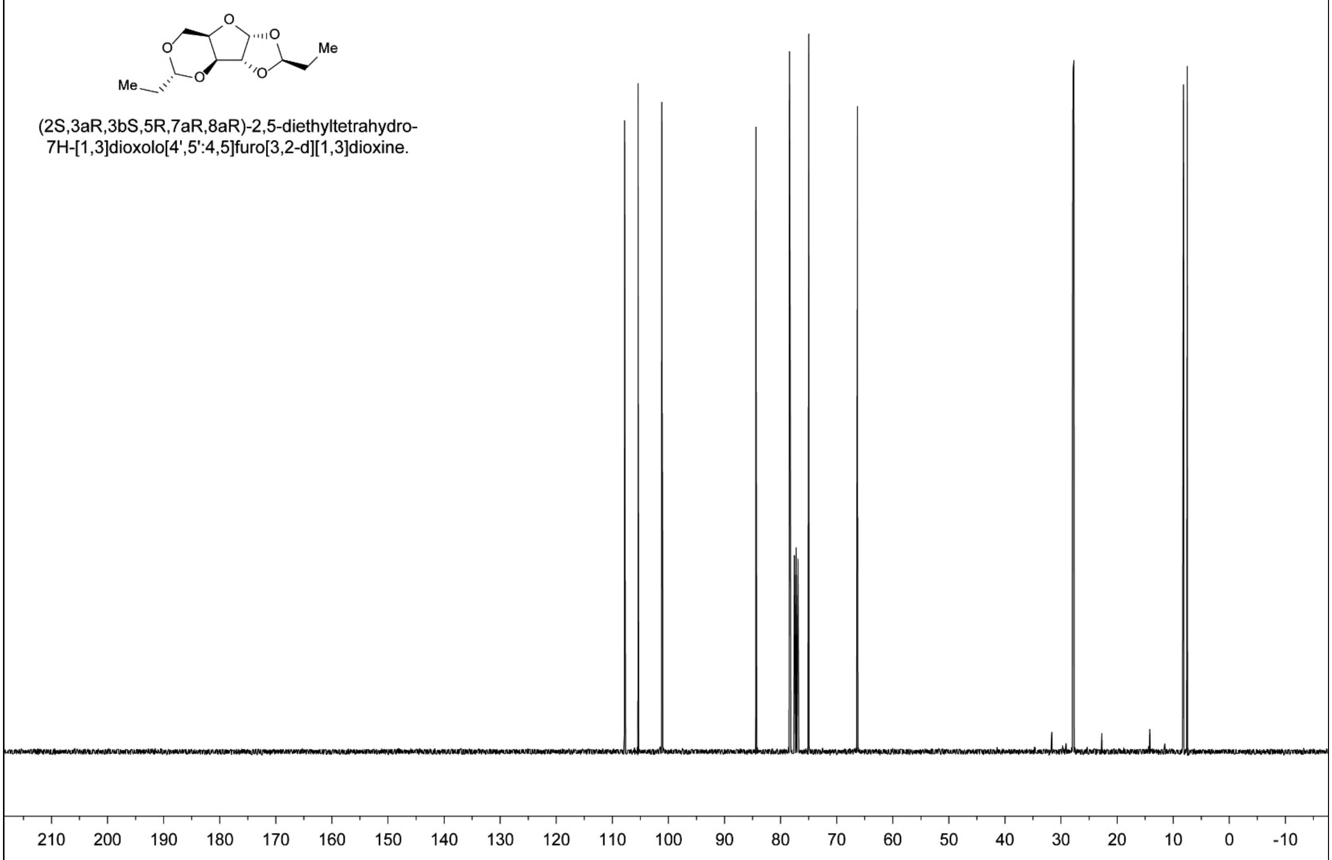


Figure B11

13C-NMR of (2S,3aR,3bS,5R,7aR,8aR)-2,5-diethyltetrahydro-7H-[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine.dioxine (Di-propylxylose) in CDCl3.

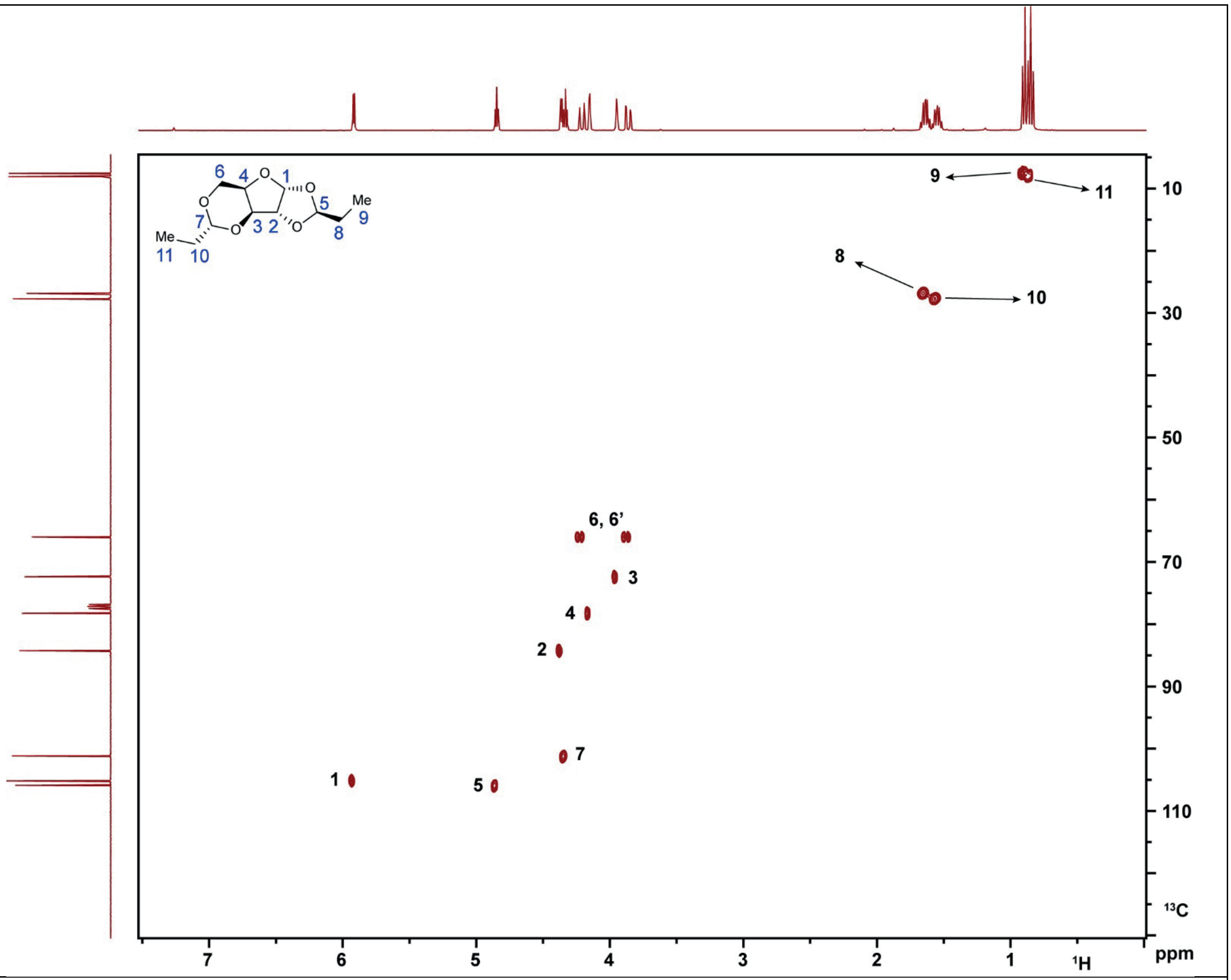


Figure B12

HSQC of Purified (2S,3aR,3bS,5R,7aR,8aR)-2,5-diethyltetrahydro-7H-[1,3]dioxolo[4',5':4,5]furo[3,2-d][1,3]dioxine.dioxine (Dipropylxylose) in CDCl3.

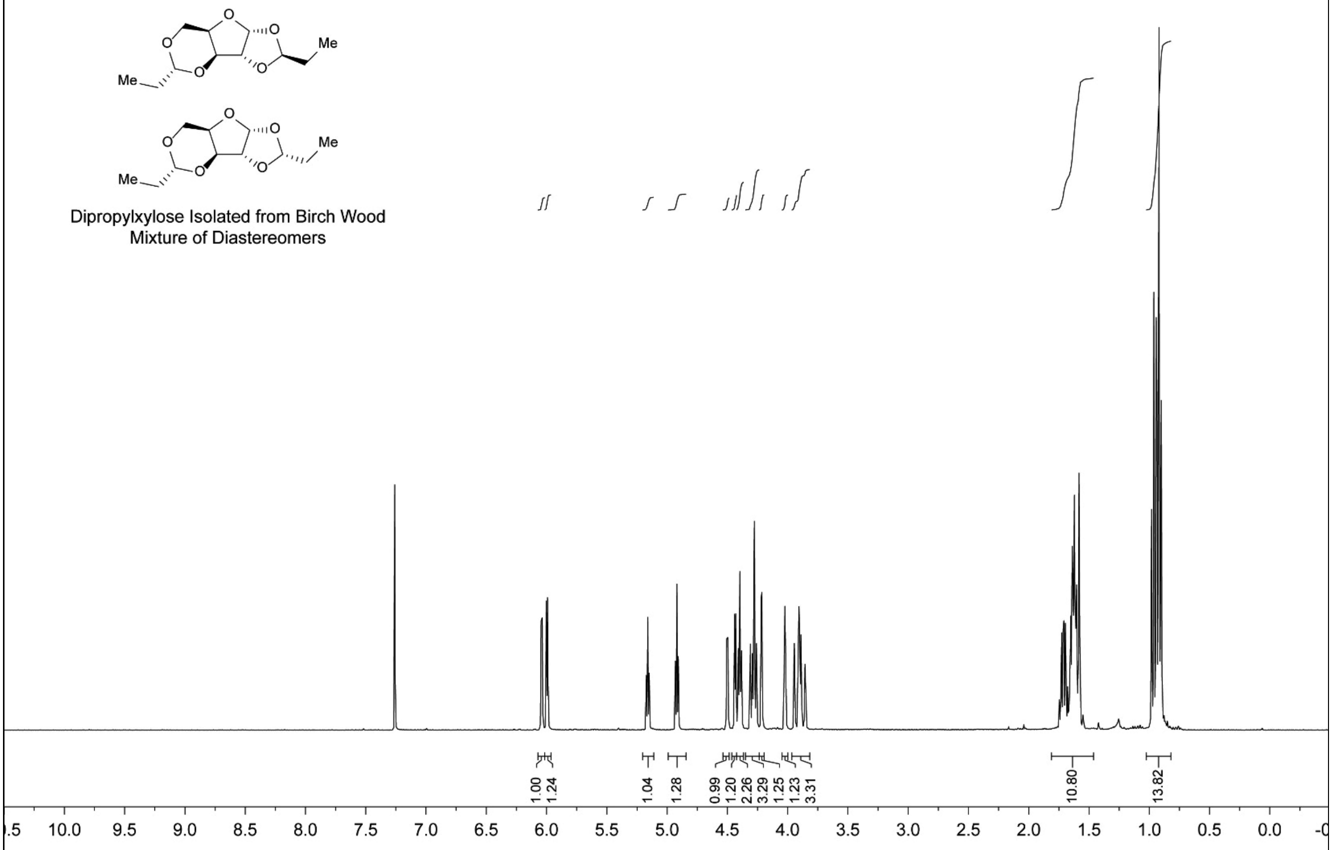


Figure B13

1H-NMR in CDCl3 of Dipropylxylose (Mixture of Isomers) Isolated from the 2018 Birch Wood Using the Propionaldehyde Fractionation Protocol.

