

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

**NOVEL PRETREATMENT CONTROL AND ANALYTICAL
TECHNIQUE MAXIMIZING MONOLIGNOLS YIELD FROM
WASTE LIGNOCELLULOSIC BIOMASS**

GUAN JIANYU

PhD

The Hong Kong Polytechnic University

2026

The Hong Kong Polytechnic University
Department of Civil and Environmental Engineering

**Novel Pretreatment Control And Analytical Technique
Maximizing Monolignols Yield From Waste Lignocellulosic
Biomass**

Guan Jianyu

A thesis submitted in partial fulfillment of the requirements for the degree
of Doctor of Philosophy

Aug 2025

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgment has been made in the text.

Signature: _____

Name of Student: Guan Jianyu

DEDICATION

To My parents

And

Those who have been aside and supported me through this journey.

ABSTRACT

Economic growth has been driven by fossil fuels and petroleum-based chemicals, yet their continued use has heightened global concerns over climate instability and resource depletion. Lignocellulosic biomass offers a renewable and carbon-neutral alternative, but its efficient utilization has been hindered by its recalcitrance due to the compact and cross-linked architecture of cellulose, hemicellulose, and lignin. This dissertation addresses innovative solutions to overcome these challenges; by elucidating biomass molecular structures using the state-of-the-art techniques, building novel numerical tools in predicting pretreatment performance, and establishing robust conversion processes for lignin valorization. Lignocellulosic structure with significant diversity among plant species was systematically reviewed, classifying hardwoods, softwoods, grasses, barks, and seeds, and correlating their structural features with recalcitrance. Biorefinery strategies were compared, emphasizing pretreatment approaches and their efficiencies in polysaccharide and lignin utilization. Based on comprehensive experimental studies and sample analyses, a novel index, Pretreatment Depolymerization Factor (PDF), was constructed through acid-organosolv pretreatments of *Camellia oleifera* shell. The model quantified the hydrolysis kinetics for different Xylan conformations, providing the mechanistic foundation with 91% accuracy. Finally, 2D-HSQC SS-NMR and hydrogenolysis experiments identified C γ -OH initiated C β -C γ cleavage pathway. This mechanism was incorporated into the PDF simulation to quantitatively clarify its treatment of lignin condensation. As a result, application of the PDF model in semi-continuous biomass pretreatments achieves 40.1 mg g⁻¹ monolignol yield with 95% delignification, nearly three-fold higher than baseline. Finally, ¹³C-labeled plant cell walls were analyzed using solid-state (SS) NMR method to elucidate biomass molecular architectures. Sequential weak-acid fractionation coupled with advanced SS-NMR reveals covalent lignin-xylan linkages and spatial proximities within secondary cell walls, explaining residual recalcitrance and potential modes of lignin intermolecular connectivity.

PUBLICATIONS ARISING FROM THE THESIS

Journal articles

1. **Guan J**, Khan, A, Patria, R. D, and Leu, S.-Y., et al. "Strategic Design Principles for Greener Biorefinery: A Substrate-Process Matrix Emphasizing Complete Lignocellulose Utilization over Variant Biomass." *Green Chemistry* 27.38 (2025): 11581-11606.
2. **Guan J**, TSE, H.-Y, Patria, R. D and Leu, S.-Y., et al. "Novel Fractionation-Depolymerization Control on Rapid Dissociation of Lignin-Associated Xylan: Complete Biomass Valorization Toward Carbon Neutrality." *Green Chemistry* 27.43 (2025): 13834-13848.
3. **Guan J**, Zhuang H, Lau C Y, et al. Energy and carbon footprint analysis of municipal wastewater treatment process integrated with food waste disposer[J]. *Journal of Cleaner Production*, 2022, 375: 134063.
4. Khan R J, **Guan J**, Lau C Y, et al. Monolignol Potential and Insights into Direct Depolymerization of Fruit and Nutshell Remains for High Value Sustainable Aromatics[J]. *ChemSusChem*, 2024, 17(7): e202301306.
5. Zhuang, H., **Guan, J**, Leu, S.-Y., Wang, Y., & Wang, H. (2020). Carbon footprint analysis of chemical enhanced primary treatment and sludge incineration for sewage treatment in Hong Kong. *Journal of Cleaner Production*, 272, 122630.
6. TSE, H.-Y. Yeung C S, Lau C Y, **Guan J**, et al. One-pot synthesis to prepare lignin/photoacid nanohybrids for multifunctional biosensors and photo-triggered singlet oxygen generation[J]. *Green Chemistry*, 2022,
7. Khan R J, Lau C Y, **Guan J**, et al. Recent advances of lignin valorization techniques toward sustainable aromatics and potential benchmarks to fossil refinery products[J]. *Bioresource Technology*, 2022, 346: 126419
8. Islam M K, Rehman S, **Guan J**, et al. Biphasic pretreatment for energy and carbon efficient conversion of lignocellulose into bioenergy and reactive lignin[J]. *Applied Energy*, 2021, 303: 117653.
9. Islam M K, Thaemngoen A, Lau C Y, **Guan J** et al. Staged organosolv pretreatment to increase net energy and reactive lignin yield in whole oil palm tree biorefinery[J]. *Bioresource Technology*, 2021, 326: 124766.

Manuscript to be submitted upon this dissertation

Guan J, Patria, R, Wang Tuo and Leu, S.-Y., et al. "Revealing Lignin Intermolecular and Intramolecular Interactions in Plant Cell Walls by Solid-State NMR"

ACKNOWLEDGMENTS

I am deeply grateful for the invaluable support and guidance I have received throughout my doctoral studies. I would like to express my sincere gratitude to my supervisor, **Prof. Shao-Yuan** (Ben) Leu. From my master's studies through to the completion of this Ph.D., he has given me invaluable opportunities, guidance, and trust. He provided expert advice, constant support, and positive motivation throughout my research and in preparing this dissertation. His devotion to research and strong sense of integrity have deeply influenced me, teaching me to value both knowledge and responsibility. I am profoundly thankful for his patience, encouragement, and unwavering support, and feel truly fortunate to have been supervised by him.

I am also sincerely grateful to our laboratory technicians and senior colleagues who guided me during my early days in research, including Dr. Chi Shun, Dr. Oliver Yeung, Dr. Ho-Yin (Leo) Tse, Dr. Huichuan Zhuang, Dr. Huaimin Wang, Dr. Md Khairul Islam, Dr. Leng Ling, Mr. Wai Shung Lam, Ms. Che Ying (Celine), Mr. Yuen Him (Jonathan), and Dr. Kenneth Yan. Their patience and generous help gave me the foundation to move forward with confidence.

Special thanks go to my lab mates. Many parts of this thesis were only possible because of teamwork, countless discussions, and shared efforts. The long hours in the lab, the struggles, and the small victories we celebrated together have become some of the warmest memories of my Ph.D. years. I would especially like to thank Dr. Chun Yin Lau, Mr. Raffel Dharma Patria, Mr. Aamir Khan, Dr. Rabia Jalil Khan, and Dr. Yanlin Zhang for their companionship and support. During my exchange at Michigan State University, I sincerely acknowledge the guidance and support provided by Prof. Tuo Wang and Dr. Wancheng Zhao, whose mentorship significantly broadened my academic perspective and strengthened my foundation in solid-state NMR.

Finally, I would like to dedicate this thesis to my parents. Their unconditional love, sacrifice, and belief in me have been the strongest source of strength throughout my journey. In my darkest and most difficult times, it was their support that kept me moving forward. This achievement belongs as much to them as it does to me.

Guan Jianyu

Aug 2025

Table of Contents

ABSTRACT	v
ACKNOWLEDGMENTS.....	vii
List of Abbreviation	xiii
CHAPTER 1. BACKGROUND AND THESIS OUTLINE	1
CHAPTER 2. WOODY BIOMASS AND ITS STRUCTURE.....	7
ABSTRACT	7
2.1. Chemical composition of lignocellulosic biomass	8
2.2. Structural diversity of lignin in different biomass.....	11
2.3. Molecular Architecture and Recalcitrance of Plant cell walls	14
CHAPTER 3. BIOREFINERY AND PRETREATMENT	19
ABSTRACT	19
3.1. Biorefinery strategies.....	20
3.1.1. Organosolv pretreatment	22
3.1.2. Deep eutectic solvent (DES) pretreatment	23
3.1.3. Pretreatment-Catalytic Transfer Hydrogenolysis (PT-CTH)	23
3.1.4. Reductive Catalytic Fractionation (RCF).....	25
3.2. Comparison between organosolv and DES pretreatment	25
3.3. Delignification mechanisms in PT& RCF	28
3.4. Pretreatment performance in different biomass.....	30
3.5. Delignification kinetics.....	33
3.6. Mass balance of PT-CTH and RCF	34
3.7. Suggestion for different biomass pretreatment.....	37
CHAPTER 4. METHODOLOGY.....	41
4.1. Biomass feed stock and chemicals reagents	41
4.2. Chemical composition analysis	42
4.3. Lignin analysis.....	42
4.3.1. Two-dimensional HSQC	42
4.3.2. Gel Permeation Chromatography (GPC)	43
4.3.3. Fourier-transform infrared spectra (FTIR)	43
4.3.4. Quantitative solid-state NMR (Multi-CP).....	43
4.3.5. Quantitative ³¹ P NMR spectroscopic characterization.....	43
4.4. Scanning electron microscopy (SEM).....	44
4.5. X-ray Diffraction (XRD).....	44
4.6. Monomer analysis.....	45
4.7. Solid-state NMR spectroscopy	45
4.7.1. One-dimensional direct polarization & cross-polarization ¹³ C spectra...45	
4.7.2. 2D INADEQUATE, PDSO & Gated-DARR ¹³ C Spectroscopy.....	46
4.8. PDF Modeling and calculation	47
4.8.1. Pretreatment depolymerization factor (PDF)	47
4.8.2. Hemicellulose hydrolysis	48
4.8.3. Lignin fractionation.....	48
4.8.4. Lignin depolymerization	50
CHAPTER 5. POLYSACCHARIDE HYDROLYSIS AND LIGNIN DISSOCIATION KINETICS.....	51
ABSTRACT	51
5.1. Introduction	52
5.2. Experiment Design	56
5.2.1. Biomass feedstock and chemicals	56

5.2.2. Biomass Pretreatment	56
5.2.3. Preparation of subtract sample for SS-NMR	57
5.2.4. Preparation of lignin sample	58
5.3. Results and discussion	58
5.3.1. Overview of pretreatment results	58
5.3.2. Characterization of hemicelluloses using solid-state NMR	60
5.3.3. X-ray diffraction (XRD) for the COS subtracts	63
5.3.4. Morphology of Pretreatment Feedstock (SEM)	64
5.3.5. Polysaccharide hydrolysis kinetic	66
Conclusion	69
CHAPTER 6. LIGNIN CONDENSATION AND DEPOLYMERIZATION	70
ABSTRACT	70
6.1. Introduction	71
6.2. Experiment Design	73
6.2.1. Pretreatment lignin separation	73
6.2.2. Preparation of lignin samples	73
6.2.3. Catalytic transfer hydrogenolysis (CTH) for lignin depolymerization ...	74
6.2.4. PDF-Controlled Semi-Continuous Flow Experiments	74
6.3. Results and discussion	75
6.3.1. Lignin structural changes at increasing PDF	75
6.3.2. Lignin Molecular Weight Distribution by GPC	78
6.3.3. Fourier-transform infrared spectroscopy of separated lignin	81
6.3.4. Results of catalytic transfer hydrogenolysis (CTH)	82
6.3.5. Analysis of lignin polarity changes during pretreatment	84
6.3.6. Discovery of new lignin condensation mechanisms	85
6.3.7. Application of PDF in lignin-first biorefinery	88
Conclusions	92
CHAPTER 7. REVEALING LIGNIN INTERMOLECULAR AND INTRAMOLECULAR INTERACTIONS IN PLANT CELL WALLS BY SOLID-STATE NMR	93
ABSTRACT	93
7.1. Introduction	94
7.2. Materials and method	97
7.2.1. Biomass feedstock and chemicals	97
7.2.2. Lignin dissociation experiment	97
7.2.3. Preparation of lignin samples	98
7.2.4. Solid-state NMR experiment	98
7.3. Results and discussion	99
7.3.1. Structural resolution of eucalyptus cell walls by ¹³ C SS-NMR	99
7.3.2. Mapping Lignin-Carbohydrate Interactions and Intramolecular Linkages	105
7.4. Perspective	112
CHAPTER 8. Conclusion	113
Appendix	115

List of Figure

Fig. 1-1. Key features and chapters of this dissertation	4
Fig. 2-1. Plant part classification and chemical composition	8
Fig. 2-2. Lignin model structures and unique lignin units investigated in this study..	12
Fig. 2-3. Molecular architecture and recalcitrance of plant cell walls.....	15
Fig. 3-1. Process flow of PT-CTH and RCF biorefinery.....	20
Fig. 3-2. Comparison of product distribution using PT-CTH and RCF strategie.....	24
Fig. 3-3. Schematic comparison of lignin removal mechanisms in acid pretreatment and reductive catalytic fractionation (RCF).	28
Fig. 3-4. Phase distribution of biomass components after PT and RCF.....	29
Fig. 3-5. Component separation performance of five representative biomass types (hardwood, softwood, bark, grass, and seed) under organosolv pretreatment.. ...	32
Fig. 3-6. Delignification kinetics in various pretreatment methods and RCF	33
Fig. 3-7. Sankey diagram depicting mass balance of (a) PT-CTH of hardwood and (b) RCF of hardwood. (c) Average values of initial composition of biomass (Raw), products composition after PT-CTH, and products composition after RCF	36
Fig. 3-8. Decision framework for biomass processing based on cellulose content and lignin depolymerization potential.....	37
Fig. 3-9. Two-dimensional compositional map showing the relationship between lignin depolymerization potential (x-axis) and cellulose content (y-axis) across diverse biomass feedstocks.....	39
Fig. 5-1. Conceptual diagrams of (a) a lignin-first biorefinery based on the newly proposed control algorithm, the pretreatment and depolymerization factor (PDF), for maximizing the monolignol production potential from <i>Camellia oleifera</i> shell (COS);.....	55
Fig. 5-2. Photo of the feedstock in this study	56
Fig. 5-3. Summary of biorefinery experiments.....	60
Fig. 5-4. Solid-State NMR spectra of biomass with detailed peak assignments and quantitative analysis..	62
Fig. 5-5. Differential spectra with peak deconvolution highlighting characteristic signals of xylan (Xn^{2f} , Xn^{3f}), glucan (Glc), arabinose (Arb), and lignin methoxyl groups, enabling quantitative analysis of hemicellulose conformations.	63
Fig. 5-6. XRD of a) untreated; b) 1,4-BDO/100mM H_2SO_4 , 121°C; c) 1,4- BDO/400mM H_2SO_4 , 121°C; d) ethanol/100mM H_2SO_4 , 165°C <i>Camellia</i> <i>Oleifera</i> shell; e) untreated pine.....	64
Fig. 5-7. SEM of untreated/treated <i>Camellia Oleifera</i> shell	65
Fig. 5-8. SEM of untreated/treated <i>Camellia Oleifera</i> shell	66
Fig. 5-9. Cross-feedstock validation of the Pretreatment Depolymerization Factor (PDF) model.	68
Fig. 6-1. ^1H - ^{13}C 2D-HSQC NMR spectra showing lignin interunit linkages (in aliphatic regions, top row) and lignin fraction compositions.	76
Fig. 6-2. Gel permeation chromatograms(GPC) of extracted lignin before and after hydrogenolysis over increased pretreatment severities	79
Fig. 6-3. GPC results of lignin liquor over increased pretreatment severities	80
Fig. 6-4. Lignin structural changes during pretreatment.....	81
Fig. 6-5. FT-IR results in different PDF fractionation condition.....	82
Fig. 6-6. (l) GC signals of the six major monolignols produced from biomass at different PDF	83
Fig. 6-7. Discovery of a new lignin depolymerization and condensation pathway. ...	86

Fig. 6-8. ^{13}C multi-CP NMR spectra of lignin at different PDF conditions (PDF=1.8 and 10.3).	87
Fig. 6-9. Comparison of monomer yields and lignin removal efficiency under different strategies.	89
Fig. 6-10. Simulation results and application of the PDF model.	91
Fig. 7-1. Structural heterogeneity of eucalyptus cell walls revealed by quantitative ^{13}C DP-35s spectra.	100
Fig. 7-2. Representative 2D ^{13}C refocused INADEQUATE spectra of eucalyptus cell walls acquired with cross polarization (CP, gray) selecting rigid domains and direct polarization (DP, orange) enhancing mobile domains.	103
Fig. 7-3. Peak deconvolution of aromatic-region difference spectra obtained from quantitative ^{13}C -DP(35s) NMR under different pretreatment conditions.	104
Fig. 7-4. Overlay of short-range (0.1 s mixing; red) and long-range (1.0s mixing; grey) dipolar-gated ^{13}C - ^{13}C correlation spectra of ^{13}C -labeled eucalyptus cell walls under different pretreatment durations.	107
Fig. 7-5. Representative dipolar-gated 2D ^{13}C - ^{13}C correlation spectra measured with short (0.1s, orange) and long (1.0s, black) mixing times for untreated eucalyptus cell walls.	108
Fig. 7-6. Representative dipolar-gated 2D ^{13}C - ^{13}C correlation spectra measured with short (0.1 s, orange) and long (1.0 s, black) mixing times for untreated eucalyptus cell walls.	109
Fig. 7-7. One-dimensional slices extracted at 149 ppm (top row, G3 of G-lignin) and 153 ppm (bottom row, S3/5 of S-lignin) from dipolar-gated 2D ^{13}C - ^{13}C correlation spectra of untreated eucalyptus (1-0), 1 h (1-1), 18 h (1-6), and 36 h (1-7) pretreated samples. Orange traces correspond to short mixing time (0.1 s)	111

List of Table

Table 2-1: Summary of Lignin Characteristics in Different Biomass Types.....	13
Table 2-2: Molecular Composition and Interaction Modes of Cellulose, Hemicellulose, and Lignin in Plant Cell Walls	17
Table 2-3: Timeline of Solid-State NMR Advances in Elucidating Biomass Cell Wall Structure across Plant Types.....	18
Table 5-1: Comparison of mass fractions measured by NREL and solid-state NMR methods.....	68
Table 6-1: Quantitative analysis of lignin characteristics, including S/G ratios (syringyl/guaiacyl ratios) in the solid substrate, dissolved lignin content, and monomer after depolymerization.....	77
Table 6-2: Yields of G-type (M1, M2, M3), S-type (M4, M5, M6) monomers, and S/G ratios at different PDFs.....	84
Table 7-1: Quantitative analysis of lignin content, delignification degree, and cellulose content in eucalyptus substrates during pretreatment, calculated from ¹³ C DP-35 s spectra.....	102

List of Abbreviation

LCB	Lignocellulosic biomass
RCF	Reductive Catalytic Fractionation
PDF	Pretreatment Depolymerization Factor
NMR	Nuclear Magnetic Resonance
SS-NMR	Solid-State Nuclear Magnetic Resonance
MAS	Magic Angle Spinning
CP	Cross Polarization
DP	Direct Polarization
DNP	Dynamic Nuclear Polarization
INADEQUATE	Incredible Natural Abundance Double Quantum Transfer Experiment
PDSD	Proton-Driven Spin Diffusion
DARR	Dipolar Assisted Rotational Resonance
Gated-PDSD	Dipolar-gated proton-driven spin
PT-CTH	Pretreatment-Catalytic Transfer Hydrogenolysis
CTH	Catalytic Transfer Hydrogenolysis
COS	Camellia oleifera shell
EA	Elemental Analysis
GC-FID/MS	Gas Chromatography - Flame Ionization Detection / Mass Spectrometry
1,4-BDO	1,4-butanediol
DES	Deep Eutectic Solvent
THF	Tetrahydrofuran
2-MeTHF	2-Methyltetrahydrofuran
p-TsOH	Para-Toluenesulfonic Acid
HCl	Hydrochloric Acid
2D-HSQC	^1H - ^{13}C Two-Dimensional Heteronuclear Single Quantum Coherence
Ara	Arabinose
Gal	Galactose
Man	Mannose
$\text{Xn}^{2\text{f}}, \text{Xn}^{3\text{f}}$	Two- and three-fold xylan ($\text{Xylan}^{2\text{f}}$ and $\text{Xylan}^{3\text{f}}$, respectively)
C_A	Acid concentration (mM)
$C_{\beta-O-4}$	Abundance of β -O-4 linkage (%)
C_r	Mass fraction of the residual cellulose over it in the raw biomass (%)
C_s	Organosolv concentration (%)
CTH	Catalytic transfer hydrogenolysis
E	Apparent activation energy (mole/J)
EA	Elemental Analysis
f_l^{H}	Mass fraction of lignin associated with specific hemicellulose i
GC-FID/MS	Gas Chromatography-Flame Ionization Detection/Mass Spectrometry

k_c	Lignin condensation factor in the dual solvent system
k_d	Solvent-specific lignin depolymerization coefficient (1/PDF)
k_i	Specific hemicellulose hydrolysis kinetic coefficient
k_p	Maximum capacity of lignin separation without acid
L_c	Condensed lignin content in the pretreated substrate (%)
L_e	Lignin removal efficiency (%)
L_h	Physical mitigation of lignin associated with the hemicellulose Hydrolysis
L_p	Lignin dissolved in the organosolv due to the physical separation
L_r	Mass fraction of the residual lignin over it in the raw biomass
M_m	Lignin monomer yield (%)
M_m^{Max}	Theoretical productivity of the monolignol of the selected Biomass (%)
M_w	Molecular weight (g/mole)

CHAPTER 1. BACKGROUND AND THESIS OUTLINE

Global industrialization was driven by fossil-based energy resources, which enabled sustained economic growth (Syvitski et al., 2020). However, the dependence on petroleum-based chemicals and energy source raising serious concerns regarding the sustainability of global energy systems and their long-term environmental impacts (Achakulwisut et al., 2023). Continued reliance on these non-renewable resources not only accelerates the depletion of finite reserves but also drives greenhouse gas (GHG) emissions, thereby exacerbating global warming and destabilizing the climate system (Welsby et al., 2021). This dependence continues not intentionally, but rather due to the inertia of entrenched technological pathways and economic structures deeply integrated within fossil-based systems across manufacturing, transportation, and chemical industries (Khan et al., 2022). Although these systems remain efficient, their accumulated environmental costs highlight the need for renewable alternatives that maintain performance and align with current infrastructure (Zhao et al., 2023).

Among various renewable feedstocks, lignocellulosic biomass (LCB) has attracted considerable attention due to its abundant availability, rapid regeneration capacity, and carbon-neutral lifecycle (Singhvi & Gokhale, 2019). Each hectare of agricultural land could yield approximately ten metric tons of dry biomass annually, offering a substantial and renewable carbon source to produce biofuels and biochemicals (Cassidy et al., 2013). The intrinsic presence of both aromatic and aliphatic structures in these resources endows the potential for conversion into a diverse range of bio-based products (Yamakawa et al., 2018). Through optimized biorefinery strategies and chemical derivatization, biomass-derived products can effectively substitute petroleum-based chemicals, thereby reduce fossil fuel dependence and substantially contribute toward achieving global carbon neutrality.

Efficient utilization of lignocellulosic biomass is essential for maximizing resource value and minimizing waste throughout the transition toward carbon neutrality. However, lignocellulosic biomass consists of highly recalcitrant bio-composites comprising cellulose, hemicelluloses, and lignin (Dong et al., 2019b). These

components are intricately interconnected at the nanoscale, forming complex anisotropic structures within plant secondary cell walls (Kang et al., 2019). The covalent linkages and extensive hydrogen-bonding interactions between lignin and polysaccharides lead to a dense and heterogeneous matrix featuring both hydrophobic and hydrophilic domains, which severely restricts biomass accessibility and impedes enzymatic efficiency in biorefinery processes (Kirui et al., 2022; Liu et al., 2016).

To overcome biomass recalcitrance, current strategies involve a sequence of steps: pretreatment (Islam et al., 2020a), enzymatic hydrolysis (Chan et al., 2021), fermentation (Lin & Tanaka, 2006), lignin depolymerization via catalytic transfer hydrogenolysis (CTH) (Das et al., 2018), and final product recovery (Alherech et al., 2021). Among these, pretreatment is particularly essential, enabling lignin-carbohydrate dissociation and improving substrate accessibility for downstream conversion. Various pretreatment approaches have been developed to disrupt the intrinsic resistance of lignocellulosic biomass, including Kraft pulping, acid or alkaline hydrolysis (Woiciechowski et al., 2020), and deep eutectic solvent (DES) systems (Shen et al., 2019) (Gierer, 1980). Kraft pretreatment, widely used in the pulp and paper industry, employs sulfide and sulfate ions to degrade lignin and bleach cellulose. However, the harsh sulfur-containing environment promotes both depolymerization and undesired condensation reactions, leading to the formation of water-soluble but structurally altered lignin with limited utility for downstream upgrading (Prasad et al., 2023). For the acid organosolv system, selectively cleaves ester and ether bonds between hemicelluloses and lignin, facilitating partial lignin removal and carbohydrate exposure. The organosolv systems promote lignin solubilization and allow for controlled reprecipitation, thereby minimizing recondensation and preserving reactive functionalities (Dong et al., 2019b). These features are critical for efficient catalytic depolymerization and high monomer yields.

Despite growing interest in whole-biomass utilization, most biorefineries remain centered on the conventional sugar-platform approach (Fasahati et al., 2022), which prioritizes cellulose valorization while neglecting lignin and hemicellulose fractions. This selective focus limits both the economic viability and resource efficiency of

biorefinery processes (Patel et al., 2025). In addition, when certain lignin structures, particularly β -O-4 linkages, are readily cleaved under harsh pretreatment conditions, unfavorable condensation reactions and the formation of stable C-C bonds can be triggered (Li et al., 2018a). These undesired transformations reduce the availability of fermentable sugars and monolignols, ultimately compromising catalytic upgrading efficiency and product yields. To avoid lignin condensation, stabilization techniques such as formaldehyde-assisted fixation and selective arylation have been explored (Li et al., 2024b; Shuai et al., 2016a). While effective in certain contexts, these methods often suffer from limited applicability across heterogeneous biomass types and face substantial barriers to industrial scalability. Meanwhile, conventional pretreatment practices typically rely on empirical adjustment of parameters-including solvent type, biomass-to-solvent ratio, temperature, and reaction time-based on the specific characteristics of each feedstock (Meng et al., 2020). Those approaches lack the precision required to simultaneously achieve efficient delignification and structural preservation of lignin.

Advanced process control aiming to balance the performance of biomass fractionation and lignin condensation. In biorefinery system, pretreatment severity normally has been optimized through biomass-specific experiments at various operational conditions, like organosolv types and doses and acid concentration, treatment time, and temperature (Islam et al., 2020a). Owing to the inherent recalcitrance and wide diversity of biomass, the construction of a universally applicable process control algorithm for a biorefinery is challenging, especially when the concerned biomass is derived from waste. The existing process control models have been developed by the pulping industry with preprocessed wood chips. These models primarily focus on predicting the hydrolysis capacity of pretreated substrate for sugar production, while neglecting the integrity of the fractionated lignin for monolignol production. The essential information on structural changes of pretreated biomass has been rarely reported, likely due to the limitations and challenges of current analytical tools. To overcome the challenge, a fundamental prerequisite for addressing this challenge is a quantitative understanding of biomass molecular architecture and its

inherent recalcitrance to chemical deconstruction. Therefore, to address these issues concerned in biorefinery the structure of the dissertation (**Fig. 1-1**) is established:

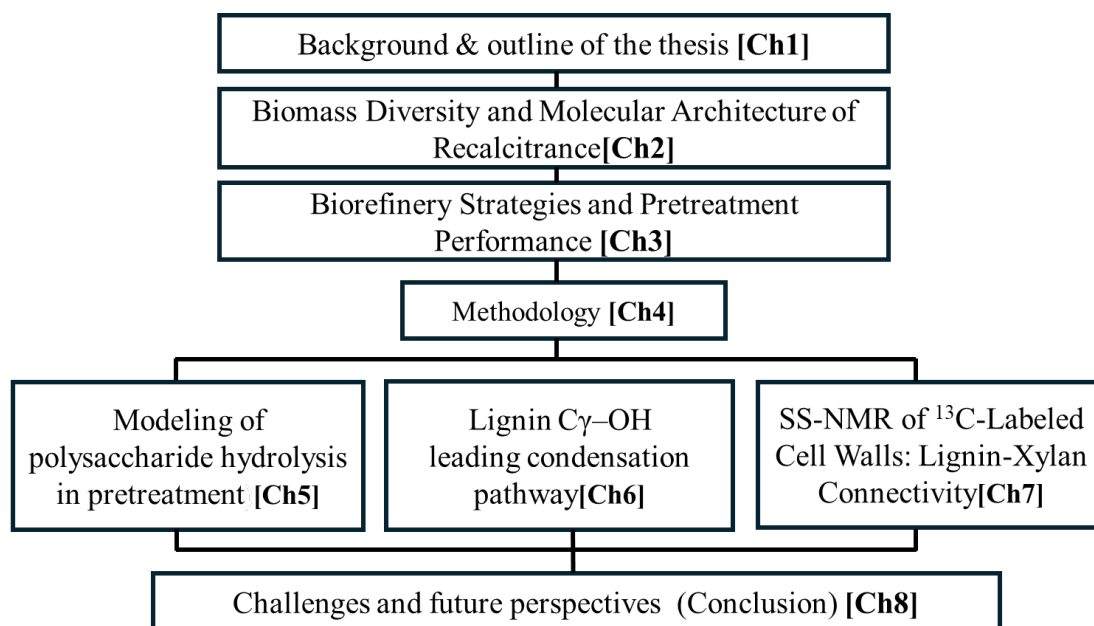


Fig. 1-1. Key features and chapters of this dissertation

Chapter-1 introduces the background of the dissertation. It outlines the global dependence on fossil fuels and petroleum-derived chemicals, the environmental and resource concerns associated with their continued use, and the potential of lignocellulosic biomass as a renewable and carbon-neutral alternative. It addresses the challenges of limited molecular-level insights into plant cell wall architecture and highlights the need for precise process control strategies to minimize side reactions.

Chapter-2 systematically reviews biomass diversity, classifying major categories such as hardwoods, softwoods, grasses, barks, and seeds. Structural differences in cellulose crystallinity, hemicellulose composition, and lignin content are examined, and their relationships with recalcitrance are discussed. This classification provides essential insights into how biomass molecular architecture dictates pretreatment outcomes and sets the stage for targeted biorefinery strategies.

Chapter-3 surveys biorefinery strategies with a focus on pretreatment methods. Conventional pulping, acid/alkali hydrolysis, organosolv, and emerging solvent systems are compared in terms of delignification efficiency, sugar release, and lignin

valorization potential. The chapter identifies technical gaps in balancing carbohydrate utilization with lignin preservation and underscores the importance of integrated process optimization.

Chapter-4 introduces the methodological foundation of this dissertation. Analytical instruments such as solid- and solution-state NMR, GC-MS/FID, and GPC are described, along with catalytic depolymerization techniques. This chapter also presents the mathematical formulation and parameters of the Pretreatment Depolymerization Factor (PDF), a kinetic model designed to simulate polysaccharide hydrolysis and lignin condensation under varying pretreatment conditions.

Chapter-5 develops the PDF model through extensive acid-organosolv pretreatments of *Camellia oleifera* shell. Hydrolysis kinetics of different xylan conformations are quantified by solid-state NMR, providing the mechanistic basis for model construction. The model achieves $R^2 > 0.91$ in simulating hemicellulose hydrolysis and is validated across pine, eucalyptus, and bamboo, demonstrating broad applicability across biomass feedstocks.

Chapter-6 investigates lignin condensation pathways during pretreatment. Using 2D-HSQC NMR and catalytic hydrogenolysis, a γ -OH initiated C β -C γ cleavage pathway are identified and incorporated into the PDF model. The refined model is then applied to semi-continuous solvent-switching pretreatments, achieving 40.1 mg g⁻¹ monolignol yield with 95% delignification nearly three-fold higher than baseline. These results provide mechanistic insight into controlling lignin reactivity.