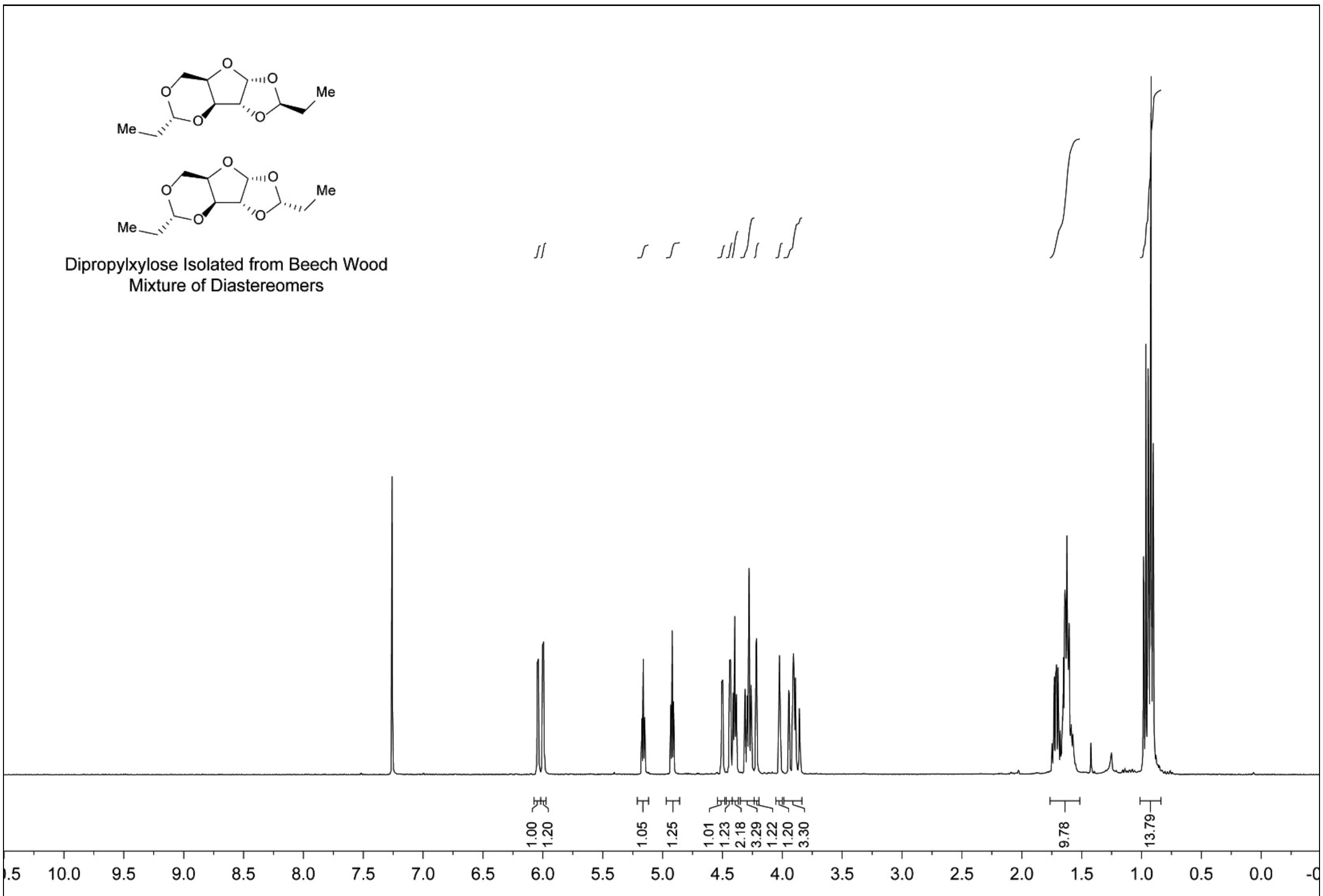


Figure B14

1H-NMR in CDCl3 of Dipropylxylose (Mixture of Isomers) Isolated from the 2018 Beech Wood Using the Propionaldehyde Fractionation Protocol.

Appendix C Appendix for chapter 4

S1 Chemicals and Materials

All commercial chemicals were analytical reagents, and were used without further purification. 5% Ru on Carbon, decane (>99%), sodium bicarbonate, propionaldehyde (≥98%), and dimethyl sulfoxide-d6 (99.9 atom % D) were purchased from Sigma Aldrich. Methanol (>99%) and 1,4-dioxane (99%) were purchased from ABCR GmbH. Fuming hydrochloric acid (37 %) was purchased from VWR. The formaldehyde solution (37%) was purchased from Roth AG.

S2 Experimental Methods

S2.0 Extraction and drying of biomass

For lignin isolation of lignin, the biomass is either used as is, without removing the extractives, or in some cases the extractives are removed, after which the biomass is further dried. In cases with extraction, the biomass is washed with an 80 wt% Ethanol aqueous solution three times and dried in the vacuum oven at 60 °C and 50 mbar for at least 16 hours. Biomass extraction is mentioned in the following section detailing each sample preparation and also listed in Table S2.

S2.1 Lignin isolation

S2.1.1 Formaldehyde-stabilized lignin (SF1-SF5)

The isolation of formaldehyde-stabilized lignin was according to 2 different approaches to provide a range of ether linkage content lignin.

For the preparation of the samples SF1, SF2, and SF5, in a 60 ml glass reactor, 1 g of biomass was mixed with 9 ml of 1,4-dioxane, 1 ml of formaldehyde solution (37 wt%), and 420 µl of HCl solution (37 wt%). The reaction was conducted in an oil bath set at 95 °C and stirred with a stir bar set at 300 rpm for 3.5 hours. After the reaction, the slurry was filtered and washed with 13.5 ml of 1,4-dioxane to separate cellulose. The filtrate was neutralized by addition of a saturated $NaHCO_3$ solution (5 ml). The solvent was then evaporated in rotavap at 40 °C and 60 mbar pressure. The lignin was then precipitated by adding 25 ml of Milli-Q water and stirring at room temperature for 1 hour in the case of SF1 and SF2, and 30 minutes for SF5. The precipitated lignin was recovered by filtration and dried overnight in a desiccator under vacuum.

For the preparation of samples SF3 and SF4, 4.5 g of extracted and dried biomass were mixed with 25 ml of 1,4-dioxane, 5.2 ml of formaldehyde solution (37 wt%), and 2.1 ml of HCl solution (37 wt%) in a 60 ml glass reactor. The reaction was conducted in an oil bath set at 95 °C and stirred with a stir bar at 300 rpm for 3.5 hours. The reaction solution was swirled by hand every 30 minutes to ensure the homogeneity. After the reaction, the slurry was filtered and washed with 1,4-dioxane and methanol to separate the cellulose. The filtrate was neutralized by addition of a saturated $NaHCO_3$ solution (35 ml) and precipitated by solvent removal in a rotary evaporator set at 35 °C and 60 mbar. The precipitated lignin was then separated by filtration and dried overnight in a desiccator under vacuum. The sources of biomass for samples SF3 and SF4 were wild birch and beech respectively.

S2.1.2 Propionaldehyde stabilized lignin (SP6 and SP7)

In a 100 ml round-bottom flask, 4.5 g of extracted and dried biomass was mixed with 25 ml of 1,4-dioxane, 4.8 ml of propionaldehyde, and 0.85 ml of HCl solution (37 wt%). A reflux condenser was used on top of the round-bottom flask to reflux the propionaldehyde. The reaction was conducted in an oil bath set at 85 °C and stirred with a stir bar set at 300 rpm for 3 hours. After the reaction, the slurry was filtered and washed with 1,4-dioxane and methanol to remove the cellulose. The filtrate was neutralized by addition of $NaHCO_3$ (1.680 g) and stirred for 30 minutes. Afterwards, the excess amount of $NaHCO_3$ and NaCl was removed by filtration followed by washing with 1,4-Dioxane. The filtrate was then concentrated using a rotavap at 40 °C and 25 mbar final pressure resulting in a dark brown oil. 10 ml of ethyl acetate were added to the oil and the lignin was precipitated by drop-wise addition of this solution to hexanes and separated by filtration. The precipitated lignin was then washed with diethyl ether for full sugar removal and dried using a rotavap at 40 ^{o}C at 25 mbar. The sources of biomass for samples SP6 and SP7 were beech and birch, respectively.

S2.1.3 Mild dilute acid-catalyzed lignin extraction (SA8 and SA9)

5.0 g of biomass and 120 mL of dioxane/H_2O (9:1, v/v) containing 0.2 mol/L HCl were added to a round bottom flask. The mixture was heated to reflux under vigorous stirring for 45 (SA8) min or 150 (SA9) min. Once cooled down to room temperature, the mixture was vacuum-filtered and the filter cake was washed with 50 mL dioxane/H_2O (9:1, v/v). The filtrate was collected and the pH valued was adjusted to 3-4 with a saturated $NaHCO_3$ solution. The solution was then concentrated at 40 °C under reduced pressure to about 25 mL. The dark brown oil was added slowly to cold water to precipitate the lignin. The resulting powder was collected by centrifugation, washed with H_2O, and dried in a vacuum oven set at 50 °C for 24 h.

S2.1.4 Non-stabilized lignin (SU10)

In a 60 ml glass reactor, 4.5 g of extracted and dried biomass was mixed with 25 ml of 1,4-dioxane, 3.3 ml of MiliQ water, and 2.1 ml of HCl solution (37 wt%). The reaction was conducted in an oil bath set at 95 ^{o}C and stirred with a stir bar at 300 rpm for 3.5 hours. The reaction solution was swirled by hand every 30 minutes to ensure the homogeneity. After the reaction, the slurry was filtered and washed with 1,4-dioxane and methanol to separate the cellulose. The lignin was precipitated by solvent removal in a rotary evaporator at 35 ºC and 60 mbar. The precipitated lignin was then separated by filtration and overnight in a desiccator under vacuum.

S2.2 Hydrogenolysis of lignin into lignin monomers

200 mg of isolated lignin was added to a 50-mL high-pressure Parr reactor along with 100 mg of catalyst (5 wt% Ru/C) and 20 ml of solvent. In the case of FA, PA, and unstabilized lignin the solvent was 1,4-Dioxane and, in the case of MDAC lignin, the solvent was methanol.