Chapter-7 employs uniformly ¹³C-labeled eucalyptus to probe biomass molecular architecture at the cell wall level. Sequential weak-acid fractionation combined with advanced SS-NMR resolves covalent lignin-xylan linkages and spatial proximities within secondary cell walls. These findings explain residual recalcitrance and provide direct molecular evidence of lignin connectivity, offering a structural basis for improved pretreatment designs.

Chapter-8 concludes the main findings and contributions of this thesis, which addresses severity-dependent polysaccharide–lignin and lignin–lignin interactions during biomass pretreatment. To achieve this objective, pretreatment experiments were systematically

conducted to establish a quantitative and mechanistic understanding of lignin–polysaccharide dissociation and structural evolution during biorefinery pretreatment processes. A PDF model was developed, enabling quantitative correlations to be established between pretreatment severity and changes in lignin inter-unit linkages, hydroxyl group distributions, and molecular characteristics. To further elucidate lignin condensation mechanisms during pretreatment, solid-state and 2D-HSQC NMR analyses combined with catalytic hydrogenolysis were employed, leading to the identification of a γ -OH-initiated C β –C γ cleavage pathway that was subsequently incorporated into the PDF model. Application of the refined model to semi-continuous pretreatment achieved 95% delignification with an approximately 30% improvement in monolignol yield. In addition, uniformly ^{13}C -labeled eucalyptus combined with weak-acid fractionation was utilized to resolve molecular architecture at the cell-wall level, revealing potential origins of biomass recalcitrance and providing direct molecular evidence for covalent bonding and spatial connectivity among lignin and polysaccharide domains. Collectively, these findings deepen the mechanistic understanding of lignin reactivity and offer guidance for severity-controlled pretreatment strategies in lignin-first biorefinery systems.

CHAPTER 2. WOODY BIOMASS AND ITS STRUCTURE

ABSTRACT

This chapter provides a systematic review of the structural and compositional characteristics of lignocellulosic biomass, classifying into five categories: hardwoods, softwoods, grasses, barks, and seeds. By linking anatomical origin with chemical composition, the analysis highlights how variations in cellulose, hemicellulose, and lignin content fundamentally shape biomass recalcitrance and valorization potential. Hardwood lignin's, enriched in β -O-4 linkages and with balanced S/G ratios, are highly depolymerizable, whereas softwood lignin's, dominated by guaiacyl units and C-C linkages, exhibit higher resistance. Grasses display more complex lignin profiles, including hydroxycinnamate residues, while bark tissues contain suberin-rich matrices that significantly impede reactivity. Seed-derived biomass reveals two distinct lignin architectures: catechyl (C) lignin in seed coats, which enables nearly quantitative monomer yields under mild catalytic conditions, and reactive G-type lignin in seed shells, which favors carbonization rather than depolymerization. This review reveals that the molecular architecture of secondary cell walls as the fundamental determinant of biomass recalcitrance: cellulose microfibrils, hemicellulose conformations, and lignin networks form highly interconnected nanoscale composites that restrict chemical and enzymatic accessibility. Recent advances in MAS-SS-NMR and isotopic labeling have enabled direct probing of these interactions, revealing lignin-xylan-cellulose associations that define biomass rigidity and conversion efficiency. Comparative structural models are proposed for each biomass category, illustrating how molecular organization governs biorefinery outcomes. By establishing the connection between biomass classification and nanoscale architecture, this chapter provides a framework for rational biomass selection and process design which is critical for developing tailored pretreatment and catalytic depolymerization strategies, ultimately advancing efficient and sustainable lignocellulosic biorefineries.

2.1. Chemical composition of lignocellulosic biomass

The chemical composition, molecular architecture, and interactions within the plant secondary cell wall remain insufficiently understood. Lignocellulosic biomass is a highly recalcitrant bio-composite comprising cellulose, hemicelluloses, and lignin (Dong et al., 2019b). These components are intricately interconnected at the nanoscale, forming complex anisotropic structures within plant secondary cell walls (Kang et al., 2019). The covalent linkages and extensive hydrogen-bonding interactions between lignin and polysaccharides lead to a dense and heterogeneous matrix featuring both hydrophobic and hydrophilic domains, which severely restricts biomass accessibility and impedes enzymatic efficiency in biorefinery processes (Kirui et al., 2022; Liu et al., 2016). The 3D matrix is responsible for the wall's mechanical strength, flexibility, and biological resistance, yet conventional analytical techniques often fail to preserve or resolve the native architecture. To facilitate the rational design of biorefinery strategies, plant biomass was classified into five distinct categories: hardwood, softwood, herbaceous plants (grasses), barks, and seeds based on their anatomical origin and physiochemical composition (Fig. 2-1).

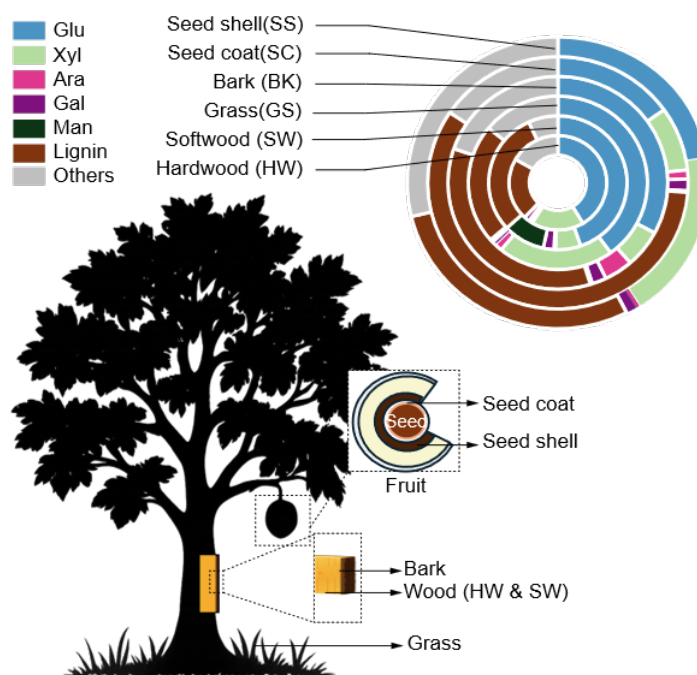


Fig. 2-1. Plant part classification and chemical composition (weight %) of biomass types,

including hardwood (HW), softwood (SW), grass (GS), bark (BK), seed coat (SC), and seed shell (SS). Data are taken from **Table S2-1**

These biomass types differ significantly in their cellulose, hemicellulose, extraction and lignin content, as well as in the structural characteristics of their lignin (Guan et al., 2025). Such variation directly influences their recalcitrance and suitability for catalytic depolymerization (Yan et al., 2020). To facilitate a comprehensive evaluation of lignocellulosic biomass in biorefinery applications, we compiled and analyzed a dataset encompassing more than 30 representative plant-derived biomass species. These were systematically categorized into five major classes based on anatomical origin and lignin composition: hardwoods, softwoods, grasses, barks, and seeds. Hardwoods including eucalyptus, aspen, and poplar, which typically possess moderate lignin content and a high proportion of β -O-4 linkages, rendering them highly favorable for depolymerization and catalytic conversion processes. Pine and spruce were classified as softwoods, characterized by higher lignin content dominated by guaiacyl units and more C-C structures that pose greater resistance to depolymerization. The grass category consisted of fast-growing herbaceous plants such as miscanthus and wheat straw, known for their diverse hemicellulose structures and incorporation of hydroxycinnamate-type lignin. Bark samples, including those from cedar and birch, were also included to reflect the complex suberin-rich tissues that often differ structurally from inner wood tissues.

Additionally, various seed-derived biomass types were considered, comprising both seed coats containing C-lignin and seed shells lacking it, each exhibiting distinct chemical architectures. These variations provide critical insight into the structural and compositional diversity of biomass resources. A complete summary of the chemical composition, β -O-4 content, and RCF monomer yields for all samples are provided in **Table S2-1**. In detail, hardwoods typically contain 37.1-45.7% cellulose and 14.6-24.3% hemicellulose, with a relatively moderate lignin content (14.2-24.9%). The dominant hemicellulose in hardwoods is glucuronoxylan, a C5-based sugar polymer. A defining structural trait of hardwood lignin is the co-existence of syringyl and guaiacyl units,

with S/G ratios ranging from 0.75 to 7.0. Notably, hardwood lignin exhibits a high proportion of β -O-4 aryl ether linkages, typically in the range of 55-65%, though some species exceed 90% which renders them highly susceptible to depolymerization under mild reductive catalytic conditions. This abundance of cleavable linkages contributes to monomer yields as high as 50% in reductive catalytic fractionation (RCF), making hardwoods a preferred substrate in lignin-first biorefinery. Softwoods have higher cellulose content (42.0-48.4%) and a significant amount of mannan-rich hemicellulose, with mannose accounting for up to 12.2%. Their lignin content is also relatively high (24.7-36.6%), composed almost exclusively of G-units. As a result, softwood lignin exhibits greater crosslinking and condensation, with a much lower β -O-4 linkage content (31.2-38.1%). This condensed architecture contributes to higher oligomer formation and lower monomer yields (typically 10-22%) during RCF. Nevertheless, the high selectivity of G-units offers opportunities for targeted chemical valorization using tailored.

Grasses or herbaceous feedstocks show moderate cellulose (31.0-52.5%) and hemicellulose (18.0-22.9%) contents, with lignin ranging from 18.5% to 27.1%. Their lignin is chemically diverse, composed of S/G/H units and acylated with hydroxycinnamic acids such as p-coumaric and ferulic acid, which esterify the hemicellulose side chains. These structural modifications enhance biomass rigidity but reduce enzymatic accessibility and monomer yields. Although β -O-4 content in grass lignin varies (51.0-67.1%), the presence of noncanonical linkages and acyl substitutions complicates depolymerization and downstream upgrading. Barks exhibit highly variable composition due to their protective physiological role. Cellulose content ranges from 14.0% to 58.5%, hemicellulose from 2.5% to 10.9%, and lignin between 30.8% and 40.0%. In addition to lignin, bark tissues contain large amounts of suberin, a hydrophobic aliphatic-aromatic polymer that forms a lamellar barrier in the cell walls. Suberin content in some species, such as cork oak and birch, can reach up to 60%. The dual presence of lignin and suberin renders bark highly recalcitrant and poorly reactive under conventional catalytic depolymerization conditions.

Seeds represent a structurally diverse biomass category with high lignin content

and are broadly divided into two valorization-relevant parts: seed coats and seed shells. the seed coat of certain species such as castor, vanilla, and candlenut is enriched with C-lignin, a recently recognized form of lignin composed almost exclusively of caffeyl alcohol-derived benzodioxane units. Unlike conventional G/S/H-type lignins, C-lignin displays an unprecedented homogeneity and linearity in its linkage pattern, enabling nearly quantitative monomer yields (up to 84%) under mild catalytic conditions. The absence of C-C condensation and the dominance of β -O-4 like benzodioxane linkages make seed coats among the most promising natural sources for selective aromatic production.

The seed shell, including biomass such as rice husks, walnut shells, or peanut shells, typically contains 30-60% Klason lignin, along with significant polysaccharide fractions (e.g., 13-34% cellulose and 8-26% hemicellulose). These tissues are rich in condensed G-type lignin and exhibit low β -O-4 content, leading to low monomer yields (9.8-38%) and increased structural rigidity. Despite their lower reactivity, seed shells remain attractive for bulk carbon applications or pyrolysis-based conversion due to their high lignin loading and dense matrix. Together, these classifications reflect the structural and compositional diversity inherent to lignocellulosic feedstocks. A clear understanding of these differences is essential for guiding biomass selection and designing biomass-specific pretreatment and depolymerization strategies.

2.2. Structural diversity of lignin in different biomass

To further elucidate the structural diversity of lignin among different biomass types, we compiled and compared the monolignol compositions of various representative lignocellulosic feedstocks. **Table2-1** summarizes the syringyl-to-guaiacyl (S/G) ratios, hydroxyphenyl content, and other relevant substructures including hydroxycinnamates and C-units across five biomass categories: hardwoods, softwoods, grasses, barks, and seed-derived tissues and show in **Fig. 2-2**.

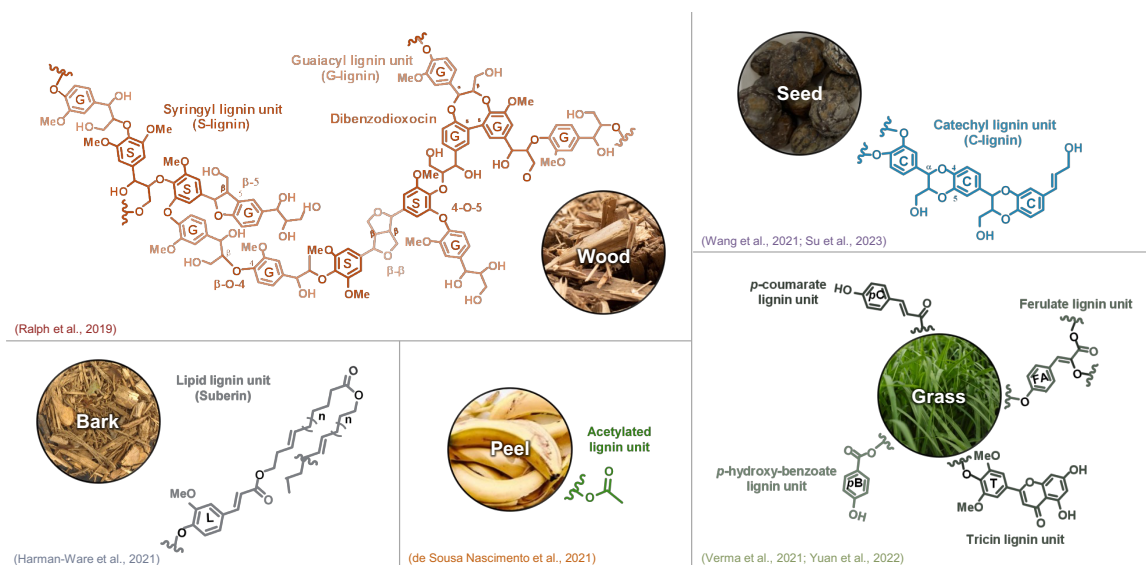


Fig. 2-2. Lignin model structures and unique lignin units investigated in this study.

Hardwoods typically exhibit both syringyl and guaiacyl lignin units, with S/G ratios ranging widely from 0.7 to 7.0, depending on species. For example, silver birch presents a notably high S/G ratio, while European beech and teak exhibit lower values. The relatively high abundance of S-units correlates with more linear and less cross-linked lignin structures, enhancing depolymerization efficiency and making hardwood lignin especially suitable for reductive catalytic fractionation (RCF) processes. Softwoods, on the other hand, predominantly consist of guaiacyl units, often with negligible syringyl and hydroxyphenyl content. The lack of S-units results in more condensed lignin networks with increased C-C linkages, contributing to higher recalcitrance. Notably, maritime pine and spruce display extremely low or undetectable S/G ratios, reflecting the structural rigidity and lower reactivity of softwood lignin. Grasses possess a more complex lignin composition, comprising not only S, G, and H units, but also significant amounts of *p*-coumaric acid (pCA), ferulic acid (FA), and triclin (T) residues.

Table 2-1: Summary of Lignin Characteristics in Different Biomass Types

Biomass Type	Main Lignin Units	S/G Ratio Range	Structural Features
Hardwood	Syringyl (S), Guaiacyl (G)	0.7 - 7.0	High content of β -O-4 linkages, less condensed; suitable for lignin depolymerization
Softwood	Mainly Guaiacyl (G)	0 - 0.048 (or H/G)	Highly condensed, lower β -O-4 content; rich in C-C linkages
Grass	S, G, H, p-coumarate, ferulate	$\sim 32:63$ (S:G, e.g. Sorghum)	Abundant hydroxycinnamate esters; complex cross-links and tricin units
Bark	S, G + Suberin	~ 0.1 (S/G, when available)	Contains suberin; highly hydrophobic, impermeable, and structurally rigid
Seed coat	Catechyl (C-lignin, nearly exclusive)	N/A	Linear benzodioxane structures, less cross-linking, homogeneous polymer

These hydroxycinnamate-derived units are ester-linked to hemicelluloses and contribute to both structural rigidity and cross-linking between lignin and polysaccharides. For example, corn cob and sorghum show particularly high pCA and FA content, making grass lignin chemically distinct and requiring specialized valorization strategies. Barks exhibit an additional layer of complexity due to the presence of suberin a hydrophobic polyester that coexists with lignin in protective tissues. Suberin content varies substantially across bark species, ranging from 1.5% to over 60%, and often coexists with lignin having low S/G ratios. For instance, cork oak and silver birch barks contain both suberin and S-lean lignin, complicating chemical fractionation and requiring selective solvent systems.