The reactor was stirred with a magnetic stir bar and heated with high-temperature heating tape (Omega) connected to a variable power supply controlled by a PID temperature controller (Omega) with a K-type thermocouple that measured the reaction temperature through a thermowell. Once closed, the reactor was purged three times and then pressurized with 40 bar of H_2. The reactor was heated to the desired temperature and then held at that temperature for the specified residence time. After reaction, the reactor was cooled with an external flow of compressed air to room temperature.

S3 Analytical Methods

S3.1 Lignin monomers yield analysis by GC

After the hydrogenolysis reaction, the reactor was cooled down to room temperature. 200 µl of internal standard solution (2 g of decane in 50 ml 1,4-Dioxane) was added to the reaction mixture and stirred with a spatula for homogeneity. The mixture was then filtered through a syringe filter. A sample of the filtrate was used for analysis with a GC (Agilent 7890B series) equipped with an HP5-column and a flame ionization detector (FID). The injection temperature was 300 °C. The column temperature program was: 40 °C (3 min), 30 °C/min to 100 °C, 40 °C/min to 300 °C and 300 °C (5 min). The detection temperature was 300 °C. The monomer yield was calculated based on the area of the monomer and the area of decane in the GC chromatogram as previously reported.[1] The detailed calculation is as follows:

$$\mathrm{n_{decane}} = \frac{\mathrm{W_{decane\ in\ sample}}}{\mathrm{MW_{decane}}} \qquad \text{(Equation S1)}$$

$$\mathrm{n_{monomer}} = \frac{\mathrm{A_{monomer\ in\ sample}}}{\mathrm{A_{decane\ in\ sample}}} \times \mathrm{n_{decane}} \times \frac{\mathrm{ECN_{decane}}}{\mathrm{ECN_{monomer}}} \qquad \text{(Equation S2)}$$

In the equations,

$W_{decane\ \mathrm{in\ sample}}$ (mg): the weight of decane used as an internal standard in each analyzed sample;

MW_{decane} (mg mmol^{-1}): the molecular weight of decane (142 mg mmol^{-1});

n_{decane} (mmol): the molar amount of decane in each analyzed sample;

$n_{monomer}$ (mmol): the molar amount of monomer in each analyzed sample;

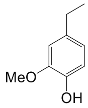
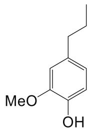
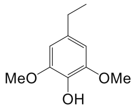
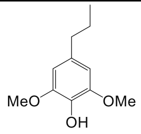
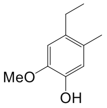
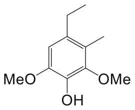
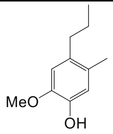
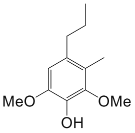
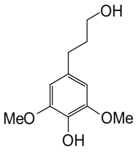
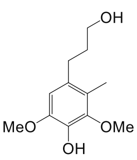
$A_{monomer\ in\ sample}$: the peak area of monomer in the GC-FID chromatogram;

$A_{decane\ in\ sample}$: the peak area of decane in the GC-FID chromatogram;

ECN_{decane}: the effective carbon number (10) of decane;

$ECN_{monomer}$: the effective carbon number of the lignin monomer molecule (Table S1);

Table S1 Effective carbon number (ECN) for lignin monomers

Lignin monomer structure	Effective carbon number calculated based on adjusted ECN rule ($\text{ECN}_{\text{monomer}}$)
	7
	8
	7
	8
	8
	8
	9
	9
	7.4
	8.4

S3.2 Lignin monomers yield prediction by 2D-HSQC$_0$ NMR

S3.2.1 Sample preparation

Approximately 50 mg sample of isolated lignin was added to NMR tube and 0.7 ml of DMSO-d was added to dissolve the lignin by using a sonication bath and vortex mixer. After making sure that the lignin was fully dissolved, a known amount of the Tetramethylsilane (TMS) was added as an internal standard to the NMR tube. The vortex mixer was used to dissolve the TMS in the solution. NMR spectra were acquired on a Bruker Avance III 500 MHz spectrometer. Table S2 presents the detailed information of the prepared samples.

Table S2 Detailed prepration of the samples

Sample	Biomass source	Extraction and drying	Isolation method	Lignin sample (mg)	TMS (mg)	Solvent (ml)
SF1	Birch	No	FA	50.0	2.6	0.7
SF2	Birch	No	FA	58.9	2.2	0.7
SF3	Birch	Yes	FA	21.5	4.1	0.7
SF4	Beech	Yes	FA	51.1	2.5	0.7
SF5	Birch	No	FA	50.6	5.3	0.7
SP6	Beech	Yes	PA	47.8	6.0	0.7
SP7	Birch	Yes	PA	53.3	7.6	0.7
SA8	Birch	Yes	MDAC	53.7	3.2	0.7
SA9	Birch	Yes	MDAC	52.6	3.2	0.7
SU10	Birch	Yes	Untabilized	51.3	2.9	0.7

S3.2.2 Pulse sequence

The main focus in gradient-selective HSQC is to ignore the signal attenuation during the coherent transfer. The detailed explanation of the pulse sequence was reported previously.[2] In this work, the strengths of the pulsed field gradients applied along the z-axis are: g_1: 80%, g_2: 20.1%, g_3: 60%, g_4: 15.075%, g_5: 40%, g_6: 10.05% of the maximum of 53 G/cm, each with a duration of 1 ms followed by a 200 μs gradient recovery period.[2]

S3.2.3 Calculation of the moles of chemical groups in the sample

Principally, this method involved performing consecutive 2D HSQC NMR data acquisitions on the same sample, which is shown as $\mathrm{HSQC_i}$ (i = 1, 2, 3 …). We performed three data acquisitions for each sample in this study. Each acquisition runs right after the next, but with a set relaxation time that was longer. This produces peak intensities for each chemical group at different relaxation times. By extrapolating the absolute peak intensities and their corresponding sequence (i), it was possible to predict the peak intensities at time zero with the equation below:

$$\ln(\mathrm{V_{i,x}}) = \ln(2\mathrm{V_{0,x}}) + \mathrm{i} * \ln(\mathrm{f_{A,x}} * \frac{1}{2}) \qquad \text{(Equation S3)}$$

Where i is the number of the corresponding sequence which is proportional to the relaxation time, $V_{i,x}$ is the absolute intensity (volume) of peak x in the corresponding spectrum i, $f_{A,x}$ is the amplitude attenuation factor specific for peak x, and $V_{0,x}$ is the intensity corresponding to chemical group x in the extrapolated signal. $V_{i,x}$ is obtained by integration of the peak volume in TopSpin 3.5. $V_{0,x}$ is directly proportional to sample concentrations due notably to the elimination of relaxation time effects. By adding a known number of moles of an internal standard to the sample, the number of moles of each chemical group can be calculated with Equation S4.

$$\mathrm{n_x} = \frac{\mathrm{V_{0,x}}}{\mathrm{C - H_x}} * \frac{\mathrm{n_{IS}} * \mathrm{C - H_{IS}}}{\mathrm{V_{0,IS}}} \qquad \text{(Equation S4)}$$

Where n_x is the predicted moles of the chemical group in sample, $V_{0,x}$ is the intensity of the chemical group x in the sample, $\mathrm{C - H_x}$ is the number of effective C—H bonds of the chemical group of interest, n_{IS} is the known number of moles of internal standard in the sample, $\mathrm{C - H_{IS}}$ is the number of C—H bonds of internal standard (12 for TMS), and $\mathrm{V_{0,IS}}$ is the intensity of the internal standard.

S3.2.4 Prediction of monomer yields and distribution

The gradient-selective HSQC$_0$ was used to quantify the absolute number of moles of protected, unprotected and C—C linkages. Knowing that only protected and unprotected ether bonds can be cleaved by hydrogenolysis, the total number of moles of the protected and unprotected lignin can be added up to estimate the maximum moles of monomers that can be produced by hydrogenolysis by assuming that each ether linkage will result in one monomer produced (Equation S5).

$$\mathrm{n}_{\text{Total Monomers}} = \mathrm{n}_{\text{Protected}} + \mathrm{n}_{\text{Unprotected}}$$ (Equation S5)

As expected from lignin’s structure, the predicted number of moles of α, β, and γ in their protected form and the moles of protection group should result in the same number as all of these atoms are part of the same chemical functionality. Therefore, the amount of protected lignin is calculated based on the average of these chemical groups:

$$\mathrm{n}_{\text{Protected Lignin}} = \text{Average}(\mathrm{n}_{\text{protected},\alpha}, \mathrm{n}_{\text{protected},\beta}, \mathrm{n}_{\text{protected},\gamma}, \mathrm{n}_{\text{protection group}})$$ (Equation S6)

The same logic applies for calculating the number of moles of the unprotected linkages. However, in the case where no unprotected form of α and/or β peaks were observed in the spectrum, the unprotected portion was assumed to be zero.

$$\mathrm{n}_{\text{Unrotected Lignin}} = \text{Average}(\mathrm{n}_{\text{unprotected},\alpha}, \mathrm{n}_{\text{unprotected},\beta}, \mathrm{n}_{\text{unprotected},\gamma})$$ (Equation S7)

These measurements could also be used to predict the distribution of monomers with various functionalities by quantification of the syringyl and guaiacyl units (along with the amount of their hydroxymethylated form in the case of formaldehyde-stabilized lignin). However, the assigned peaks for syringyl and guaiacyl units (G2 and S2/6 in Figure 2) represent all of these units whether they are in the form of protected, unprotected or linked to lignin by C—C linkages. However, we assume that there is a uniform distribution of these groups across these different units, regardless of their linkages, which is not necessarily true, especially given that G-units may have a greater propensity to condense due to their open positions on the aromatic ring. Nevertheless, this leads to accurate predictions, as discussed in the main manuscript. As a result, the amount of syringyl and guaiacyl units are calculated indirectly. First, we calculate the ratio of the total guaiacyl to syringyl in the lignin named as G/S:

$$\mathrm{G/S} = \frac{n_{\text{Total G units}}}{n_{\text{Total S units}}} \qquad \text{(Equation S8)}$$

We then use this ratio to calculate the number of each type of monomer:

$$n_{\text{syringyl monomers}} = \frac{n_{\text{Total Monomers}}}{1 + \mathrm{G/S}} \qquad \text{(Equation S9)}$$

$$n_{\text{guaiacyl monomers}} = n_{\text{Total Monomers}} - n_{\text{monomers from syringyl}} \qquad \text{(Equation S10)}$$

As a general rule for the integration of the peaks, the integration boundary is defined by auto integration of the software based on the lowest contour level that does not cause overlapping with neighboring peaks, especially in the case of aldehyde-stabilized lignin where traces of sugars can be found in the spectrum. Because the tale of the neighboring peaks can create big errors in the integration, the oblique integration mode can be used as a guide for setting the contour level to avoid the interference of the tale of neighboring peaks. For example in the case of aldehyde-stabilized lignin, traces of sugars in the isolated lignin interfered with one of the peaks from protected γ. Therefore, only one of the C—H bonds was used for integration of the peak corresponding to the Cγ as explicitly depicted on Figure 3 by the label “Signal used for integration”.