Seed-derived biomass, especially seed coats, displays the most unique lignin profile, characterized by the presence of C-lignin. Unlike S/G/H-based lignins, C-lignin consists primarily of benzodioxane structures formed from caffeyl alcohol, offering a linear and highly homogeneous polymer. Vanilla and castor seed coats contain exceptionally high proportions of C-lignin, with some species showing nearly complete

substitution of conventional monolignols. This structural simplicity facilitates selective depolymerization to catechol-based monomers, providing a promising route for bio-aromatic production. Together, these results highlight the considerable variation in lignin monomer composition across biomass types and underscore the importance of tailoring pretreatment and catalytic depolymerization strategies according to lignin structure. A complete list of S/G ratios, H and C contents, and suberin levels is provided in **Table 2-1**.

2.3. Molecular Architecture and Recalcitrance of Plant cell walls

In lignocellulosic biomass secondary cell wall, the cellulose forms the backbone of the cell wall architecture. It consists of β -1,4 linked glucose chains that are tightly packed into crystalline microfibrils through intra- and intermolecular hydrogen bonding (Chen, 2014). These microfibrils provide structural rigidity and are organized into hydrophilic and hydrophobic faces, which influence their interaction with surrounding matrix polymers. Surrounding the cellulose scaffold is hemicellulose, a diverse group of branched heteropolysaccharides. In hardwoods and grasses, xylan is the dominant hemicellulose, composed of a β -1,4-xylose backbone substituted with side groups such as glucuronic acid, arabinose, or acetyl groups (Guan et al., 2025). Xylan can adopt two distinct conformations: a flat two-fold screw (Xn^{2f}) that closely associates with the hydrophilic surface of cellulose, and a twisted three-fold screw (Xn^{3f}) that is more loosely bound. (Simmons et al., 2016) These conformations are not just structural variants they directly regulate how hemicellulose interacts with both cellulose and lignin within the wall matrix. Lignin is an irregular polyphenolic network formed via radical coupling of monolignols-primarily guaiacyl (G), syringyl (S), and p-hydroxyphenyl (H) units (Boerjan et al., 2003). These units are interconnected by various ether and carbon-carbon linkages, with the β -O-4 aryl ether bond being the most abundant and chemically accessible for depolymerization. Lignin fills the interstitial spaces of the polysaccharide network, imparting hydrophobicity and mechanical strength. SS-NMR studies have shown that lignin preferentially associates with Xn^{3f} -type xylan regions and

is often colocalized with the cellulose microfibril interface, forming lignin-xylan-cellulose complexes that limit solvent penetration and enzyme accessibility (Kang et al., 2019). Based on these cumulative insights, comparative 3D structural models have been proposed for the cell walls (**Fig.2-3**).

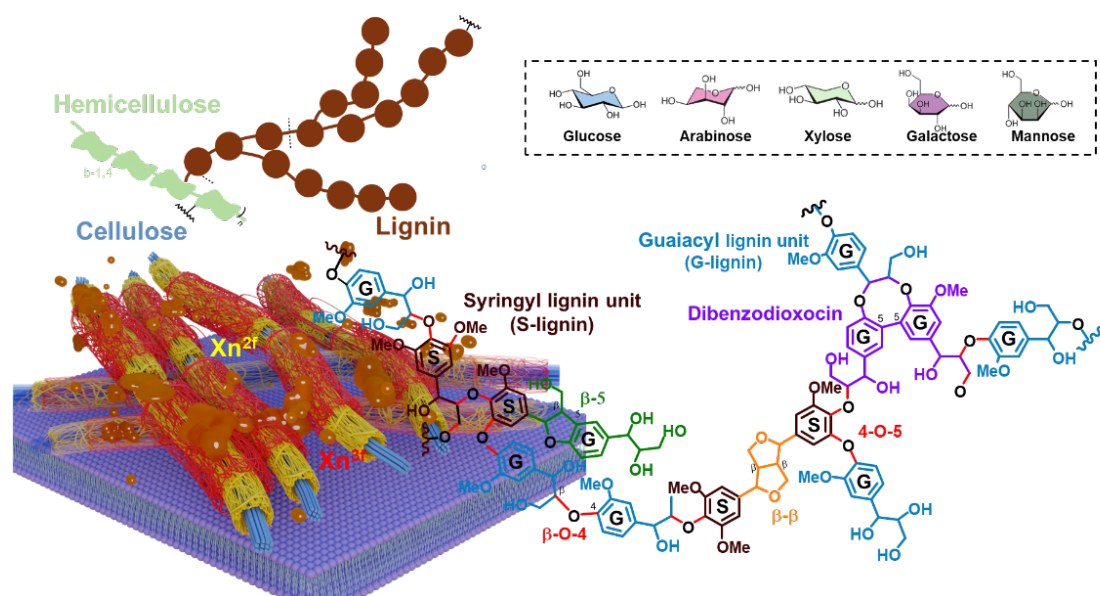


Fig.2-3. Molecular architecture and recalcitrance of plant cell walls. Schematic representation of cellulose microfibrils, hemicellulose chains, and lignin polymers forming a compact and cross-linked matrix. Major lignin subunits (S- and G-units) and key inter-unit linkages (β -O-4, β -5, β - β , 4-O-5, and dibenzodioxin) are illustrated.

In hardwoods, cellulose microfibrils are coated with Xn^{2f} forming non-crystalline interfaces that provide flexibility for deformation. Xn^{3f} layers further mediate lignin anchoring, primarily on the outer regions of microfibrils. Grass cell walls display a similar configuration, though GAX introduces additional crosslinking complexity. In softwoods, GGM functions as a structural mediator, bridging lignin with cellulose bundles, while Xn^{3f} and extended GGM chains define the matrix architecture (Kirui et al., 2022). In contrast, seed biomass, a structurally diverse category can be divided into seed coats and seed shells. Seed coats in species such as castor, vanilla, and candlenut are enriched in C-lignin, composed almost entirely of caffeyl alcohol-derived benzodioxane units. This polymer is chemically homogeneous, linear, and devoid of C-C condensations, making it highly amenable to near-complete depolymerization into

catechols under mild catalytic conditions. Li et al., 2018b) (Li et al., 2018b) On the other hand, seed shells such as rice husks or walnut shells contain 30-60% condensed G-type lignin and exhibit disordered cellulose chains and low crystallinity, characteristics that complicate efficient depolymerization but favor carbonization applications. Similarly, bark tissues exhibit unique features, including low cellulose order and high suberin content, which further impede enzymatic and catalytic accessibility. These traits were inferred from post-treatment measurements, compositional analyses, and biophysical assessments such as water retention and fiber tensile strength (Guan et al., 2025). Despite these advancements, current NMR-based models are largely constrained by the requirement for ^{13}C -labeling, limiting their applicability to fast-growing plants or seedlings. Studies on more recalcitrant tissues such as bark or seed biomass remain sparse. A notable exception is Kirui's work on unlabeled cotton, which used DNP-enhanced 2D-INADEQUATE analysis to gain structural insights, though with fewer cross peaks than labeled samples (Kirui et al., 2022). These findings underscore both the achievements and ongoing challenges in decoding plant SCW architecture, offering a roadmap for engineering biomass more amenable to deconstruction and conversion.

These three components are not organized as separate layers but rather form an intertwined, nanoscale composite where the physical proximity and chemical interactions among polymers result in a highly recalcitrant structure. The structural anisotropy and chemical heterogeneity of this composite matrix represent a key bottleneck in the bioconversion of lignocellulosic biomass.

Table 2-2: Molecular Composition and Interaction Modes of Cellulose, Hemicellulose, and Lignin in Plant Cell Walls

Component	Primary Structural Units	Intramolecular Linkages	Intermolecular Interactions	Most Probable Interconnection Sites
Cellulose	β -D-glucopyranose (glucose) units	β -1,4-glycosidic bonds forming linear chains; stabilized by intra- and inter-chain hydrogen bonds	Forms hydrogen bonds with hemicellulose; physically embedded in lignin matrix via van der Waals forces	Hydrophilic surface interacts with two-fold screw xylan through multiple hydrogen bonds
Hemicellulose	Mainly β -D-xylopyranose (xylan backbone), with side groups such as 4-O-methyl-D-glucuronic acid, L-arabinofuranose, and acetyl groups	β -1,4-glycosidic bonds in the backbone; α -linkages in side chains (e.g., α -1,2, α -1,3)	Forms hydrogen bonds with cellulose; covalent bonds (ester or ether) with lignin	Side chains (e.g., glucuronic acid, arabinose) form ester or ether linkages with lignin phenylpropane units
Lignin	Phenylpropanoid monomers: guaiacyl (G), syringyl (S), and p-hydroxyphenyl (H) units	Linked via β -O-4 aryl ether bonds, β - β resinol, β -5 phenylcoumaran, and carbon-carbon linkages	Covalently linked to hemicellulose via ester and ether bonds; physically entangled with cellulose	γ -hydroxyl group of side chain esterified with hemicellulose side chains;

Table 2-3: Timeline of Solid-State NMR Advances in Elucidating Biomass Cell Wall Structure across Plant Types

Year	Species	Classification	Key Experiment	Main result	Mark	Reference
2014	<i>b. distachyon</i>	Grass	CP DP INADEQUATE PDS	Molecular-level insights into the polysaccharide organization of grass cell walls, the unique interaction of GAX with cellulose	¹³ C Labeled Biomass	(Wang et al., 2014)
2016	<i>I: Arabidopsis</i> <i>II: Brachypodium</i>	Hardwood	DP INADEQUATE PDS	Cellulose presence of inaccessible core, intermediate interior layer, and surface layer have at least 24 cellulose chains	¹³ C Labeled Biomass	(Wang & Hong, 2016)
2019	<i>Norway spruce</i>	Softwood	INADEQUATE PDS	Identified lignin surrounding cellulose clusters and GGM as a key linker in softwood, proposing an organized molecular model of softwood.	¹³ C Labeled Biomass	(Terrett et al., 2019)
2019	<i>I: Arabidopsis thaliana</i> <i>II: Rice steam</i> <i>III: Switchgrass</i> <i>IV: Maize</i>	Grass	DP CP INADEQUATE CORD PDS	Lignin forms hydrophobic nanodomains and extensively interacts with xylan through electrostatic forces, highlighting its preference for binding xylans with 3-fold or distorted 2-fold helical conformation	¹³ C Labeled Biomass & DNP	(Kang et al., 2019)
2022	<i>I: Eucalyptus</i> <i>II: Poplar</i> <i>III: Spruce</i>	Hardwood & softwood	INADEQUATE RFDR PDS DP CP	Glycan-aromatic associations across grasses, hardwoods, and softwoods, with lignin binding to non-flat Xylan (Xylan3f) and partially interacting with cellulose junctions,	¹³ C Labeled Biomass & DNP	(Kirui et al., 2022)
2024	<i>Hybrid poplar (DN34)</i>	Hardwood	Multi-CP-DARR INADEQUATE	Refined the understanding of plant secondary cell walls by focusing on the structural configuration and interactions among cellulose, hemicelluloses, and lignin integrated NMR data into an atomistic macromolecular model	¹³ C Labeled Biomass	(Addison et al., 2024)

CHAPTER 3. BIOREFINERY AND PRETREATMENT

ABSTRACT

Lignocellulosic biomass presents a promising feedstock for sustainable biofuel and bioproduct production. Despite its abundance, only a fraction of lignocellulosic biomass is currently utilized, highlighting its complex compositions and the need for advanced biorefinery technologies. In this chapter, we systematically review various pretreatment strategies employed within biorefinery processes, focusing on their mechanisms, efficiency, and effectiveness in biomass component recovery. Various pretreatment techniques, including organosolv methods, and DES, are analyzed for their capabilities in selectively fractionating lignocellulosic biomass into cellulose, hemicellulose, and lignin. The chapter further examines and compares two prominent integrated strategies: Pretreatment-Catalytic Transfer Hydrogenolysis (PT-CTH) and Reductive Catalytic Fractionation (RCF). The mechanisms of lignin extraction, preservation of structural motifs, and the subsequent production of aromatic monomers under these methodologies are critically discussed. Additionally, a comparative analysis is provided for five representative biomass (hardwood, softwood, grass, bark, and seed) assessing their responses to different pretreatment conditions and highlighting the implications for biomass valorization. Our analysis reveals that while RCF converts lignin to higher monomer yield and exhibits higher economic feasibility, PT-CTH is more sustainable and utilizes the whole biomass more efficiently due to its efficient pretreatment. Hardwood, softwood, and grass demonstrate high resource efficiency with hardwood producing the highest yield of monomers in RCF. Seed coats, rich in C-lignin, offer high potential for aromatic monomer production, but demonstrate low resource efficiency while barks require tailored approaches due to their complex lignin units. This chapter proposes recommendations to advance efficient, sustainable biorefinery operations, leading the structural diversity within lignocellulosic biomass for optimized biomass utilization, and supports the development of more economic and environmentally friendly biorefinery techniques toward carbon neutrality.

3.1. Biorefinery strategies

Biorefinery techniques have been broadly conceptualized using several organizational frameworks, among which sugar-based (biological) and syngas-based (thermochemical) platforms are the most widely discussed. The sugar-based platform has focused on how to effectively hydrolyze cellulose and hemicellulose into sugars for fermentation to bioethanol and many other bio-products (Robak & Balcerek, 2020).

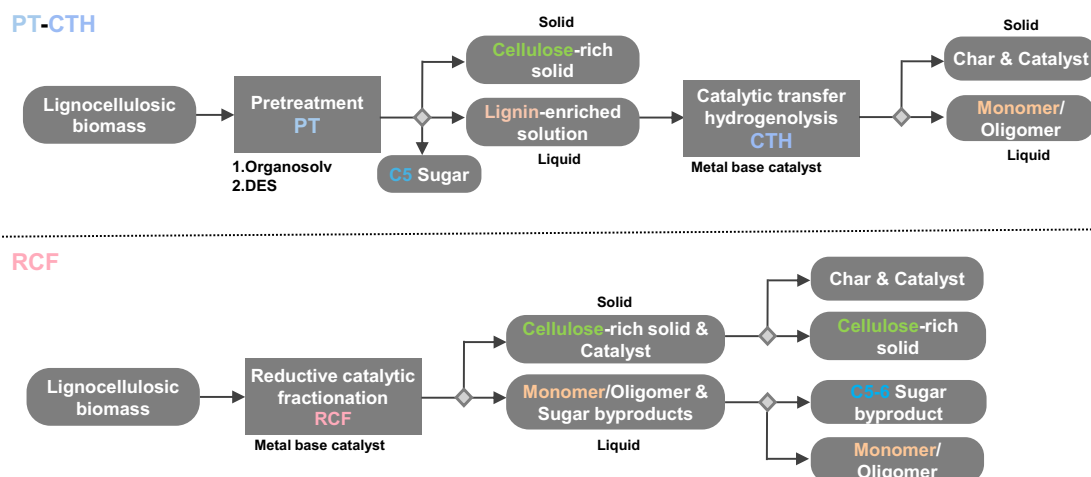


Fig. 3-1. Process flow of PT-CTH and RCF biorefinery.

The syngas- platform, such as gasification and pyrolysis, has been designed for rapid production of syngas, bio-oil, and biochar from lignocellulosic biomass (Wang et al., 2017). Based on the two basic principle, the "lignin-first" concept has been introduced as an emerging approach on the valorization of lignin to produce high-value aromatic monomers for downstream applications, such as food, biopolymer, and pharmaceuticals (Li et al., 2020; Ren et al., 2021). This approach is crucial for balancing the low profitability of biorefinery practices that use non-food feedstock compared to conventional refineries and food-based biorefinery products. By doing so, it enhances the competitiveness of lignocellulosic biomass in establishing a bio-economy (Abu-Omar et al., 2021; Ragauskas et al., 2014). The high complexity of lignocellulosic biomass has, however, presented significant technical challenges that must be addressed to fully realize the benefits of biorefinery technologies. Biomass is plant-based composite materials with physical and chemical properties vary significantly from the

harvesting plant species, parts, and conditions of the plants. This variability can lead to a wide range of products/by-products and yields in biorefinery, impacting its economic viability of the selected techniques. As researchers and industries continue to explore innovative solutions, the effective utilization of lignin and the advancement of biorefinery technologies will be key to unlocking the full potential of lignocellulosic biomass. A comprehensive understanding of the relationship between biomass characteristics and processing methods is essential to fully clarify these benefits.