S3.2.5 Eliminating the error from t_1 noise

The t_1 noise is a ridge of noise around large peaks that is parallel to the F_1 axis in two-dimensional spectra. Figure S1 shows the t_1 noise of internal standard (TMS) along the F_1 axis.

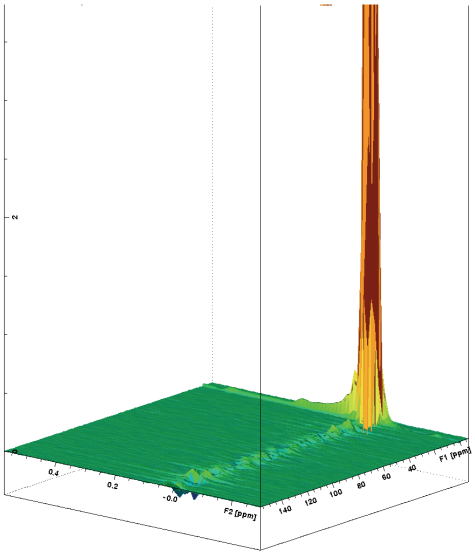


Figure S1 The t_1 noise caused by internal standard (TMS)

The other peaks in the spectra, which were significantly large to create significant noise, included the solvent (DMSO) and the methoxy group peaks from the lignin structure (Figure S2). Even though the intensity of these two peaks are lower than that of the TMS, they are visible in the spectrum.

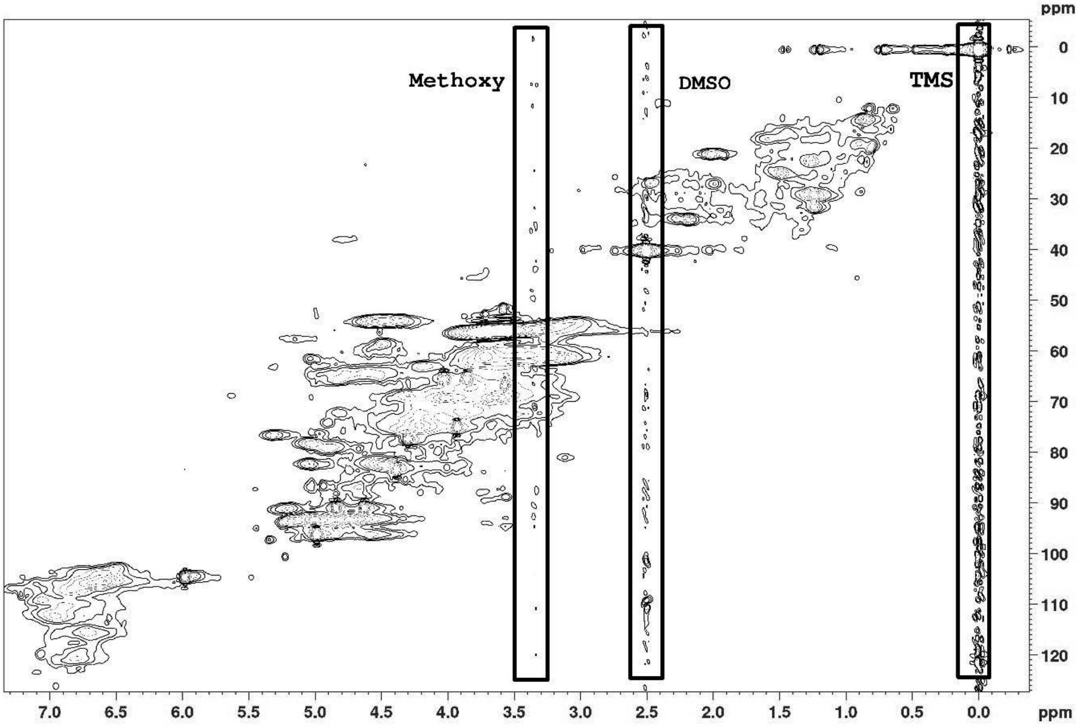


Figure S2 The main sources of t_1 noise in lignin samples

Importantly, the t_1 noise from the methoxy group interferes with one of the C—H bonds of γ carbon. This interference is a source of error in integration of the peak intensity of this chemical group, and for this reason (and as previously mentioned), the integration of the γ position is done only with the peak, for which no overlapping with the t_1 noise occurs (see Figure 2, peak marked with "Signal used for integration").

S3.2.6 Monomer yield prediction for formaldehyde-stabilized lignin

Considering the negligible amount of unprotected lignin, the predicted total amount of monomers was calculated directly from the amount of protected β—O—4 linkages:

$$n_{\text{Total Monomers}} = \text{Average}(n_{MA_\alpha}, n_{MA_\beta}, n_{MA_\gamma}, n_{MA_1}) \quad \text{(Equation S11)}$$

The peak that is used for calculation of MA_{γ} is marked with a hashed circle in Figure 3 labeled "Signal used for integration". The prediction of the monomer distribution in the case of formaldehyde-stabilized was not straight forward due to the overlap of the peaks for syringyl units with its hydroxymethylated form (S2/6 and S2 in Figure 3b). The different number of effective C—H bonds for these peaks creates more complexity as there is 1 C—H bond for S2 and 2 for S2/6. However, the peak of the guaiacyl unit does not overlap with its hydroxymethlated structure and they both have the same effective C—H bonds number that equals 1. On the other hand, the total amount of hydroxymethylated units (syringyl and guaiacyl) can be predicted by integration of the HM peak HM (see Figure 3b). Therefore, we first calculated the amount of hydroxymethylated (HM) units:

$$\mathrm{n}_{\mathrm{HM\ units}} = \mathrm{Eq.\ S3}\{\mathrm{V}_{\mathrm{i,x}} = \mathrm{V}_{\mathrm{HM}}\}, \mathrm{Eq.\ S4}\{\mathrm{C} - \mathrm{H}_{\mathrm{x}} = 1\} \qquad \text{(Equation S12)}$$

And the number of hydroxymethylated guaiacyl units were calculated by integration of the G2 peak (marked in green in Figure 3b):

$$\mathrm{n}_{\mathrm{HM\ Guaiacyl}} = \mathrm{Eq.\ S3}\{\mathrm{V}_{\mathrm{i,x}} = \mathrm{V}_{\mathrm{G2(green)}}\}, \mathrm{Eq.\ S4}\{\mathrm{C} - \mathrm{H}_{\mathrm{x}} = 1\} \qquad \text{(Equation S13)}$$

And finally the amount of hydroxymethylated syringyl were calculated as shown below:

$$\mathrm{n}_{\mathrm{HM\ Syringyl}} = \mathrm{n}_{\mathrm{HM\ units}} - \mathrm{n}_{\mathrm{HM\ Guaiacyl}} \qquad \text{(Equation S14)}$$

The total amount of guaiacyl units in the lignin structure were calculated as shown below:

$$\mathrm{n}_{\mathrm{Total\ Guaiacyl}} = \mathrm{Eq.\ S3}\{\mathrm{V}_{\mathrm{i,x}} = \mathrm{V}_{\mathrm{G2(green)}} + \mathrm{V}_{\mathrm{G2(blue)}}\}, \mathrm{Eq.\ S4}\{\mathrm{C} - \mathrm{H}_{\mathrm{x}} = 1\} \qquad \text{(Equation S15)}$$

To calculate the total amount of syringyl units, we first integrated the peak of syringyl units and its hydroxymethylated structure together:

$$n_{S2/6,S2} = Eq.\,S3\{V_{i,x} = V_{S2/6}\;\&\;V_{S2}\}, Eq.\,S4\{C - H_x = 1\} \qquad \text{(Equation S16)}$$

Then the total number of syringyl units was calculated by subtracting the number of hydroxymethylated syringyl units:

$$n_{\text{Total Syringyl}} = \frac{(n_{S2/6,S2} - n_{\text{HM Syringyl}})}{2} + n_{\text{HM Syringyl}} \qquad \text{(Equation S17)}$$

The G/S ratio was then calculated as described below:

$$\frac{G}{S} = \frac{n_{\text{Total Guaiacyl}}}{n_{\text{Total Syringyl}}} \qquad \text{(Equation S18)}$$

The amount of syringyl and guaiacyl units were then calculated as previously described, based on the equations S9 and S10, respectively.

Table S4 Number of effective C—H bonds for integration of peaks for formaldehyde-stabilized lignin

Peak annotation*	Effective C—H bonds
TMS	12
HM	2
MA_γ	1
MA_α	1
MA_β	1
MA_1	2
S2/6	2
G2 (Green)	1
G2 (Blue)	1

* Based on Figure 32b

S3.2.6 Monomer yield prediction for propionaldehyde-stabilized lignin

In the case of propionaldehyde-stabilized lignin, there were three peaks that were identified as part of the protection group which are shown as PA_1, PA_2, and PA_3 (Figure 3). However, PA_2 and PA_3 are not propotional to the quantity of protection group. Similar functionality in the structure of the lignin overlaps with these peaks, which can cause error in the integration (this mainly the case for PA_2). Therefore, the calculation of the number of protected β—O—4 linkages ignores these two signals and uses:

$$n_{\mathrm{Protected}} = \mathrm{Average}(n_{\mathrm{PA}_\alpha}, n_{\mathrm{PA}_\beta}, n_{\mathrm{PA}_\gamma}, n_{\mathrm{PA}_1}) \qquad \text{(Equation S19)}$$

The number of unprotected β—O—4 linkages was calculated with:

$$n_{\mathrm{Unprotected}} = \mathrm{Average}(n_{\mathrm{A}_\alpha}, n_{\mathrm{A}_\beta}, n_{\mathrm{A}_\gamma}) \qquad \text{(Equation S20)}$$

And the number of resinol linkages (C—C linked) was calculated with:

$$n_{\mathrm{Resinol}} = \mathrm{Average}(n_{\mathrm{C}_\alpha}, n_{\mathrm{C}_\beta}, n_{\mathrm{C}_\gamma}) \qquad \text{(Equation S21)}$$

The total number of monomers was then calculated using the equation S5. The total number of syringyl and guaiacyl units was calculated based on the integration of the peaks S2/6 and G2 in Figure 3c. The monomer distribution was then predicted based on equations S8 to S10.

Table S3 Number of effective C—H bonds for integration of peaks for propionaldehyde-stabilized lignin

Peak annotation*	Effective C—H bonds
TMS	12
C_γ	2
C_α	1
C_β	1
A_γ	1
A_α	1
A_β	1
PA_γ	1
PA_α	1
PA_β	1
PA_1	1
S2/6	2
G2	1

* Based on Figure 32c

S3.2.6 Monomer yield prediction for mild dilute-acid catalyzed and non-stabilized lignin

Taking into account that there is no protected linkages in these samples, the total number of monomers was calculated based on the number of unprotected β—O—4 linkages:

$$n_{\text{Total Monomers}} = n_{\text{unprotected}} = \text{Average}(n_{A_\alpha}, n_{A_\beta}, n_{A_\gamma}) \qquad \text{(Equation S21)}$$

The amount of resinol linkages (C—C linked) and the monomer distribution were calculated as described in previous section for propionaldehyde-protected lignin.

Table S4 Number of effective C—H bonds for integration of peaks for MDAC and non-stabilized lignin

Peak annotation*	Effective C—H bonds
TMS	12
C_γ	2
C_α	1
C_β	1
A_γ	1
A_α	1
A_β	1
S2/6	2
G2	1

* Based on Figure 32a and 32d

References

1. Annual Statistical Bulletin 2005. (2007).
2. Daniel A Lashof & R. Ahuja Dilip. Relative contributions of greenhouse gas emissions to global warming. *Nature* **344**, 529 (1990).
3. Intergovernmental Panel on Climate Change. *Climate Change 2014: Mitigation of Climate Change*. **3**, (Cambridge University Press, 2015).
4. Edenhofer, Ottmar *et al.* IPCC special report on renewable energy sources and climate change mitigation. (2011).
5. Global, B.P. Statistical Review of World Energy, June 2016. (2016).
6. Tester, J. W., Drake, E. M., Driscoll, M. J., Golay, M. W. & Peters, W. A. *Sustainable Energy: Choosing Among Options*. (MIT Press, 2012).
7. Chuang, C.-H. M., Brown, P. R., Bulović, V. & Bawendi, M. G. Improved performance and stability in quantum dot solar cells through band alignment engineering. *Nat. Mater.* **13**, 796–801 (2014).
8. Saliba, M. *et al.* Cesium-containing triple cation perovskite solar cells: improved stability, reproducibility and high efficiency. *Energy Environ. Sci.* **9**, 1989–1997 (2016).
9. Rong, Y., Liu, L., Mei, A., Li, X. & Han, H. Beyond Efficiency: the Challenge of Stability in Mesoscopic Perovskite Solar Cells. *Adv. Energy Mater.* **5**, 1501066 (2015).
10. Lewis, N. S. Toward Cost-Effective Solar Energy Use. *Science* **315**, 798–801 (2007).
11. Lewis, N. S. & Nocera, D. G. Powering the planet: Chemical challenges in solar energy utilization. *Proc. Natl. Acad. Sci.* **103**, 15729–15735 (2006).
12. Morton, O. Solar energy: A new day dawning?: Silicon Valley sunrise. *Nature* **443**, 19–22 (2006).

13. Hammer, A. *et al.* Solar energy assessment using remote sensing technologies. *Remote Sens. Environ.* **86**, 423–432 (2003).
14. Herzog, Antonia V., Timothy E. Lipman & Daniel M. Kammen. Renewable energy sources. *Univ. Calif. Berkeley USA* (2001).
15. Long, H., Li, X., Wang, H. & Jia, J. Biomass resources and their bioenergy potential estimation: A review. *Renew. Sustain. Energy Rev.* **26**, 344–352 (2013).
16. Archer, C. L. & Jacobson, M. Z. Evaluation of global wind power. *J. Geophys. Res. Atmospheres* **110**, (2005).
17. Bartle, A. Hydropower potential and development activities. *Energy Policy* **30**, 1231–1239 (2002).
18. Lund, J. W. & Freeston, D. H. World-wide direct uses of geothermal energy 2000. *Geothermics* **30**, 29–68 (2001).
19. Klass, D. L. *Biomass for Renewable Energy, Fuels, and Chemicals*. (Elsevier, 1998).
20. Jacobson, M. Z. & Delucchi, M. A. Providing all global energy with wind, water, and solar power, Part I: Technologies, energy resources, quantities and areas of infrastructure, and materials. *Energy Policy* **39**, 1154–1169 (2011).
21. Corporan, E. *et al.* Chemical, Thermal Stability, Seal Swell, and Emissions Studies of Alternative Jet Fuels. *Energy Fuels* **25**, 955–966 (2011).
22. Aresta, M., Dibenedetto, A. & Angelini, A. Catalysis for the Valorization of Exhaust Carbon: from CO2 to Chemicals, Materials, and Fuels. Technological Use of CO2. *Chem. Rev.* **114**, 1709–1742 (2014).
23. Huang Chih-Hung & Chung-Sung Tan. A review: CO2 utilization. *Aerosol Air Qual. Res* **14**, 480–499 (2014).

24. Song, C. Global challenges and strategies for control, conversion and utilization of CO2 for sustainable development involving energy, catalysis, adsorption and chemical processing. *Catal. Today* **115**, 2–32 (2006).
25. Sims, R. E. H., Hastings, A., Schlamadinger, B., Taylor, G. & Smith, P. Energy crops: current status and future prospects. *Glob. Change Biol.* **12**, 2054–2076 (2006).
26. Chakar, F. S. & Ragauskas, A. J. Review of current and future softwood kraft lignin process chemistry. *Ind. Crops Prod.* **20**, 131–141 (2004).
27. A. Sheldon, R. Green and sustainable manufacture of chemicals from biomass: state of the art. *Green Chem.* **16**, 950–963 (2014).
28. Vassilev, S. V., Baxter, D., Andersen, L. K. & Vassileva, C. G. An overview of the chemical composition of biomass. *Fuel* **89**, 913–933 (2010).
29. Jensen, C. U., Rodriguez Guerrero, J. K., Karatzos, S., Olofsson, G. & Iversen, S. B. Fundamentals of Hydrofaction$^{\mathrm{TM}}$: Renewable crude oil from woody biomass. *Biomass Convers. Biorefinery* **7**, 495–509 (2017).
30. Vassilev, S. V., Baxter, D., Andersen, L. K., Vassileva, C. G. & Morgan, T. J. An overview of the organic and inorganic phase composition of biomass. *Fuel* **94**, 1–33 (2012).
31. Schutyser, W. *et al.* Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem. Soc. Rev.* **47**, 852–908 (2018).
32. Bajpai, P. Structure of Lignocellulosic Biomass. in *Pretreatment of Lignocellulosic Biomass for Biofuel Production* (ed. Bajpai, P.) 7–12 (Springer Singapore, 2016). doi:10.1007/978-981-10-0687-6_2
33. Brown, R. M. Cellulose structure and biosynthesis: What is in store for the 21st century? *J. Polym. Sci. Part Polym. Chem.* **42**, 487–495 (2004).
34. Desvaux, M. Clostridium cellulolyticum: model organism of mesophilic cellulolytic clostridia. *FEMS Microbiol. Rev.* **29**, 741–764 (2005).

35. Klemm, D., Heublein, B., Fink, H.-P. & Bohn, A. Cellulose: Fascinating Biopolymer and Sustainable Raw Material. *Angew. Chem. Int. Ed.* **44**, 3358–3393 (2005).

36. Ciolacu, Diana, Florin Ciolacu & Valentin I. Popa. Amorphous cellulose—structure and characterization. *Cellulose chemistry and technology* **45**, 13 (2011).

37. Laureano-Perez, L., Teymouri, F., Alizadeh, H. & Dale, B. E. Understanding factors that limit enzymatic hydrolysis of biomass. *Appl. Biochem. Biotechnol.* **124**, 1081–1099 (2005).

38. H. Isikgor, F. & Remzi Becer, C. Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polym. Chem.* **6**, 4497–4559 (2015).

39. Saha, B. C. Hemicellulose bioconversion. *J. Ind. Microbiol. Biotechnol.* **30**, 279–291 (2003).

40. Hendriks, A. T. W. M. & Zeeman, G. Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresour. Technol.* **100**, 10–18 (2009).

41. Palmqvist, E. & Hahn-Hägerdal, B. Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. *Bioresour. Technol.* **74**, 25–33 (2000).

42. Chandra, R. P. *et al.* Substrate Pretreatment: The Key to Effective Enzymatic Hydrolysis of Lignocellulosics? in *Biofuels* (ed. Olsson, L.) 67–93 (Springer Berlin Heidelberg, 2007). doi:10.1007/10_2007_064

43. Crestini, C., Melone, F., Sette, M. & Saladino, R. Milled Wood Lignin: A Linear Oligomer. *Biomacromolecules* **12**, 3928–3935 (2011).

44. Grabber, J. H. How Do Lignin Composition, Structure, and Cross-Linking Affect Degradability? A Review of Cell Wall Model Studies. *Crop Sci.* **45**, 820–831 (2005).

45. Vanholme, R. *et al.* Metabolic engineering of novel lignin in biomass crops. *New Phytol.* **196**, 978–1000 (2012).

46. Ralph, J. Hydroxycinnamates in lignification. *Phytochem. Rev.* **9**, 65–83 (2010).

47. Lan, W. *et al.* Tricin, A Flavonoid Monomer in Monocot Lignification. *Plant Physiol.* pp.114.253757 (2015). doi:10.1104/pp.114.253757

48. Morreel, K. *et al.* Profiling of Oligolignols Reveals Monolignol Coupling Conditions in Lignifying Poplar Xylem. *Plant Physiol.* **136**, 3537–3549 (2004).

49. Del Río, J. C., Marques, G., Rencoret, J., Martínez, Á. T. & Gutiérrez, A. Occurrence of Naturally Acetylated Lignin Units. *J. Agric. Food Chem.* **55**, 5461–5468 (2007).

50. Parthasarathi, R., Romero, R. A., Redondo, A. & Gnanakaran, S. Theoretical Study of the Remarkably Diverse Linkages in Lignin. *J. Phys. Chem. Lett.* **2**, 2660–2666 (2011).

51. Vanholme, R., Morreel, K., Ralph, J. & Boerjan, W. Lignin engineering. *Curr. Opin. Plant Biol.* **11**, 278–285 (2008).

52. Tolbert, A., Akinosho, H., Khunsupat, R., Naskar, A. K. & Ragauskas, A. J. Characterization and analysis of the molecular weight of lignin for biorefining studies. *Biofuels Bioprod. Biorefining* **8**, 836–856 (2014).

53. Edwards, M. C. & Doran-Peterson, J. Pectin-rich biomass as feedstock for fuel ethanol production. *Appl. Microbiol. Biotechnol.* **95**, 565–575 (2012).

54. Kumar, R., Singh, S. & Singh, O. V. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *J. Ind. Microbiol. Biotechnol.* **35**, 377–391 (2008).

55. Raphael Slade, Robert Saunders, Robert Gross & Ausilio Bauen. Energy from biomass: the size of the global resource. (2011).

56. Biofuels - Energy - European Commission. *Energy* Available at: /energy/en/topics/renewable-energy/biofuels. (Accessed: 22nd October 2018)

57. Tilman, D. *et al.* Beneficial Biofuels—The Food, Energy, and Environment Trilemma. *Science* **325**, 270–271 (2009).