Pretreatment is the most essential unit process in the biorefinery (Joshi et al., 2015). It aims to fractionate the lignin aromatics from the polysaccharide-based fibers formed by cellulose and various types of hemicelluloses (Terrett et al., 2019). Pretreatment includes two major components, i.e., physical and chemical processing. The physical treatment involves various approach to break down biomass particles, such as ultrasound, microwave, extrusion, and milling, while chemical treatment are critical to cleave the linkages among the building block chemicals (Islam et al., 2021b). As the effectiveness of physical separation rely heavily on the chemical reactions in biorefinery, we emphasized the chemistry of biorefinery in this review. The goal of pretreatment is to increase accessibility of the substrate to allow chemicals or enzymes for effective conversion. This process reduces cellulose polymerization, destroys cellulose crystallinity, and increases the surface area and pore size of the substrates (Zhao et al., 2017). Pretreatment also separates cellulose from lignin, preventing the inhibition caused by lignin during enzymatic hydrolysis of cellulose into residual sugars (Leu & Zhu, 2013b). Various chemical(s) (Kuglarz et al., 2018; Leu et al., 2013) and solvent(s) (Islam et al., 2021a; Martínez et al., 2018) have been applied to catalyze the reaction and to dissolve the fractionated lignin and hemicellulose, respectively. Kraft pulping process (Na_2S and NaOH in 145-170 °C) can be considered as the most representative pretreatment technique for delignification (Fernández-Rodríguez et al., 2019). It has outstanding lignin removal efficiency (about 90%) and hence is widely applied in the pulping industry (Kim et al., 2016). In kraft and the other alkali pretreatment processes, lignin is removed from the cellulosic/hemicellulose bundles via significant either bond breakage, lignin degradation and condensation (Abolore et al., 2024). Kraft and alkali-

based pretreatment processes, including ammonia fiber explosion (AFEX) process (Zhang et al., 2022a), provide high quality polysaccharides to favor paper, polymer, and sugar production after further processing. However, as those processes can result in severer and uncontrollable condensation issues on the fractionated lignin, they are considered less favorable in lignin-first application. Acid-catalyzed pretreatment process include a few different variations, e.g., organosolv(Xu et al., 2023), deep eutectic solvents (DES) (Han et al., 2022), and co-solvent enhanced lignocellulosic fractionation (Zhao et al., 2022). In this process, easily hydrolysable acetyl groups are removed from hemicellulose branches linking between cellulose and lignin, providing opportunities to better preserve useful C-O-C structure in lignin for depolymerization. We discussed mainly the acid-based pretreatment processes in this chapter.

3.1.1. Organosolv pretreatment

In organosolv pretreatment, numerous types of organic solvents have been used with acid catalysts for lignocellulosic biomass fractionation. The most applied organosolv include glycerol, ethylene glycol, butanol, methanol, ethanol, and other alcohols. Organic acids (such as oxalic acid, formic acid, acetylsalicylic, acetic acid, and salicylic acid) and mineral acids, such as HCl, H₂SO₄, and H₃PO₄, are often employed as catalysts to improve delignification and hemicellulose removal (Zhao et al., 2017). Other organosolv pretreatment approaches include the utilization of Tetrahydrofuran (THF), γ -valerolactone (GVL), and 2-methyltetrahydrofuran (2-MeTHF), which show even better delignification performance (Xu et al., 2023). Some organosolv can prevent lignin condensation. For example, formaldehyde protects α -hydroxyl group by forming 1,3-dioxane in C α and C γ , and diol forms ether linkages with C α (Dong et al., 2019b; Shuai et al., 2016a). Organosolv pretreatment has been studied extensively as shown in **Table S3-1** to **Table S3-3**. In those processes, diluted acid is applied to cleave the relatively ether linkages between hemicelluloses and lignin, and organosolv serves as a carrier of the fractionated lignin fragments. However, serious side effects have also been reported in these processes. Acid catalysts can also attack the β -O-4 linkage in lignin, an essential structure to produce valuable lignin monomers. This reaction would result

in the generation of reactive benzylic carbocation intermediates at the C α position of lignin (Li et al., 2018a). The intermediates could couple with other benzoic rings and form a highly condensed polymer with stable C-C linkage, which severely affects the productivity of both fermentable sugars and monolignols (Islam et al., 2020b).

3.1.2. Deep eutectic solvent (DES) pretreatment

DES is generally produced through mixing two components, i.e. hydrogen bond donor and hydrogen bond acceptor. Examples of HBD include amines, carboxylic acids, alcohol, and carbohydrates, while examples of HBA include quaternary salts like choline chloride (ChCl). HBA and HBD are solid forms at room temperature and become liquid when two reagents are combined. The eutectic temperature of various acidic DES can be found in a review conducted by Qin et al (Qin et al., 2020). DES has similar properties with ionic liquid (IL), but with greener properties, i.e. non-toxic, biodegradable, and biocompatible (Das et al., 2018; Hansen et al., 2021; Zhang et al., 2022b). Recent studies on acid-based DES pretreatment of hardwood, grass, and seed coats were compiled in **Table S3-1**. Like organosolv, DES can be classified into three pHs, i.e., acid, alkali, and neutral. DES was found to be a suitable solvent for pretreatment of lignocellulosic biomass due to its high lignin dissolution ability (Hansen et al., 2021). Interestingly, DES has also been utilized in RCF as a substitute solvent.

3.1.3. Pretreatment-Catalytic Transfer Hydrogenolysis (PT-CTH)

Lignin fractions derived from organosolv, or DES pretreatment exhibit considerable potential for value-added upgrading. Among various lignin valorization strategies, the PT-CTH process has emerged as a promising approach for the selective production of aromatic monomers. In this two-step route, lignin is first isolated using organosolv or DES-based pretreatment methods, both of which are effective in preserving a high proportion of β -O-4 aryl ether linkages the most labile and chemically accessible bonds within the lignin polymer. The preservation of these linkages is critical, as they serve as the primary cleavage sites in subsequent catalytic depolymerization.

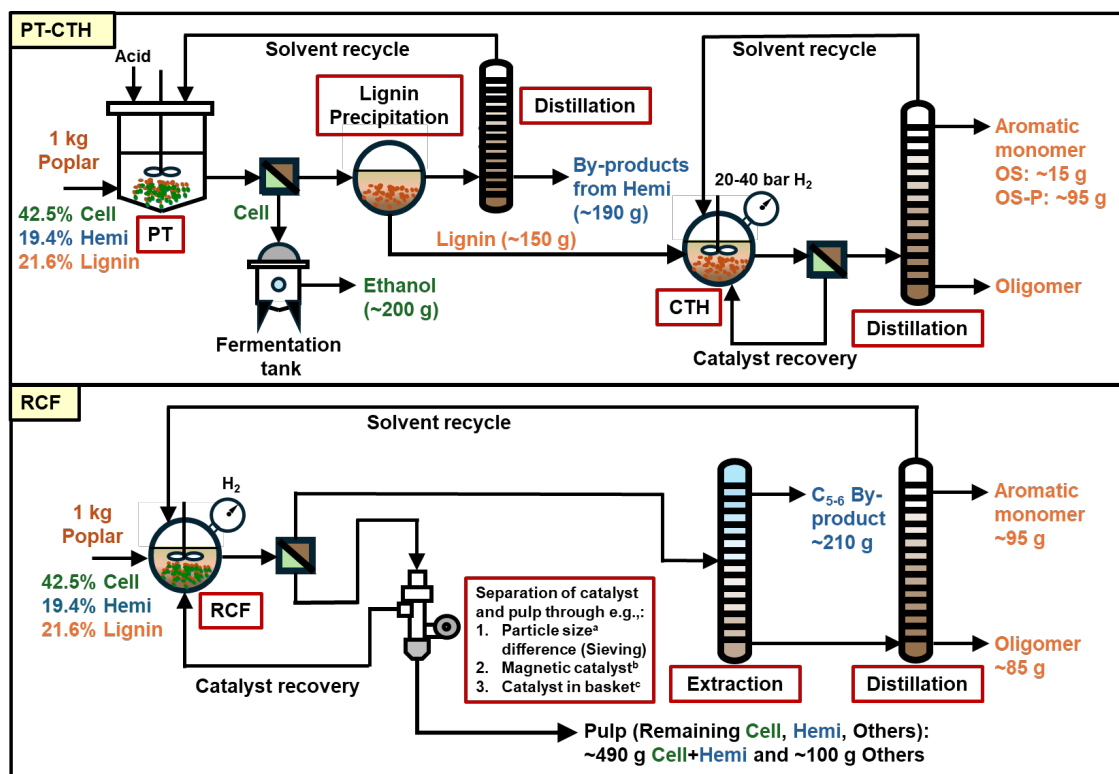


Fig. 3-2. Comparison of product distribution using PT-CTH and RCF strategies. PT-CTH yields ethanol, lignin, and hemicellulose by-products further converted to aromatic monomers and oligomers.

Following extraction, the lignin undergoes catalytic transfer hydrogenolysis under mild reductive conditions, typically using hydrogen-donor solvents such as isopropanol or ethanol and supported metal catalysts (e.g., Ru/C, Pd/C, or Ni-based systems). This process selectively cleaves the β -O-4 linkages while minimizing undesired condensation reactions, thereby enabling the efficient release of phenolic monomers with defined substitution patterns. The PT-CTH strategy thus provides a coherent platform for transforming structurally intact lignin into valuable aromatic building blocks, linking upstream biomass fractionation with downstream chemical valorization in a unified, scalable framework.

3.1.4. Reductive Catalytic Fractionation (RCF)

RCF is a promising approach for the conversion of biomass lignin into value-added aromatics. Lignin's physical location and activity in biomass composites makes it possible to extract and depolymerize it into phenolic monomers, dimers, and oligomers. Furthermore, cellulose may be retained in solid biomass for additional treatment. The performance of β -O-4 units cleavage and its stabilization are catalyst-specific (Liu et al., 2022). RCF consists of three significant stages: solve-thermal fractionation of lignin from biomass, depolymerization of the fractionated lignin into phenolic monomers, and monomers stabilization (Jindal et al., 2023). Studies have proposed that the fractionation primarily occurred through solvolysis (e.g. methanol in 200-250°C), while the depolymerization and monomer stabilization occurred through hydrogenation by catalysts (e.g., Ru/C) and H₂ gas (Huang et al.; Stone et al., 2018). Without the presence of reductive catalysts, β -O-4 cleavage leads to repolymerization and stable C-C bonds. After repolymerization, the condensed lignin is intended to be inert and incapable of being cleaved under RCF conditions (Klein et al., 2015). This section will first discuss the solvents effect of RCF, then evaluate the RCF performance on hardwood, softwood, grass, barks, and seed.

3.2. Comparison between organosolv and DES pretreatment

The performances of organosolv and DES pretreatment are shown in **Table 3-1**, over different process control parameters including reaction time, reaction temperature, solid/liquid ratio, delignification after PT, and monomer yield after PT-CTH. All the values are collected from existing literature (**Table S3-1** to **Table S3-3**). In summary, average reaction time of organosolv (1.77 h) is shorter than DES pretreatment (9.21 h), most likely due to higher average reaction temperature of organosolv (147.8 °C) compared with DES (112.1 °C). Organosolv generally utilizes mineral acids or Lewis acids, such as dilute H₂SO₄, HCl, and FeCl₃, and DES utilize organic carboxylic acids which decompose at lower temperatures, ranging from 125 °C and 235 °C (Delgado-Mellado et al., 2018; Rodriguez Rodriguez et al., 2019). Recent research also showed

that esterification between carboxylic acid and OH moiety, such as choline chloride, already happened at even lower temperatures, i.e., at storage temperature of 25 °C after 11 months, and during DES preparation through heating method at 80°C (Marchel et al., 2022; Rodriguez Rodriguez et al., 2019). Due to the possible decomposition and esterification, the DES may lose its acidity. This is possibly the reason behind the lower reaction temperature for acid-based DES pretreatment with an average of 112.1 °C.

Solid loading ratio (S:L) of organosolv (1:10) is generally higher than DES pretreatment (1:15), possibly because DES typically has higher viscosity than organosolv. Many ternary DESs have recently been developed, which uses ethylene glycol (EG) to decrease its viscosity while protecting Cα bond and preventing cleavage of β-O-4 (Han et al., 2022; Liu et al., 2021; Shen et al., 2023; Su et al., 2023b; Zhai et al., 2020). By using a ternary ChCl/EG/OA with mole ratio of 1:2:0.2, Shen et al. successfully pretreated corn stover with a solid loading of 50%, achieving 70% delignification, 95% cellulose recovery, 98% β-O-4 preservation, and 5.37% monomer yield (in g/g-lignin) in the subsequent CTH (Shen et al., 2023). From industrial and economic point of view, higher solid loading is preferred since it can reduce reactor volume (lower investment cost) and reduce energy consumption during solvent recovery. Solvent recovery is necessary since the pretreatment generally utilizes high amount of solvent, which can contribute to high cost if it is not recycled.

Table 3-1. Comparison of different plant species subjected to Deep Eutectic Solvent (DES) and Organic Solvent (OS) pretreatments

	Pretreatment	Temp(oC)	SL	Time(h)	Composition			MY	Ref
					CR	HY	LY		
Hardwood	OS	149.00	1.60	0.10	85.00%	73.00%	71.00%	10.40%	Table S2-1
	DES	95.70	6.93	0.06	92.69%	68.49%	57.56%	16.00%	Table S3
Softwood	OS	177.00	0.92	0.15	82.00%	84.00%	64.00%	3.68%	Table S2-2
	DES	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
Grass	OS	147.00	1.40	0.10	89.00%	71.00%	66.00%	3.20%	Table S2-2
	DES	108.08	7.86	0.11	79.74%	75.76%	64.77%	15.97%	Table S3
Bark	OS	145.00	2.40	0.09	84.00%	84.00%	48.27%	3.90%	Table S2-3
	DES	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
Seed	OS	112.50	3.80	0.09	68.50%	64.20%	34.10%	6.08%	Table S2-3
	DES	156.00	6.00	0.10	85.00%	59.00%	29.00%	10.13%	Table S3
Total	OS	146.10	2.02	0.11	81.70%	75.24%	56.67%	5.45%	Table S2
	DES	71.96	4.16	0.05	51.49%	40.65%	30.27%	8.42%	Table S2-1

SL=solid loading, Values are in unit of g-initial-dry-weight-of-biomass/volume of solvent;

CR = cellulose remain; HY = hemicellulose separation yield; LY = lignin yield; MY=Monomer yield.

All values are calculated based on the initial composition of biomass.

3.3. Delignification mechanisms in PT& RCF

Acid pretreatment and RCF represent two mechanistically distinct pathways for lignin removal from lignocellulosic biomass. Both approaches involve disruption of covalent linkages within the plant cell wall matrix, yet they differ markedly in terms of reaction environment, bond specificity, and the structural fate of lignin-derived fragments (**Fig. 3-3**). Acid pretreatment is typically conducted under hydrothermal conditions ($<160^{\circ}\text{C}$) using dilute inorganic acids (e.g., H_2SO_4 , HCl) or organic acids (e.g., maleic acid, oxalic acid).