58. Fargione, J., Hill, J., Tilman, D., Polasky, S. & Hawthorne, P. Land Clearing and the Biofuel Carbon Debt. *Science* **319**, 1235–1238 (2008).

59. H. Khoo, H., L. Ee, W. & Isoni, V. Bio-chemicals from lignocellulose feedstock: sustainability, LCA and the green conundrum. *Green Chem.* **18**, 1912–1922 (2016).

60. Nhu Quynh Diep *et al.* Biorefinery: concepts, current status, and development trends. *Int J Biomass Renew* **2**, 1–8 (2012).
61. Berntsson, T., Sandén, B., Olsson, L. & Åsblad, A. What is a biorefinery? *Syst. Perspect. Biorefineries 2012* 16–25 (2012).
62. Cherubini, F. The biorefinery concept: Using biomass instead of oil for producing energy and chemicals. *Energy Convers. Manag.* **51**, 1412–1421 (2010).
63. Cherubini, F. & Jungmeier, G. LCA of a biorefinery concept producing bioethanol, bioenergy, and chemicals from switchgrass. *Int. J. Life Cycle Assess.* **15**, 53–66 (2010).
64. Nigam, P. S. & Singh, A. Production of liquid biofuels from renewable resources. *Prog. Energy Combust. Sci.* **37**, 52–68 (2011).
65. Taylor, G. Biofuels and the biorefinery concept. *Energy Policy* **36**, 4406–4409 (2008).
66. Kobayashi, H. & Fukuoka, A. Synthesis and utilisation of sugar compounds derived from lignocellulosic biomass. *Green Chem.* **15**, 1740–1763 (2013).
67. Werpy, T. *et al. Top Value Added Chemicals From Biomass. Volume 1 - Results of Screening for Potential Candidates From Sugars and Synthesis Gas*. (DEPARTMENT OF ENERGY WASHINGTON DC, DEPARTMENT OF ENERGY WASHINGTON DC, 2004).
68. Gallezot, P. Conversion of biomass to selected chemical products. *Chem. Soc. Rev.* **41**, 1538–1558 (2012).
69. 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 Chem.* **12**, 539–554 (2010).
70. Zviely, M. Converting Lignocellulosic Biomass to Low-Cost Fermentable Sugars. in *Pretreatment Techniques for Biofuels and Biorefineries* (ed. Fang, Z.) 133–150 (Springer Berlin Heidelberg, 2013). doi:10.1007/978-3-642-32735-3_7

71. Mosier, N. *et al.* Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour. Technol.* **96**, 673–686 (2005).
72. Ragauskas, A. J. *et al.* Lignin Valorization: Improving Lignin Processing in the Biorefinery. *Science* **344**, 1246843 (2014).
73. Chang, V. S. & Holtzapple, M. T. Fundamental Factors Affecting Biomass Enzymatic Reactivity. in *Twenty-First Symposium on Biotechnology for Fuels and Chemicals: Proceedings of the Twenty-First Symposium on Biotechnology for Fuels and Chemicals Held May 2–6, 1999, in Fort Collins, Colorado* (eds. Finkelstein, M. & Davison, B. H.) 5–37 (Humana Press, 2000). doi:10.1007/978-1-4612-1392-5_1
74. Öhgren, K., Bura, R., Saddler, J. & Zacchi, G. Effect of hemicellulose and lignin removal on enzymatic hydrolysis of steam pretreated corn stover. *Bioresour. Technol.* **98**, 2503–2510 (2007).
75. Alvira, P., Tomás-Pejó, E., Ballesteros, M. & Negro, M. J. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. *Bioresour. Technol.* **101**, 4851–4861 (2010).
76. Lynd, L. R., Weimer, P. J., Zyl, W. H. van & Pretorius, I. S. Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiol Mol Biol Rev* **66**, 506–577 (2002).
77. Studer, M. H. *et al.* Lignin content in natural Populus variants affects sugar release. *Proc. Natl. Acad. Sci.* **108**, 6300–6305 (2011).
78. Pandey, M. P. & Kim, C. S. Lignin Depolymerization and Conversion: A Review of Thermochemical Methods. *Chem. Eng. Technol.* **34**, 29–41 (2011).
79. Anwar, Z., Gulfraz, M. & Irshad, M. Agro-industrial lignocellulosic biomass a key to unlock the future bio-energy: A brief review. *J. Radiat. Res. Appl. Sci.* **7**, 163–173 (2014).
80. Sun, Y. & Cheng, J. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresour. Technol.* **83**, 1–11 (2002).

81. Taherzadeh, M., Karimi, K., Taherzadeh, M. J. & Karimi, K. Pretreatment of Lignocellulosic Wastes to Improve Ethanol and Biogas Production: A Review. *Int. J. Mol. Sci.* **9**, 1621–1651 (2008).
82. Jørgensen, H., Kristensen, J. B. & Felby, C. Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuels Bioprod. Biorefining* **1**, 119–134 (2007).
83. Sluiter, A. *et al.* Determination of structural carbohydrates and lignin in biomass. (2010).
84. S. Luterbacher, J., Alonso, D. M. & A. Dumesic, J. Targeted chemical upgrading of lignocellulosic biomass to platform molecules. *Green Chem.* **16**, 4816–4838 (2014).
85. A. Sluiter *et al.* Determination of Structural Carb ohydrates and Lignin in Biomass. (2012).
86. Lenihan, P. *et al.* Dilute acid hydrolysis of lignocellulosic biomass. *Chem. Eng. J.* **156**, 395–403 (2010).
87. Sannigrahi, P., Ho Kim, D., Jung, S. & Ragauskas, A. Pseudo- lignin and pretreatment chemistry. *Energy Environ. Sci.* **4**, 1306–1310 (2011).
88. Yan, L., Zhang, L. & Yang, B. Enhancement of total sugar and lignin yields through dissolution of poplar wood by hot water and dilute acid flowthrough pretreatment. *Biotechnol. Biofuels* **7**, 76 (2014).
89. Luterbacher, J. S. *et al.* Nonenzymatic Sugar Production from Biomass Using Biomass-Derived γ-Valerolactone. *Science* **343**, 277–280 (2014).
90. Martin Alonso, D., G. Wettstein, S., A. Mellmer, M., I. Gurbuz, E. & A. Dumesic, J. Integrated conversion of hemicellulose and cellulose from lignocellulosic biomass. *Energy Environ. Sci.* **6**, 76–80 (2013).
91. Sousa, L. da C. *et al.* Isolation and characterization of new lignin streams derived from extractive-ammonia (EA) pretreatment. *Green Chem.* **18**, 4205–4215 (2016).
92. S. Luterbacher, J. *et al.* Lignin monomer production integrated into the γ-valerolactone sugar platform. *Energy Environ. Sci.* **8**, 2657–2663 (2015).

93. Li, C., Wang, Q. & K. Zhao, Z. Acid in ionic liquid : An efficient system for hydrolysis of lignocellulose. *Green Chem.* **10**, 177–182 (2008).

94. Li, B., Filpponen, I. & Argyropoulos, D. S. Acidolysis of Wood in Ionic Liquids. *Ind. Eng. Chem. Res.* **49**, 3126–3136 (2010).

95. Kaufman Rechulski, M. D., Käldström, M., Richter, U., Schüth, F. & Rinaldi, R. Mechanocatalytic Depolymerization of Lignocellulose Performed on Hectogram and Kilogram Scales. *Ind. Eng. Chem. Res.* **54**, 4581–4592 (2015).

96. Käldström, M., Meine, N., Farès, C., Rinaldi, R. & Schüth, F. Fractionation of 'water-soluble lignocellulose' into C 5 /C 6 sugars and sulfur-free lignins. *Green Chem.* **16**, 2454–2462 (2014).

97. Calvaruso, G., T. Clough, M. & Rinaldi, R. Biphasic extraction of mechanocatalytically-depolymerized lignin from water-soluble wood and its catalytic downstream processing. *Green Chem.* **19**, 2803–2811 (2017).

98. Sannigrahi, P., Pu, Y. & Ragauskas, A. Cellulosic biorefineries—unleashing lignin opportunities. *Curr. Opin. Environ. Sustain.* **2**, 383–393 (2010).

99. Sannigrahi, P. & Ragauskas, A. J. Characterization of Fermentation Residues from the Production of Bio-Ethanol from Lignocellulosic Feedstocks. (2011). doi:info:doi/10.1166/jbmb.2011.1170

100. Pu, Y., Hu, F., Huang, F. & Ragauskas, A. J. Lignin Structural Alterations in Thermochemical Pretreatments with Limited Delignification. *BioEnergy Res.* **8**, 992–1003 (2015).

101. Vanneste, J., Ennaert, T., Vanhulsel, A. & Sels, B. Unconventional Pretreatment of Lignocellulose with Low-Temperature Plasma. *ChemSusChem* **10**, 14–31 (2017).

102. Zhang, Q. *et al.* Pretreatment of microcrystalline cellulose by ultrasounds: effect of particle size in the heterogeneously-catalyzed hydrolysis of cellulose to glucose. *Green Chem.* **15**, 963–969 (2013).

103. Boudet, A.-M. Lignins and lignification: Selected issues. *Plant Physiol. Biochem.* **38**, 81–96 (2000).

104. Liu, C., Wang, H., M. Karim, A., Sun, J. & Wang, Y. Catalytic fast pyrolysis of lignocellulosic biomass. *Chem. Soc. Rev.* **43**, 7594–7623 (2014).

105. Bridgwater, A. V. Review of fast pyrolysis of biomass and product upgrading. *Biomass Bioenergy* **38**, 68–94 (2012).

106. Wang, H., Male, J. & Wang, Y. Recent Advances in Hydrotreating of Pyrolysis Bio-Oil and Its Oxygen-Containing Model Compounds. *ACS Catal.* **3**, 1047–1070 (2013).

107. Zakzeski, J., Bruijnincx, P. C. A., Jongerius, A. L. & Weckhuysen, B. M. The Catalytic Valorization of Lignin for the Production of Renewable Chemicals. *Chem. Rev.* **110**, 3552–3599 (2010).

108. Pu, Y., Hu, F., Huang, F., Davison, B. H. & Ragauskas, A. J. Assessing the molecular structure basis for biomass recalcitrance during dilute acid and hydrothermal pretreatments. *Biotechnol. Biofuels* **6**, 15 (2013).

109. Liu, C. & Wyman, C. E. The Effect of Flow Rate of Compressed Hot Water on Xylan, Lignin, and Total Mass Removal from Corn Stover. *Ind. Eng. Chem. Res.* **42**, 5409–5416 (2003).

110. Yan, L., Ma, R., Li, L. & Fu, J. Hot Water Pretreatment of Lignocellulosic Biomass: An Effective and Environmentally Friendly Approach to Enhance Biofuel Production. *Chem. Eng. Technol.* **39**, 1759–1770 (2016).

111. Trajano, H. L. *et al.* The fate of lignin during hydrothermal pretreatment. *Biotechnol. Biofuels* **6**, 110 (2013).