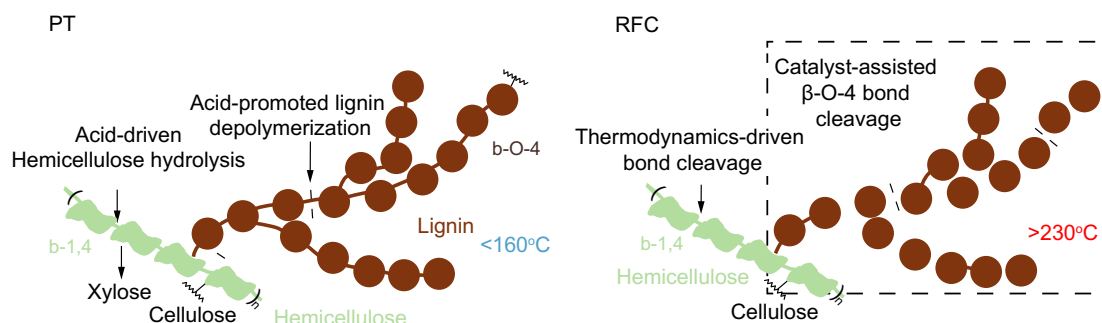


Fig.3-3. Schematic comparison of lignin removal mechanisms in acid pretreatment and reductive catalytic fractionation (RCF).

These acids catalyze the hydrolysis of ester and α -ether bonds that covalently link lignin to hemicellulose, particularly at the g -position of lignin side chains. This facilitates the partial release of lignin from the carbohydrate matrix. However, due to the non-selective nature of acid-catalyzed reactions, $\beta\text{-O-4}$ linkages critical to lignin depolymerization are also susceptible to cleavage. This undesired reaction pathway is often initiated by acid-promoted dehydration and protonation at the benzylic C_α position, resulting in carbocation intermediates or unsaturated C1-C3 fragments. These reactive intermediates readily participate in electrophilic substitution or radical-induced condensation, forming C-C coupled oligomers and reducing the $\beta\text{-O-4}$ content of the recovered lignin. Consequently, the isolated lignin is often highly condensed, chemically heterogeneous, and less amenable to subsequent catalytic depolymerization. RCF employs a controlled reductive environment to simultaneously solubilize and

depolymerize lignin. The process is carried out at elevated temperatures (200-260 °C) in the presence of hydrogen gas or hydrogen-donor solvents such as ethanol, isopropanol, or formic acid. Supported metal catalysts most commonly Ru/C, Pd/C, Ni/C, or Cu/Al₂O₃ enable selective hydrogenolysis of β -O-4 linkages while suppressing carbocation formation and condensation. The cleavage proceeds via metal-assisted hydride transfer or heterolytic C-O bond scission at the benzylic position, yielding monomeric phenolic compounds with fully preserved aliphatic side chains (e.g., propyl, ethyl). Unlike acid pretreatment, RCF avoids the formation of electrophilic intermediates, thereby maintaining structural integrity and enabling higher monomer yields with minimal byproduct formation. The distribution and separability of biomass-derived products are closely influenced by the processing pathway.

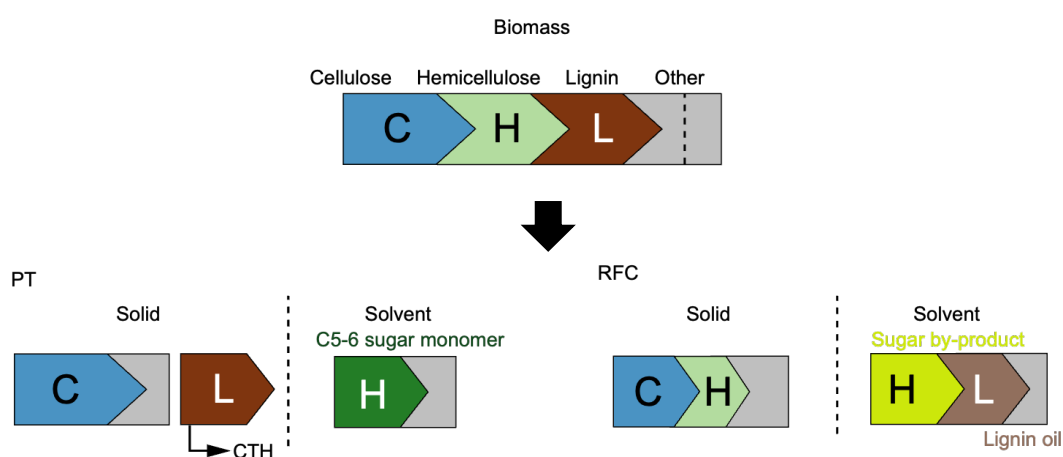


Fig.3-4. Phase distribution of biomass components after PT and RCF. PT enables full separation of cellulose, hemicellulose sugars, and lignin; RCF yields lignin oil and sugar by-products in the solvent phase.

In acid pretreatment systems, the application of dilute inorganic acids (e.g., H₂SO₄, HCl) or organic acids (e.g., maleic acid, oxalic acid) in aqueous-organic or deep eutectic solvent (DES) media facilitates the hydrolysis of hemicelluloses and the cleavage of lignin-carbohydrate ester linkages. As a result, hemicelluloses are converted into soluble mono- and oligosaccharides, while lignin, due to its solubility and reprecipitation behavior in the solvent system, can be selectively isolated in solid form. This enables physical separation of all three biomass fractions cellulose, hemicellulose, and lignin

into distinct phases. The cellulose-rich residue, largely free from catalyst contamination, is suitable for direct enzymatic hydrolysis or fermentation. Reductive catalytic fractionation (RCF) solubilizes both hemicellulose-derived saccharides and lignin-derived monomers into a common organic phase. This co-solubilization complicates product isolation due to overlapping polarities and reactivities. Furthermore, RCF employs finely powdered heterogeneous catalysts (e.g., Ru/C, Ni/C, Pd/C), which intimately associate with the solid cellulose fraction. These particles, although essential for catalytic activity, must be subsequently removed to enable further utilization of the cellulose. The mixed-phase nature of RCF hydrolysates and the catalyst entrapment in solids introduce technical challenges for downstream fraction separation and valorization, particularly when aiming for integrated biorefinery deployment.

3.4. Pretreatment performance in different biomass

After pretreatment, hydrogenolysis (CTH) or solvolysis is used to stabilize the lignin and oligomer ether bonds by cleaving them into monomers (Sun et al., 2018). Primarily, CTH was added to enable utilization of remaining lignin after pretreatment and thus, it aims to produce the highest monomer yield from lignin. Lignin fractionated from the biomass is suitable for depolymerization into phenolic monomers, if the pretreatment is efficient enough to extract lignin with maximum amount of native linkages, particularly β -O-4 and/or benzodioxane (Islam et al., 2021a).

CTH efficiency is dependent on reaction conditions like temperature, pressure, catalyst, and solvent. Cellulose recovery, hemicellulose dissolution, delignification, and monomer yield of PT-CTH are listed in **Table S3-2** for organosolv (from 32 studies). The monomer yield and carbohydrate recovery of PT-CTH of five lignocellulosic biomass are shown in **Fig. 3-5e**. Note that some of the monomer yields are not given in the literature and thus approximated from β -O-4 content using the equation from Phongpreecha et al. (Phongpreecha et al., 2017), and the carbohydrate recovery here includes the amount of carbohydrate remaining in the solid residue and the hemicellulose derivatives dissolved in the pretreatment liquor. Some calculations were

performed to convert the units to g/g-biomass and to determine the carbohydrate recovery (Appendix 5 in the Supplementary Material). Since the main objective of PT-CTH is whole-biomass utilization, we evaluated cellulose yield, hemicellulose yield, and monomer yield to evaluate the total utilization in each biomass (**Fig. 3-5**). Across all biomasses, PT exhibits high efficiency in extracting cellulose and hemicellulose, but CTH efficiency is still low, only utilizing 3.2-10.4% of lignin. Seed coats with C-lignin have the highest average monomer yield of 7.0% g/g-biomass, compared to all other biomass types. This is because benzodioxane is preserved during acid pretreatment, while β -O-4 content generally decreased. However, the PT-CTH data in this analysis only includes acid pretreatment. Under base pretreatment, benzodioxane structure may be partially cleaved (Wang et al., 2020b).

It is also observed that the range of monomer yields of seed coats is very broad, from 0.02-0.13 g/g-biomass. This is mainly because the monomer yield is highly dependent on C-lignin content, but C-lignin content in seed coats (in g/g-biomass) have large margins e.g., 4.6% in castor, 13% in vanilla, and 15.4% in Chinese tallow from the ^{13}C NMR results, as shown in Table 1. When the monomer yield is calculated based on g/g-C-lignin, the range of monomer yield is 0.71-0.84 (Li et al., 2018c; Su et al., 2023a; Su et al., 2021; Wang et al., 2025; Wang et al., 2020b; Wang et al., 2021; Yin et al., 2024). Across all biomass except seed coats, low monomer yields are observed after PT-CTH, but utilization of protection agents like formaldehyde during PT (OS-P) can significantly increase the monomer yield. For PT-CTH of barks, only the phenolic monomer yield from lignin is included.

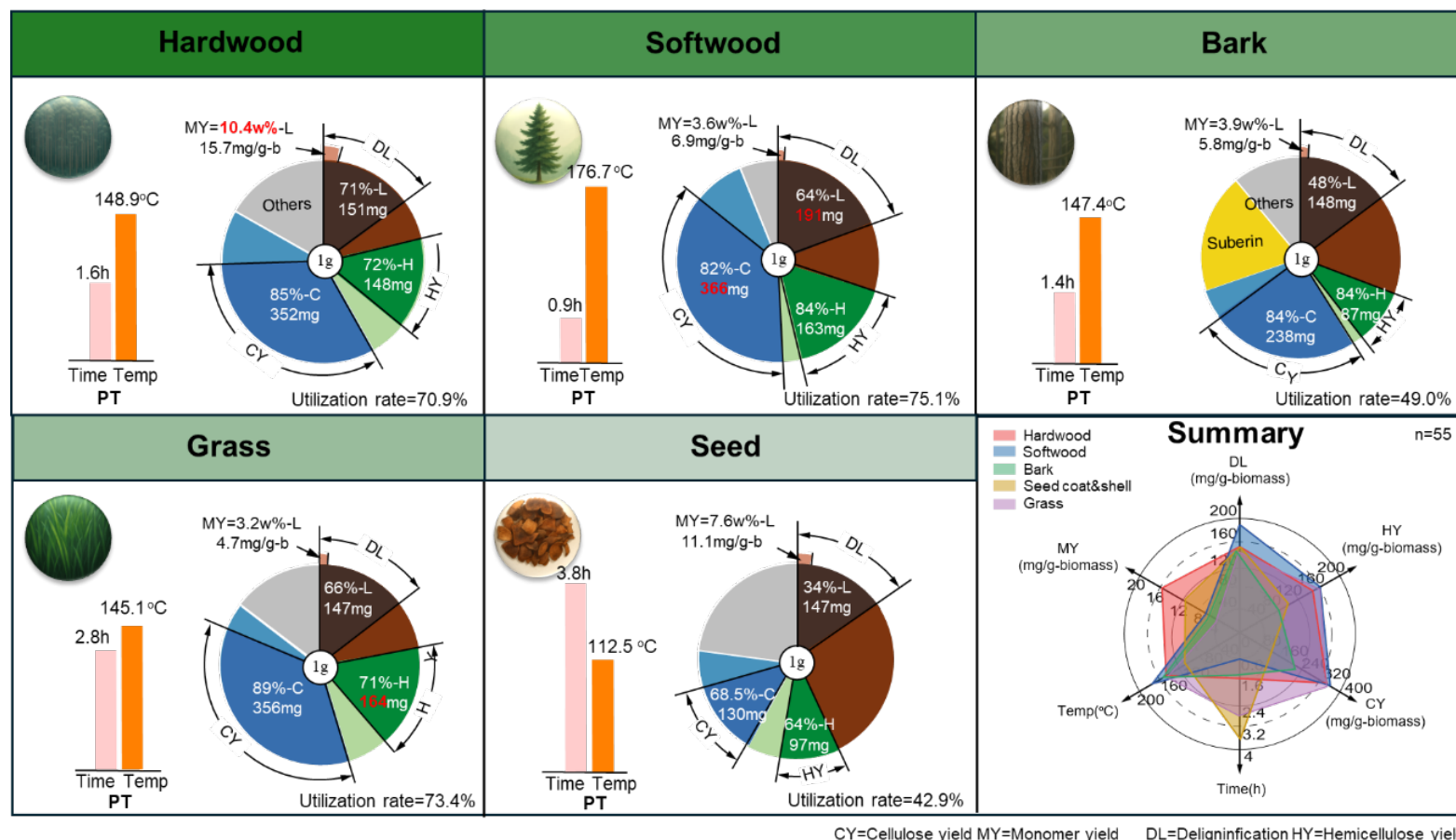


Fig. 3-5. Component separation performance of five representative biomass types (hardwood, softwood, bark, grass, and seed) under organosolv pretreatment. Pie charts illustrate the distribution and recovery of lignin (L), hemicellulose-derived sugars (H), cellulose (C), and lignin-derived monomer yield (MY) per 1 g of raw biomass. Bar charts present the corresponding organosolv pretreatment conditions. DES-based pretreatment showed comparable performance to organosolv methods, with detailed comparisons available in Table S4 (not shown here). The radar chart summarizes the biomass utilization potential and separation difficulty across different feedstocks. A total of $n = 55$ data points are included in this figure. Additional experimental conditions and Pretreatment performance, including solid loading, acid type, and acid dosage, are provided in Table S3-1 to 3.

3.5. Delignification kinetics

In PT-CTH and RCF, temperature plays a major role in influencing the reaction kinetics. For instance, Dong's study on diol pretreatment in eucalyptus at 160°C achieved 93% lignin removal, (Dong et al., 2019a) while Wang's study using *p*-Toluenesulfonic acid with poplar at 80 °C achieved only 54.8% delignification within one hour (Wang et al., 2019a). This discrepancy underscores the substantial impact of temperature on lignin removal efficiency. To better illustrate the comparison, pretreatment conditions were categorized into high-temperature (HT, Temp > 170 °C), medium-temperature (MT, 100 °C < Temp < 170 °C), and low-temperature (LT, Temp < 100 °C) groups (**Fig. 3-6**). RCF reactions are also compared here. RCF reactions are relatively standardized, typically conducted at 200-250 °C (HT) under 20-40 bar H₂, yielding consistent kinetic profiles.

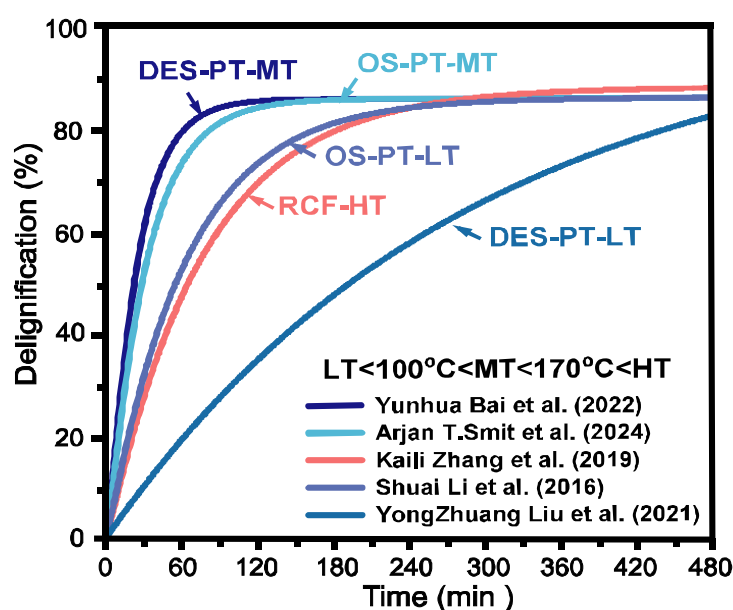


Fig. 3-6. Delignification kinetics in various pretreatment methods and RCF (Appendix 6 in the Supplementary Materials) (Bai et al., 2022; Liu et al., 2021; Shuai et al., 2016a; Smit et al., 2024; Zhang et al., 2019)

OS-PT-MT and DES-PT-MT pretreatment achieves the fastest lignin removal rates, consistently reaching over 80% removal within 60 minutes, as observed across multiple studies (Islam et al., 2024; Islam et al., 2021a; Ouyang et al., 2018). As

temperature decreases, lignin removal efficiency gradually declines. Nonetheless, the presence of acid in organosolv pretreatment enhances polysaccharide hydrolysis, allowing medium-temperature organosolv treatments to outperform RCF in lignin removal speed. When RCF operates under elevated pressures and temperatures, it typically requires around 3 hours to reach 80% lignin removal (Zhang et al., 2019). Notably, DES-PT-LT treatments exhibit slower lignin removal kinetics; for instance, the study by Yu et al. reported only 37% removal after 6 hours (Yu et al., 2018). In summary, medium-temperature pretreatment demonstrates the most rapid lignin removal kinetics, while low-temperature DES pretreatments offer potential for more controlled and gradual delignification.

3.6. Mass balance of PT-CTH and RCF

The mass balance analysis was conducted based on average value of the data for monomer yield, oligomer yield, carbohydrate retention, and by-product. To show clearer comparison, the results of the mass balance are illustrated in **Fig. 3.7**. Sankey diagrams were drawn to illustrate the fate of each component in hardwood after PT-CTH and RCF (**Fig. 3.7a** and **Fig. 3.7 b**). The mass balance analysis (**Fig. 3.7c**) showed that RCF of hardwood has the highest 9.3% monomer yield (in unit of %g/g-biomass), followed by barks (2.2% phenolic and 5.3% aliphatic), seed coats (6.1%), grass (4.5%), and softwood (4.4%). Similar patterns were observed for oligomer yield except for softwood obtaining high oligomer yield (in g/g-biomass) of 7.6%, comparable to the highest 8.3% in hardwood. Highest carbohydrate retention (in g/g-biomass) was observed for hardwood (47.6%), followed by softwood (46.5%) and grass (40.2%), while significantly lower values were observed for barks (22.2%) and seed coats (12.0%). This is mainly because barks and seed coats have inherently low carbohydrate content (24.1% cellulose and 11.4% hemicellulose for barks; 13.8% cellulose and 7.9% hemicellulose for seed coats). Barks have high lignin content left behind as by-product in the RCF solid residue (pulp), i.e., out of 36% lignin (g/g-biomass) in black locust bark, 21% lignin (g/g-biomass) was left behind in the bark residue after RCF (Vangeel et al.,

2019b). As for seed coats, the high lipid content was mostly extracted during the pre-extraction (e.g., with water/acetone) or converted into lipid derivatives after RCF, which is usually unquantified and unaccounted for as product. These are also the reasons behind the 66.7% unavailable by-products generated from RCF of seed coats (Fig. 4d). After considering all aspects of the mass balance in this study, it is concluded that hardwood is the best feedstock choice for RCF. Mass balance for PT-CTH shows that seed coats have the highest monomer yield of 7.0% g/g-biomass, compared to all other lignocellulosic biomass (**Fig.3-7c**). This is because benzodioxane is preserved during acid pretreatment, while β -O-4 content generally decreased. However, it is important to note the PT-CTH data in this analysis only includes acid pretreatment. Under base pretreatment, benzodioxane structure may be partially cleaved (Wang et al., 2020b). The mass balance of PT-CTH also does not include the PT-CTH with protection of β -O-4 (e.g., formaldehyde (Shuai et al., 2016a), propionaldehyde (Lan et al., 2018b), and p-Toluenesulfonic acid (Wang et al., 2019b)) where the monomer yield will be close to RCF value, as shown in **Fig. 3-3e**. Lastly, due to limited data, oligomers and aliphatic monomer from suberin are not available in the mass balance of PT-CTH.

RCF produces higher monomer yield, except for seed coats with benzodioxane, because β -O-4 is cleaved, and condensation might occur during pretreatment. Separation of the dissolved lignin and hemicellulose derivatives can simply be performed via lignin precipitation or extraction prior to the hydrogenolysis process in PT-CTH. This is followed by the efficient hydrolysis of cellulose in the solid residue of PT-CTH into glucose. Due to these reasons, the C5-6 sugar derivatives are considered as utilizable in PT-CTH, resulting in higher total utilizable products in PT-CTH. On the other hand, C5-6 sugar derivatives are treated as a by-product (usable with limitations in **Fig. 3-7d**) in RCF, because it is required to separate the derivatives from the lignin oil first before it can be utilized. Nevertheless, it is important to note that although most RCF studies do not evaluate the potential utilization of C5-6 sugar derivatives, some studies have successfully extracted and utilized these derivatives from RCF (Brienza et al., 2023; Brienza et al., 2022; Bugli et al., 2024; Nicolai et al., 2024; Renders et al., 2018; Van den Bosch et al., 2015).

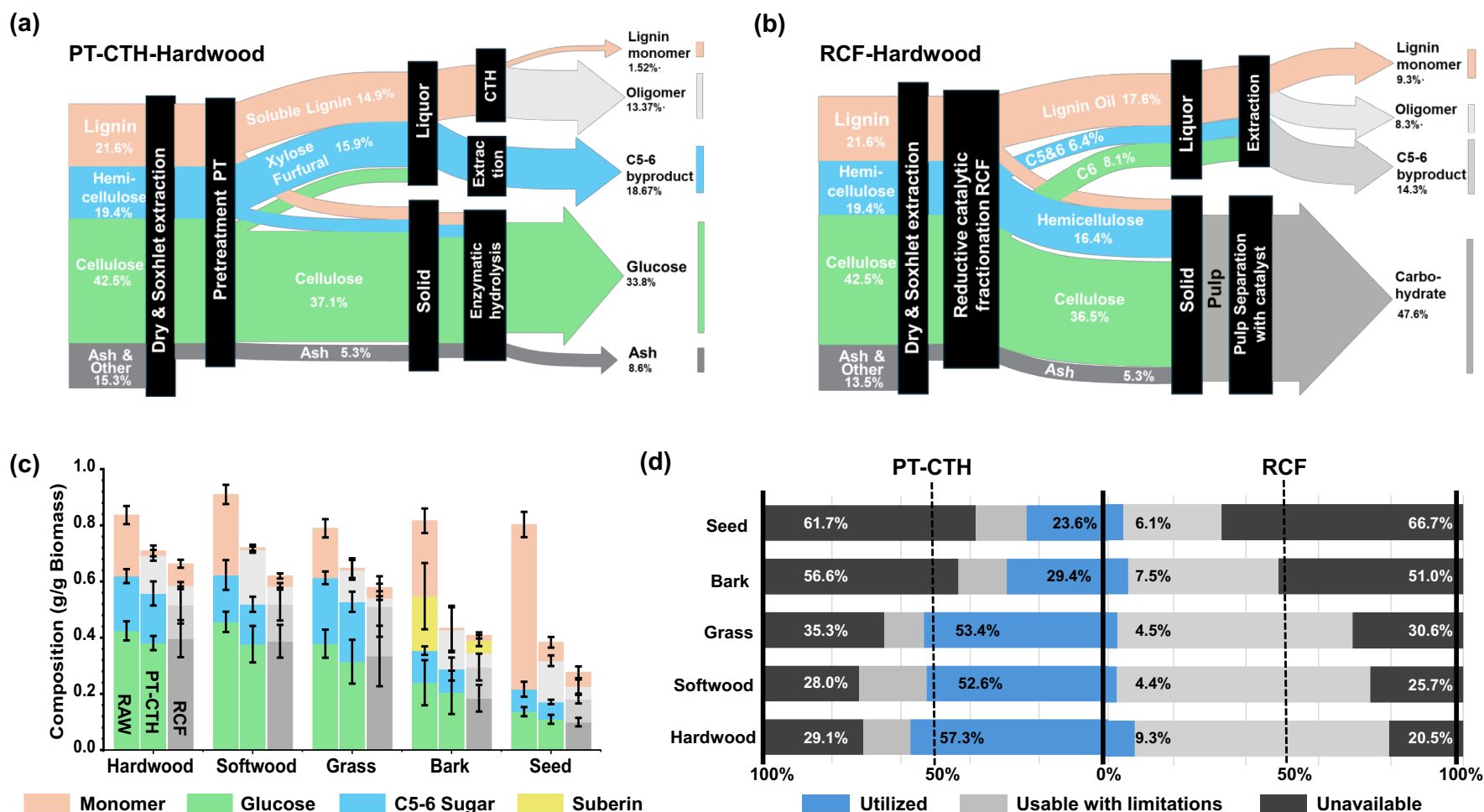


Fig.3-7. Sankey diagram depicting mass balance of (a) PT-CTH of hardwood and (b) RCF of hardwood. (c) Average values of initial composition of biomass (Raw), products composition after PT-CTH, and products composition after RCF. For the initial composition of biomass (Raw), the four components are the lignin (converted to monomer), hemicellulose (converted to C5 sugar), cellulose (converted to glucose and C6 sugar), and suberin (converted to aliphatic monomer). (d) Average values of utilizable products (utilized), possibly usable products (usable with limitations), and lost by-products (unavailable) after PT-CTH and RCF. Note that the mass balance does not include the PT-CTH with protection agents (i.e., OS-P) to better represent the general PT-CTH in the literature. The mass balance is calculated based on Table 3 for RCF and the average values obtained from Table S3 2-3 for PT-CTH. Calculation methods for the mass balance are shown in Appendix 5 in the Supplementary Materials.

3.7. Suggestion for different biomass pretreatment

Due to the unique chemical composition of each biomass, biomass potential for monomer production was evaluated for each species. Based on the component-specific product demands, current pretreatment approaches can be broadly classified into three strategic categories: 1) fractionation-based pretreatment, primarily designed for the separation and utilization of carbohydrates, without consideration of lignin valorization or preservation; 2) pretreatment followed by catalytic transfer hydrogenolysis (PT-CTH), which seeks to extract lignin while preserving its cleavable motifs for subsequent depolymerization into monomers; 3) direct reductive catalytic fractionation (RCF), which integrates lignin-first depolymerization into a one-step solvolysis and catalytic process.

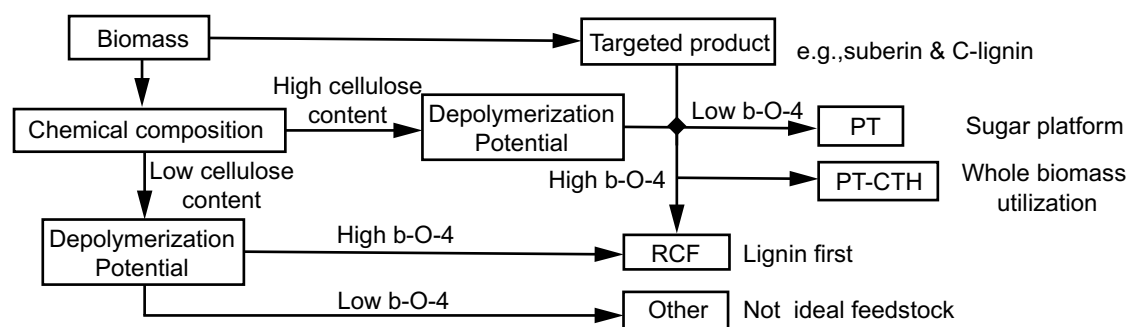


Fig.3-8. Decision framework for biomass processing based on cellulose content and lignin depolymerization potential.

The first category includes conventional processes such as kraft pulping and alkali-based treatments, which aim to remove lignin and facilitate cellulose and hemicellulose conversion. These methods are highly effective for carbohydrate recovery but often result in heavily condensed lignin fractions that are unsuitable for high-value aromatic monomer production. As such, this pathway is most appropriate when lignin valorization is not prioritized and carbohydrate conversion is the main objective. The PT-CTH pathway is appropriate when lignin contains a substantial proportion of cleavable β -O-4 linkages and exhibits relatively low condensation tendency. In such cases, acid pretreatment using organosolv or DES systems can selectively solubilize or precipitate lignin while preserving depolymerization-relevant motifs. The extracted

lignin is subsequently subjected to catalytic transfer hydrogenolysis under mild reductive conditions, enabling monomer production with minimal C-C coupling. Biomass types such as softwoods or selected grasses that exhibit moderate β -O-4 abundance but stable lignin networks benefit from this two-stage process. In contrast, the RCF route is preferable when β -O-4 content is high and spontaneous lignin depolymerization can be achieved in situ. Under reductive conditions (e.g., using Ru/C or Ni/C catalysts and hydrogen donors such as isopropanol), lignin is simultaneously extracted and cleaved into phenolic monomers. This one-pot method avoids the formation of reactive intermediates and suppresses recondensation, making it particularly suitable for hardwoods (e.g., poplar, birch) and herbaceous crops (e.g., miscanthus) with high carbohydrate availability and accessible lignin architectures.

However, for biomass types characterized by low cellulose content and highly condensed or structurally inert lignin, such as seed coats, shells, and some barks neither pathway is optimal for monomer production. These feedstocks may instead be better valorized via alternative strategies, such as suberin extraction, lipid recovery, or thermochemical conversion routes. To support systematic selection of pretreatment strategies, a decision tree framework (**Fig. 3-8**) was developed based on lignin reactivity and carbohydrate content. To facilitate the strategic selection of pretreatment routes, a two-dimensional compositional map (**Fig.3-9**) was constructed to visualize the relationship between biomass feedstock properties and their valorization potential. The x-axis reflects lignin depolymerization potential incorporating both β -O-4 ether abundance and overall lignin content while the y-axis indicates total carbohydrate content, primarily cellulose. This framework divides the biomass samples into four quadrants, each corresponding to distinct valorization strategies. In the first quadrant (upper right), feedstocks such as poplar, birch, eucalyptus, aspen, and miscanthus possess both high β -O-4 rich lignin and abundant carbohydrates. These species are ideal for PT-CTH, which combines controlled lignin extraction with catalytic monomer production, while enabling downstream sugar valorization. The second quadrant (upper left) includes biomass like candlenut, castor, and camellia seed coats, which are rich in depolymerizable lignin but low in carbohydrates. These are more suitable for direct RCF,

prioritizing aromatic monomer yield over sugar recovery.

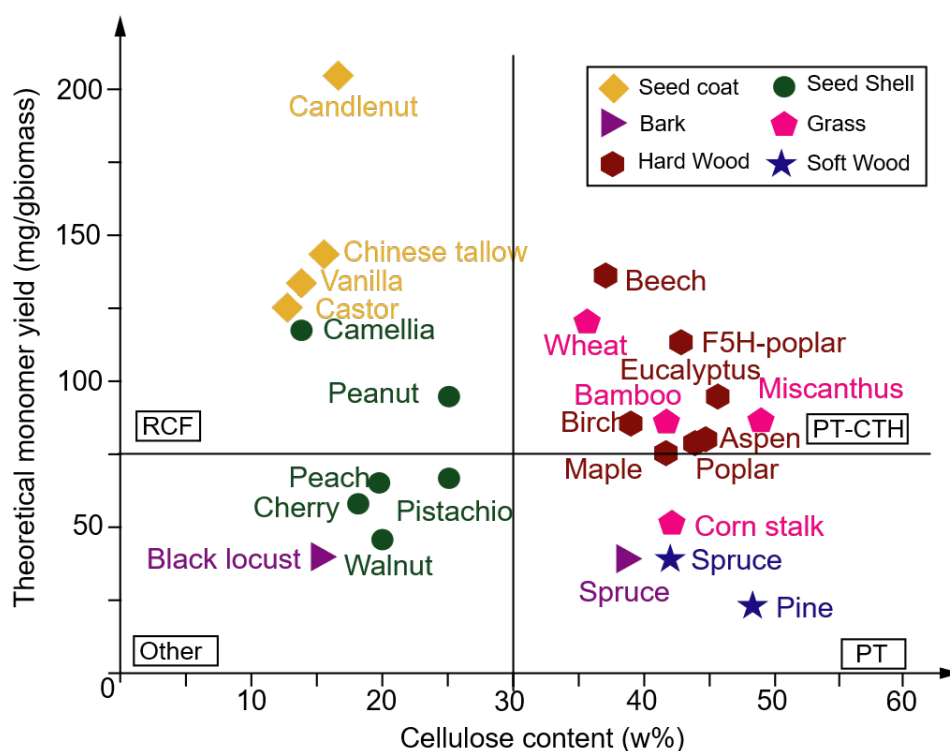


Fig.3-9. Two-dimensional compositional map showing the relationship between lignin depolymerization potential (x-axis) and cellulose content (y-axis) across diverse biomass feedstocks. Samples are grouped into four quadrants, each suggesting a preferred valorization route: PT-CTH for carbohydrate- and β -O-4-rich species (Quadrant I, e.g., poplar, miscanthus); direct RCF for lignin-rich, low-carbohydrate feedstocks (Quadrant II, e.g., candlenut, castor); thermochemical or extractive routes for low-reactivity, low-carbohydrate biomass (Quadrant III, e.g., walnut shell, black locust); and cellulose-focused pretreatment for softwoods with recalcitrant lignin (Quadrant IV, e.g., pine, spruce). Data can be referred to **Table S1-1**

In the third quadrant (lower left) lie feedstocks such as black locust bark, walnut shell, and peanut shell, characterized by more C-C lignin structures and low cellulose content. These materials are not well-suited for lignin-first strategies and may be better valorized through thermochemical routes or targeted extraction of suberin and lipophilic compounds. Finally, the fourth quadrant (lower right) contains softwoods such as pine and spruce, which are rich in cellulose but contain structurally recalcitrant lignin

structure with limited depolymerization potential. These species are appropriate for traditional pretreatment methods, especially when targeting cellulose utilization in pulp, paper, or sugar platforms. Notably, despite the low monomer yield from softwood lignin, the resulting phenolics often exhibit high chemical uniformity, which may be advantageous in specialized applications. By integrating structural reactivity and carbohydrate availability, this quadrant-based framework offers a rational guide for matching pretreatment strategies with biomass composition, supporting the development of efficient and feedstock-specific biorefinery schemes.