112. Laskar, D. D., Zeng, J., Yan, L., Chen, S. & Yang, B. Characterization of lignin derived from water-only flowthrough pretreatment of Miscanthus. *Ind. Crops Prod.* **50**, 391–399 (2013).

113. Li, J., Henriksson, G. & Gellerstedt, G. Lignin depolymerization/repolymerization and its critical role for delignification of aspen wood by steam explosion. *Bioresour. Technol.* **98**, 3061–3068 (2007).

114. Li, M.-F., Yang, S. & Sun, R.-C. Recent advances in alcohol and organic acid fractionation of lignocellulosic biomass. *Bioresour. Technol.* **200**, 971–980 (2016).

115. Zhao, X., Cheng, K. & Liu, D. Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis. *Appl. Microbiol. Biotechnol.* **82**, 815 (2009).

116. S. Lancefield, C., Panovic, I., J. Deuss, P., Barta, K. & J. Westwood, N. Pre-treatment of lignocellulosic feedstocks using biorenewable alcohols: towards complete biomass valorisation. *Green Chem.* **19**, 202–214 (2017).

117. Pye, E. K. (Alcell T. of R. T. I. & Lora, J. H. The Alcell process: A proven alternative to kraft pulping. *Tappi J. USA* (1991).

118. Rinaldi, R. *et al.* Paving the Way for Lignin Valorisation: Recent Advances in Bioengineering, Biorefining and Catalysis. *Angew. Chem. Int. Ed.* **55**, 8164–8215 (2016).

119. Azadi, P., Inderwildi, O. R., Farnood, R. & King, D. A. Liquid fuels, hydrogen and chemicals from lignin: A critical review. *Renew. Sustain. Energy Rev.* **21**, 506–523 (2013).

120. Calvo-Flores, F. G., Dobado, J. A., Isac-García, J. & Martín-Martínez, F. J. *Lignin and Lignans as Renewable Raw Materials: Chemistry, Technology and Applications*. (John Wiley & Sons, 2015).

121. Aro, T. & Fatehi, P. Production and Application of Lignosulfonates and Sulfonated Lignin. *ChemSusChem* **10**, 1861–1877 (2017).

122. Linger, J. G. *et al.* Lignin valorization through integrated biological funneling and chemical catalysis. *Proc. Natl. Acad. Sci.* **111**, 12013–12018 (2014).

123. S. Chundawat, S. P. *et al.* Multi-scale visualization and characterization of lignocellulosic plant cell wall deconstruction during thermochemical pretreatment. *Energy Environ. Sci.* **4**, 973–984 (2011).

124. Sousa, L. da C. *et al.* Next-generation ammonia pretreatment enhances cellulosic biofuel production. *Energy Environ. Sci.* **9**, 1215–1223 (2016).

125. Kim, J. S., Lee, Y. Y. & Kim, T. H. A review on alkaline pretreatment technology for bioconversion of lignocellulosic biomass. *Bioresour. Technol.* **199**, 42–48 (2016).

126. Kim, T. H., Kim, J. S., Sunwoo, C. & Lee, Y. Y. Pretreatment of corn stover by aqueous ammonia. *Bioresour. Technol.* **90**, 39–47 (2003).

127. P. Bouxin, F. *et al.* Catalytic depolymerisation of isolated lignins to fine chemicals using a Pt/alumina catalyst: part 1—impact of the lignin structure. *Green Chem.* **17**, 1235–1242 (2015).

128. Brandt, A., Gräsvik, J., P. Hallett, J. & Welton, T. Deconstruction of lignocellulosic biomass with ionic liquids. *Green Chem.* **15**, 550–583 (2013).

129. Rencoret, J. *et al.* Isolation and structural characterization of the milled-wood lignin from Paulownia fortunei wood. *Ind. Crops Prod.* **30**, 137–143 (2009).

130. Lin, S. Y. & Dence, C. W. *Methods in Lignin Chemistry*. (Springer Science & Business Media, 2012).

131. Schutyser, W. *et al.* Influence of bio-based solvents on the catalytic reductive fractionation of birch wood. *Green Chem.* **17**, 5035–5045 (2015).

132. Galkin, M. V. & Samec, J. S. M. Lignin Valorization through Catalytic Lignocellulose Fractionation: A Fundamental Platform for the Future Biorefinery. *ChemSusChem* **9**, 1544–1558 (2016).

133. Bosch, S. V. den *et al.* Reductive lignocellulose fractionation into soluble lignin-derived phenolic monomers and dimers and processable carbohydrate pulps. *Energy Environ. Sci.* **8**, 1748–1763 (2015).

134. Renders, T., Bosch, S. V. den, Koelewijn, S.-F., Schutyser, W. & F. Sels, B. Lignin-first biomass fractionation: the advent of active stabilisation strategies. *Energy Environ. Sci.* **10**, 1551–1557 (2017).

135. Anderson, E. M. *et al.* Flowthrough Reductive Catalytic Fractionation of Biomass. *Joule* **1**, 613–622 (2017).

136. Ferrini, P., Rezende, C. A. & Rinaldi, R. Catalytic Upstream Biorefining through Hydrogen Transfer Reactions: Understanding the Process from the Pulp Perspective. *ChemSusChem* **9**, 3171–3180 (2016).

137. Bosch, S. V. den *et al.* Tuning the lignin oil OH-content with Ru and Pd catalysts during lignin hydrogenolysis on birch wood. *Chem. Commun.* **51**, 13158–13161 (2015).

138. Ferrini, P. & Rinaldi, R. Catalytic Biorefining of Plant Biomass to Non-Pyrolytic Lignin Bio-Oil and Carbohydrates through Hydrogen Transfer Reactions. *Angew. Chem.* **126**, 8778–8783 (2014).

139. Anderson, E. M. *et al.* Reductive Catalytic Fractionation of Corn Stover Lignin. *ACS Sustain. Chem. Eng.* **4**, 6940–6950 (2016).

140. Grabber, J. H., Quideau, S. & Ralph, J. p-coumaroylated syringyl units in maize lignin: Implications for β-ether cleavage by thioacidolysis. *Phytochemistry* **43**, 1189–1194 (1996).

141. Luo, H. *et al.* Total Utilization of Miscanthus Biomass, Lignin and Carbohydrates, Using Earth Abundant Nickel Catalyst. *ACS Sustain. Chem. Eng.* **4**, 2316–2322 (2016).

142. Sturgeon, M. R. *et al.* A Mechanistic Investigation of Acid-Catalyzed Cleavage of Aryl-Ether Linkages: Implications for Lignin Depolymerization in Acidic Environments. *ACS Sustain. Chem. Eng.* **2**, 472–485 (2014).

143. Deuss, P. J. *et al.* Aromatic Monomers by in Situ Conversion of Reactive Intermediates in the Acid-Catalyzed Depolymerization of Lignin. *J. Am. Chem. Soc.* **137**, 7456–7467 (2015).

144. Ek, M., Gellerstedt, G. & Henriksson, G. *Wood Chemistry and Biotechnology*. (Walter de Gruyter, 2009).

145. Wyman, C. E. *Aqueous Pretreatment of Plant Biomass for Biological and Chemical Conversion to Fuels and Chemicals*. (John Wiley & Sons, 2013).

146. Heitner, C., Dimmel, D., Schmidt, J., Dimmel, D. & Schmidt, J. *Lignin and Lignans : Advances in Chemistry*. (CRC Press, 2016). doi:10.1201/EBK1574444865

147. Saake, B. & Lehnen, R. Lignin. in *Ullmann's Encyclopedia of Industrial Chemistry* (American Cancer Society, 2007). doi:10.1002/14356007.a15_305.pub3

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

149. Gierer, J. Chemistry of delignification. *Wood Sci. Technol.* **19**, 289–312 (1985).

150. Li, C., Zhao, X., Wang, A., Huber, G. W. & Zhang, T. Catalytic Transformation of Lignin for the Production of Chemicals and Fuels. *Chem. Rev.* **115**, 11559–11624 (2015).

151. Wang, S., Dai, G., Yang, H. & Luo, Z. Lignocellulosic biomass pyrolysis mechanism: A state-of-the-art review. *Prog. Energy Combust. Sci.* **62**, 33–86 (2017).

152. Chu, S., V. Subrahmanyam, A. & W. Huber, G. The pyrolysis chemistry of a β-O-4 type oligomeric lignin model compound. *Green Chem.* **15**, 125–136 (2013).

153. Holmelid, B., Kleinert, M. & Barth, T. Reactivity and reaction pathways in thermochemical treatment of selected lignin-like model compounds under hydrogen rich conditions. *J. Anal. Appl. Pyrolysis* **98**, 37–44 (2012).

154. S. Choi, Y. *et al.* Pyrolysis reaction networks for lignin model compounds: unraveling thermal deconstruction of β-O-4 and α-O-4 compounds. *Green Chem.* **18**, 1762–1773 (2016).

155. Dellon, L. D., Yanez, A. J., Li, W., Mabon, R. & Broadbelt, L. J. Computational Generation of Lignin Libraries from Diverse Biomass Sources. *Energy Fuels* **31**, 8263–8274 (2017).

156. Lu, J., Wang, M., Zhang, X., Heyden, A. & Wang, F. β-O-4 Bond Cleavage Mechanism for Lignin Model Compounds over Pd Catalysts Identified by Combination of First-Principles Calculations and Experiments. *ACS Catal.* **6**, 5589–5598 (2016).

157. H. Parsell, T. *et al.* Cleavage and hydrodeoxygenation (HDO) of C–O bonds relevant to lignin conversion using Pd/Zn synergistic catalysis. *Chem. Sci.* **4**, 806–813 (2013).

158. Bosch, S. V. den *et al.* Integrating lignin valorization and bio-ethanol production: on the role of Ni-Al 2 O 3 catalyst pellets during lignin-first fractionation. *Green Chem.* **19**, 3313–3326 (2017).

159. Roberts, V. M. *et al.* Towards Quantitative Catalytic Lignin Depolymerization. *Chem. – Eur. J.* **17**, 5939–5948 (2011).

160. Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J. & Mansfield, S. D. Designer lignins: harnessing the plasticity of lignification. *Curr. Opin. Biotechnol.* **37**, 190–200 (2016).

161. Bonawitz, N. D. & Chapple, C. Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? *Curr. Opin. Biotechnol.* **24**, 336–343 (2013).

162. Gellerstedt, G. & Henriksson, G. Chapter 9 - Lignins: Major Sources, Structure and Properties. in *Monomers, Polymers and Composites from Renewable Resources* (eds. Belgacem, M. N. & Gandini, A.) 201–224 (Elsevier, 2008). doi:10.1016/B978-0-08-045316-3.00009-0

163. Rolando, C., Monties, B. & Lapierre, C. Thioacidolysis. in *Methods in Lignin Chemistry* (eds. Lin, S. Y. & Dence, C. W.) 334–349 (Springer Berlin Heidelberg, 1992). doi:10.1007/978-3-642-74065-7_23

164. Pinto, P. C. R., Costa, C. E. & Rodrigues, A. E. Oxidation of Lignin from Eucalyptus globulus Pulping Liquors to Produce Syringaldehyde and Vanillin. *Ind. Eng. Chem. Res.* **52**, 4421–4428 (2013).

165. J. Deuss, P. *et al.* Phenolic acetals from lignins of varying compositions via iron(iii) triflate catalysed depolymerisation. *Green Chem.* **19**, 2774–2782 (2017).

166. Constant, S. *et al.* New insights into the structure and composition of technical lignins: a comparative characterisation study. *Green Chem.* **18**, 2651–2665 (2016).

167. Cass, C. L. *et al.* Cell Wall Composition and Biomass Recalcitrance Differences Within a Genotypically Diverse Set of Brachypodium distachyon Inbred Lines. *Front. Plant Sci.* **7**, (2016).

168. Wen, J.-L., Sun, S.-L., Xue, B.-L. & Sun, R.-C. Recent Advances in Characterization of Lignin Polymer by Solution-State Nuclear Magnetic Resonance (NMR) Methodology. *Materials* **6**, 359–391 (2013).

169. Zhang, L. & Gellerstedt, G. Quantitative 2D HSQC NMR determination of polymer structures by selecting suitable internal standard references. *Magn. Reson. Chem.* **45**, 37–45 (2007).

170. Heikkinen, S., Toikka, M. M., Karhunen, P. T. & Kilpeläinen, I. A. Quantitative 2D HSQC (Q-HSQC) via Suppression of J-Dependence of Polarization Transfer in NMR Spectroscopy: Application to Wood Lignin. *J. Am. Chem. Soc.* **125**, 4362–4367 (2003).

171. Rahimi, A., Ulbrich, A., Coon, J. J. & Stahl, S. S. Formic-acid-induced depolymerization of oxidized lignin to aromatics. *Nature* **515**, 249–252 (2014).

172. Shimada, K., Hosoya, S. & Ikeda, T. Condensation Reactions of Softwood and Hardwood Lignin Model Compounds Under Organic Acid Cooking Conditions. *J. Wood Chem. Technol.* **17**, 57–72 (1997).

173. Yan, N. *et al.* Selective degradation of wood lignin over noble-metal catalysts in a two-step process. *ChemSusChem* **1**, 626–629 (2008).

174. Pepper, J. M. & Lee, Y. W. Lignin and related compounds. I. A comparative study of catalysts for lignin hydrogenolysis. *Can. J. Chem.* **47**, 723–727 (1969).

175. Wang, H., Tucker, M. & Ji, Y. Recent Development in Chemical Depolymerization of Lignin: A Review. *Journal of Applied Chemistry* (2013). doi:10.1155/2013/838645

176. Karlen, S. D. *et al.* Monolignol ferulate conjugates are naturally incorporated into plant lignins. *Sci. Adv.* **2**, e1600393 (2016).

177. Furikado, I. *et al.* Catalytic performance of Rh/SiO 2 in glycerol reaction under hydrogen. *Green Chem.* **9**, 582–588 (2007).

178. Dasari, M. A., Kiatsimkul, P.-P., Sutterlin, W. R. & Suppes, G. J. Low-pressure hydrogenolysis of glycerol to propylene glycol. *Appl. Catal. Gen.* **281**, 225–231 (2005).

179. Parsell, T. *et al.* A synergistic biorefinery based on catalytic conversion of lignin prior to cellulose starting from lignocellulosic biomass. *Green Chem.* **17**, 1492–1499 (2015).

180. Sluiter, Justin & Amie Sluiter. Summative Mass Closure: Laboratory Analytical Procedure (LAP) Review and Integration: Pretreated Slurries. (2011).

181. Galkin, M. V. & Samec, J. S. M. Selective Route to 2-Propenyl Aryls Directly from Wood by a Tandem Organosolv and Palladium-Catalysed Transfer Hydrogenolysis. *ChemSusChem* **7**, 2154–2158 (2014).

182. Phongpreecha, T. *et al.* Predicting lignin depolymerization yields from quantifiable properties using fractionated biorefinery lignins. *Green Chem* **19**, 5131–5143 (2017).

183. Chang, H., Cowling, E. B. & Brown, W. Comparative Studies on Cellulolytic Enzyme Lignin and Milled Wood Lignin of Sweetgum and Spruce. *Holzforsch. - Int. J. Biol. Chem. Phys. Technol. Wood* **29**, 153–159 (2009).

184. Hu, K., Westler, W. M. & Markley, J. L. Simultaneous Quantification and Identification of Individual Chemicals in Metabolite Mixtures by Two-Dimensional Extrapolated Time-Zero 1H−13C HSQC (HSQC0). *J. Am. Chem. Soc.* **133**, 1662–1665 (2011).

185. Hu, K., Ellinger, J. J., Chylla, R. A. & Markley, J. L. Measurement of Absolute Concentrations of Individual Compounds in Metabolite Mixtures by Gradient-Selective Time-Zero 1H–13C HSQC with Two Concentration References and Fast Maximum Likelihood Reconstruction Analysis. *Anal. Chem.* **83**, 9352–9360 (2011).

186. Shuai, L. *et al.* Formaldehyde stabilization facilitates lignin monomer production during biomass depolymerization. *Science* **354**, 329–333 (2016).

187. Lan, W., Amiri, M. T., Hunston, C. M. & Luterbacher, J. S. Protection Group Effects During α,γ-Diol Lignin Stabilization Promote High-Selectivity Monomer Production. *Angew. Chem.* **130**, 1370–1374

188. Das, A. *et al.* Lignin Conversion to Low-Molecular-Weight Aromatics via an Aerobic Oxidation-Hydrolysis Sequence: Comparison of Different Lignin Sources. *ACS Sustain. Chem. Eng.* **6**, 3367–3374 (2018).

189. Phongpreecha, T. *et al.* Predicting lignin depolymerization yields from quantifiable properties using fractionated biorefinery lignins. *Green Chem.* **19**, 5131–5143 (2017).

190. Franke, R. *et al.* Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase. *Plant J.* **22**, 223–234 (2000).

191. Kim, H. & Ralph, J. Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d6/pyridine-d5. *Org. Biomol. Chem.* **8**, 576–591 (2010).

192. Stewart, J. J., Akiyama, T., Chapple, C., Ralph, J. & Mansfield, S. D. The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiol.* **150**, 621–635 (2009).

193. TAPPI method. www.tappi.org/content/SARG/T222.pdf.

194. Scanlon, J. T. & Willis, D. E. Calculation of flame ionization detector relative response factors using the effective carbon number concept. *J. Chromatogr. Sci.* **23**, 333–340 (1985).

195. Shuai, L. & Luterbacher, J. Organic solvent effects in biomass conversion reactions. *ChemSusChem* **9**, 133–155 (2016).

196. Mellmer, M. A. *et al.* Solvent Effects in Acid-Catalyzed Biomass Conversion Reactions. *Angew. Chem. Int. Ed.* **53**, 11872–11875 (2014).

197. Shuai, L., Questell-Santiago, Y. M. & Luterbacher, J. S. A mild biomass pretreatment using γ-valerolactone for concentrated sugar production. *Green Chem.* **18**, 937–943 (2016).

198. Luterbacher, J. S. *et al.* Nonenzymatic sugar production from biomass using biomass-derived γ-valerolactone. *Science* **343**, 277–280 (2014).

199. Yuan, G., Qi, C., Wu, W. & Jiang, H. Recent advances in organic synthesis with CO2 as C1 synthon. *Curr. Opin. Green Sustain. Chem.* **3**, 22–27 (2017).

Curriculum Vitae

NAME: Masoud TALEBI AMIRI

DATE, PLACE OF BIRTH: 31.10.1989, Babol, Iran

EMAIL: masoud.t.amiri@gmail.com

EDUCATION

- **École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.**
  Ph.D. in Chemistry and Chemical Engineering (Oct. 2014- Dec. 2017)
- **École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.**
  M.Sc. in Chemical Engineering and Biotechnology (Sept. 2012- Jul. 2014)
- **Sharif University of Technology, Tehran, Iran.**
  B.Sc. in Chemical Engineering (Sept. 2008- Jun. 2014)
- **National Organization for Development of Exceptional Talents (NODET), Babol, Iran.**
  High school diploma in Mathematics and Pphysics (Sept. 2004- Jun. 2008)

PROFESSIONAL HISTORY

- **BASF, Plant Support Department – Monthey, Switzerland.**
  Internship: Sept. 2013- Feb. 2014
- **Group of Catalytic Reaction Engineering, EPFL – Lausanne, Switzerland.**
  Internship: Jul. 2013- Sept. 2013
- **Tehran Oil Refining Company, R&D Department – Tehran, Iran.**
  Internship: Jul. 2011- Oct. 2011

PUBLICATIONS

- Shuai, Li, Masoud Talebi Amiri, Ydna M. Questell-Santiago, Florent Héroguel, Yanding Li, Hoon Kim, Richard Meilan, Clint Chapple, John Ralph, and Jeremy S. Luterbacher. "Formaldehyde stabilization facilitates lignin monomer production during biomass depolymerization." Science 354, no. 6310 (2016): 329-333.

- Lan, Wu, Masoud Talebi Amiri, Christopher M. Hunston, and Jeremy S. Luterbacher. "Protection Group Effects During α, γ-Diol Lignin Stabilization Promote High-Selectivity Monomer Production." *Angewandte Chemie* 130, no. 5 (2018): 1370-1374.

- Shuai, Li, Masoud Talebi Amiri, and Jeremy S. Luterbacher. "The influence of interunit carbon–carbon linkages during lignin upgrading." *Current Opinion in Green and Sustainable Chemistry* 2 (2016): 59-63.

- Saeidi, Samrand, Masoud Talebi Amiri, Nor Aishah Saidina Amin, and Mohammad Reza Rahimpour. "Progress in reactors for high-temperature Fischer–Tropsch process: determination place of intensifier reactor perspective." *International Journal of Chemical Reactor Engineering* 12, no. 1 (2014): 639-664.

CONFERENCES

- Masoud Talebi Amiri, Li Shuai, Ydna M. Questell-Santiago, Florent Héroguel, Yanding Li, Hoon Kim, Richard Meilan, Clint Chapple, John Ralph and Jeremy S.Luterbacher. " Formaldehyde-based stabilization facilitates liquid fuel production from lignin during biomass depolymerization", Biomass for Swiss Energy Future Conference (SCCER BIOSWEET), Brugg, Switzerland, Sept. 5, 2016.

- Masoud Talebi Amiri, Li Shuai, Ydna M. Questell-Santiago and Jeremy S. Luterbacher, "Quantitative depolymerization of uncondensed technical lignin under mild condition", SCS Fall Meeting, Bern, Switzerland, Aug. 22, 2017.

BOOKS

PROFESSIONAL MEMBERSHIPS

INTERESTS

Ce document a été imprimé au Centre d’impression EPFL,
imprimerie climatiquement neutre, certifiée myClimate.