CHAPTER 4. METHODOLOGY

This chapter summarizes all analytical techniques and experimental procedures employed in this thesis. In addition, the PDF modeling approaches used to evaluate pretreatment performance and lignin quality calculation are also presented in this section.

4.1. Biomass feed stock and chemicals reagents

The *Camellia oleifera shell* (COS) samples were harvested from the tea oil camellia plantation operated by Guangdong Shanma Agricultural and Forestry Development Co., Ltd. (Lianzhou, China). The Unlabeled eucalyptus chips were provided by the Leizhou Forestry Bureau (Guangdong, China). The chips were manually debarked, Wiley-milled, and screened to collect fractions smaller than 2 mm. The resulting material was Soxhlet-extracted with ethanol/hexane (2:1, v/v) for 48 hours to remove extractives. The extraction-free samples were then dried in an oven at 80°C for 48 hours prior to further use. The uniformly ^{13}C -enriched stems (97% ^{13}C) of *Eucalyptus grandis* (16weeks old) were obtained from IsoLife (The Netherlands) for solid-state NMR analysis. The plants were cultivated from seeds under controlled conditions. Upon harvesting, shoots were separated into leaves and stems. The stems were freeze-dried, longitudinally split, sectioned, and carefully debarked by incising the bark and peeling it from the xylem at the cambium interface. Debarked stems were coarsely cut with scissors and gently ground using a mortar and pestle into fragments a few millimeters in size. Ball milling was intentionally avoided to preserve the native ultrastructure of the biomass. All reagents and chemicals were of analytical grade (Sigma-Aldrich, MO, USA; and J&K Acros Organics, Beijing, China).

4.2. Chemical composition analysis

The compositional analysis of the extractive-free COS/Eucalyptus and the pretreated samples was performed according to the NREL protocol (Sluiter et al., 2008). In short, 150 mg of extractive-free biomass samples were placed in 50 mL Falcon tubes and mixed with 1.5 mL of 72% sulfuric acid. The suspension was continuously stirred at 30 °C for 1 hour, and then 42 mL of distilled water was added to the solution before autoclaving at 121°C for 1 hour. The solutions were then allowed to cool to room temperature and neutralize with barium carbonate. The sugar composition was determined using high-performance liquid chromatography (HPLC, Shimadzu) equipped with a CHO-782 column (Supelcosil™). The remaining solutions were filtered using pre-weighed glass crucibles, washed with distilled water, and oven-dried to determine the lignin content. The solid residues were placed in a furnace at 575°C for 6 to 8 hours to determine the ash content. In addition to the NREL method, EA was conducted to determine the chemical formula of the solid samples using a Perkin Elmer 2400 Series II CHNS/O Elemental Analyzer (Waltham, MA, USA). GPC was performed to determine the molecular weight distribution of the dissociated components in the pretreatment spent liquor. Briefly, the purified solid lignin sample was dissolved in THF (0.002 mg/ml) and analyzed by HPLC (Shimadzu) equipped with an Agilent GPC/SEC column (PLaquagel-OH MIXED-M 8 µm 300 × 7.5 mm) at a flow rate of 1.0 mL/min and with detection by an RID detector.

4.3. Lignin analysis

4.3.1. Two-dimensional HSQC

2D-¹³C-¹H heteronuclear single quantum coherence (HSQC) NMR spectra were measured to elucidate the molecular-level structure of the pretreated lignin samples. Approximately 50 mg of lignin samples were dissolved in 0.5 mL of DMSO-d₆ for each analysis. The experiments were conducted on a JEOL ECZR 500 MHz spectrometer equipped with a 5 mm ROYAL probe. The experimental setup followed JEOL's standard pulse sequence with the following parameters: a spectral width of 11 ppm in

F2 (^1H) with 1024 data points and 190 ppm in F1 (^{13}C) with 256 data points; 64 scans and a 1.5 s recycle delay. The volume integration of contours in the HSQC spectra was performed using JEOL's Delta 5.0 software.

4.3.2. Gel Permeation Chromatography (GPC)

Gel permeation chromatography was performed as in our previous work (Dong et al., 2019a). Briefly, purified solid lignin sample was dissolved in THF (0.002mg/ml). The sample were analyzed by HPLC (Shimazu) equipped with a Agilent GPC/SEC Columns (PL aquagel-OH MIXED-M 8 μm 300 * 7.5mm) at a flow rate of 1.0ml/min, Room Temp, and detection by RID detector.

4.3.3. Fourier-transform infrared spectra (FTIR)

FTIR spectra of lignin were obtained using a PerkinElmer spectrometer equipped with a diamond attenuated total reflectance (ATR) unit. Finely milled lignin powders were placed directly on the diamond surface, and spectra were recorded from 4000 to 400 cm^{-1} with 12 scans at a resolution of 4 cm^{-1} . The infrared beam interacted with the sample via an evanescent wave through the diamond crystal. Measurements were conducted at room temperature and 34% relative humidity.

4.3.4. Quantitative solid-state NMR (Multi-CP)

The biomass/lignin sample before and after treatment weighing 30-70 mg were filled into 3.2-mm zirconium rotors and analyzed using 400 MHz (9.4 Tesla) and 600 MHz (14.1 Tesla) Bruker Avance Neo spectrometers. The Multi-CP pulse sequence was applied to collect quantitative spectra under 15 kHz magic-angle spinning (MAS) at 296 K. Each experiment employed 7 CP blocks (1.1 ms contact time each), with a 0.9s z-filter delay between blocks to allow for repolarization(Johnson & Schmidt-Rohr, 2014). The recycle delay was set as 2s. The ^{13}C chemical shifts were calibrated by the adamantane CH₂ peak at 38.48 ppm on the TMS scale. Spectrum deconvolution was carried out using the DMfit software (Massiot et al., 2002).

4.3.5. Quantitative ^{31}P NMR spectroscopic characterization

To quantify hydroxyl groups in the lignin samples, phosphorylation of each sample was carried out using 2-chloro-4,4,5,5-tetramethyl-1,3,2-dioxaphospholane (TMDP) in a solvent mixture of pyridine/ CDCl_3 (1.6/1.0, v/v). The stock solution was prepared with pyridine/ CDCl_3 (500 μL) and 5mg/mL internal standard (cyclohexanol). Approximately 20.0mg of lignin sample was dissolved in the stock solution and phosphorylated with 100 μL of TMDP. The ^{31}P NMR spectra were recorded for 32 scans on the 500 MHz JEOL ECZR NMR spectrometer using an inverse-gated decoupling pulse sequence, and with 25s delay. All reported chemical shifts were referenced to the internal standard cyclohexanol-OH peak at 145 ppm. The quantification of hydroxyl groups was referred to those of the internal standard.

4.4. Scanning electron microscopy (SEM)

Freeze-dried lignin samples were examined using a TescanMAIA3 ultra-high-resolution SEM. Samples were affixed onto aluminum stubs with conductive carbon tape to minimize charging effects, followed by a thin gold coating applied via sputter coating. Micrographs were acquired at an accelerating voltage of 25 kV

4.5. X-ray Diffraction (XRD)

The crystalline or carbon structures of the substrates were analyzed using a Bruker D8-Advance diffractometer. Measurements were performed with $\text{Cu K}\alpha$ radiation under operating conditions of 40 kV and 30 mA. Data were collected over a 2θ range of 10-50° at a step size of 0.020° and a scan rate of 3°/min. The crystallinity index (CrI) was calculated from diffraction peak intensities using the formula:

$$\text{CrI}(\%) = \frac{I_c - I_{am}}{I_c} \times 100$$

where I_{am} is the intensity of the amorphous peak at $2\theta \approx 18.5$, and I_c is the maximum crystalline peak intensity near $2\theta \approx 22.5$.

4.6. Monomer analysis

The monomers were analyzed using a Gas chromatography-flame ionization detection/mass spectrometry (GC-FID) (Agilent 7890B series) equipped with an HP-5 column and a flame ionization detector (FID). Briefly, 1 mL of the resultant solution was mixed with 100 μ L of internal standard (8 mg of decane in 5 mL THF) and analyzed directly in the GC-FID. The injection temperature was set to 300°C. The column temperature program was 40°C (3 min), ramping at 30°C/min to 100°C, ramping at 40°C/min to 300°C, and holding at 300°C (5 min). The detection temperature was 300°C. The sensitivity factors of product were estimated by the effective carbon number (ECN) due to a lack of commercial standards. Identification of monomer peaks in the GC-FID spectra was performed by an Agilent 7890B GC equipped with a HP5-MS capillary column followed by an Agilent 5977A series mass spectrometry (MS) detector. The operating conditions for GC-MS were the same as those in GC-FID as the peaks of both chromatograms appear in the same order.

4.7. Solid-state NMR spectroscopy

Solid-state NMR spectroscopy was employed to elucidate the structural organization and dynamics of lignin-carbohydrate assemblies in plant cell walls. Uniformly ^{13}C -labeled and unlabeled stem materials were packed into MAS rotors and analyzed using Bruker 400 MHz and 600 MHz spectrometers, equipped with 4 mm and 3.2 mm MAS probes, respectively. Unless otherwise specified, all experiments were performed under MAS frequencies of 10-14 kHz at 293 K. ^{13}C chemical shifts were externally calibrated to tetramethylsilane (TMS), using the adamantane secondary reference (38.5 ppm). During acquisition, the RF field strengths were set at 80-100 kHz for ^1H decoupling, 62.5 kHz for the ^1H contact pulse in CP experiments, and 50-62.5 kHz for ^{13}C pulses.

4.7.1. One-dimensional direct polarization & cross-polarization ^{13}C spectra

One-dimensional ^{13}C NMR spectra were recorded under different polarization

conditions to investigate molecular structure and dynamics. The ^1H – ^{13}C cross-polarization (CP) method was applied to emphasize rigid domains, given that CP transfer efficiency is influenced by both protonation level and molecular mobility; thus, compositional information obtained from CP should be interpreted qualitatively. For quantitative analysis, direct polarization (DP) with an extended recycle delay of 35s was employed to ensure complete ^{13}C magnetization recovery, allowing reliable detection of all carbons. To enhance signals from more mobile fractions, the DP experiment was also conducted with a shortened recycle delay of 2s, which permits preferential recovery of carbons with rapid ^{13}C – T_1 relaxation. CP acquisitions were performed with a 1ms contact time using a ramped ^1H pulse and a constant-amplitude ^{13}C pulse to enhance signals from immobile constituents such as lignin and crystalline polysaccharides. Spectral deconvolution was performed using DMfit software (Massiot et al., 2002).

4.7.2. 2D INADEQUATE, PDSO & Gated-DARR ^{13}C Spectroscopy

Two-dimensional NMR spectra were acquired using CP-based methods with phase-sensitive quadrature detection (States acquisition). Chemical shift assignments were made using double-quantum-single-quantum (DQ-SQ) correlation spectroscopy via the refocused Incredible Natural Abundance Double QUAntum Transfer Experiment (INADEQUATE) pulse sequence. The DQ–SQ experiment improves spectral resolution by eliminating diagonal resonances and redistributing isotropic chemical shift information into the indirect DQ dimension, where the chemical shift corresponds to the sum of the isotropic shifts of two coupled ^{13}C nuclei. Because coherence transfer in INADEQUATE relies on J-coupling, the method detects exclusively directly bonded carbon pairs, thereby allowing unambiguous assignments of carbon–carbon connectivities. To highlight lignin aromatic signals, which are typically masked by intense carbohydrate resonances, a dipolar-gated sequence was applied. This approach suppresses signals from protonated carbons, effectively reducing carbohydrate contributions while retaining signals from non-protonated aromatic carbons. Subsequently, homonuclear ^{13}C – ^{13}C correlation spectra were acquired using Proton-Driven Spin Diffusion (PDSO) and Dipolar-Assisted Rotational Resonance

(DARR). DARR experiments used a 100 ms mixing period with low-power proton irradiation, optimizing detection of intramolecular contacts within ~ 6 Å. PDS experiments, which do not require proton irradiation, employed a 1 s mixing period to enable detection of long-range intermolecular interactions up to ~ 10 Å via relayed spin diffusion. Gated-DARR and PDS experiments were conducted to provide spatial constraints for lignin and its interactions with surrounding cell wall components.

4.8. PDF Modeling and calculation

The basic concept of the PDF model, including the simulated building block chemicals, their linkages, fates, and reaction mechanisms has been demonstrated in **Fig.5-1**. Three hypothetical reactions are of concern: (a) hydrolysis of hemicellulose in dilute acid; (b) dissolution of lignin in the organosolv; and (c) damage and re-precipitation of lignin due to condensation under over-reacted condition (note that all reactions can happen simultaneously during pretreatment).

The model is validated based on the following sequence: [1] quantification of the pretreatment severity functions; [2] dynamic simulation to quantify residual lignin and hemicellulose under gradually increased pretreatment severities; and [3] simulation of the dynamic changes in aromatic inter-linkage depolymerization rates to calculate the monolignol productivity and yield. The specific procedures and functions are summarized as follows.

4.8.1. Pretreatment depolymerization factor (PDF)

The PDF model incorporates the basic concept of the Arrhenius first-order kinetics, which facilitates a rigorous analysis of the hydrolysis process under various pretreatment conditions (Sella Kapu & Trajano, 2014). It includes four principal factors of the organosolv pretreatment process as follows:

$$\ln \left[\frac{PDF}{C_A \cdot t} \right] = \alpha + \frac{E}{R \cdot T} + \beta \cdot C_S + \gamma \cdot C_A \quad [1]$$

where C_A is the acid concentration (mM); T is the reaction temperature (K); t is the

reaction time (h); E is the apparent activation energy; R is the ideal gas constant (8.314 J/mole K); CS is the organosolv concentration (%); α , β , and γ are ionization potential factors across various solvents.

4.8.2. Hemicellulose hydrolysis

The PDF is associated with various hemicelluloses in the pretreated substrate over the original biomass (X_r) to simulate their hydrolysis rates, based on the following function:

$$X_r = \sum_{i=1}^n k_i \cdot X_i \cdot \left[\theta_i^H \cdot e^{-f_i^H \times PDF} \right] \quad [2]$$

Where k_i and X_i are the specific hemicellulose hydrolysis kinetics and the mass fraction of selected hemicellulose in pretreatment, i ; θ_i^H and f_i^H are the theoretical yield and normalized hydrolysis rate of the concerned hemicelluloses, respectively. In our experiments, X_i was determined by chemical composition analysis; θ_i^H of five hemicelluloses ($n=5$), including Xylan^{2f} (Xn^{2f}), Xylan^{3f} (Xn^{3f}), arabinose (Ara), mannose (Man), and galactose (Gal), was quantified by solid-state NMR; and f_i^H was analyzed by High-Performance Liquid Chromatography (HPLC) and compared with the results of solid-state NMR.

4.8.3. Lignin fractionation

PDF was further derived to simulate the status of lignin dissociation during pretreatment. In this process, the lignin removal efficiency (L_e) was simulated by the following equation:

$$L_e = 1 - L_r = L_p + L_h - L_c \quad [3]$$

where L_r is the mass fraction of the residual lignin over the lignin content in the raw

biomass; L_p represents the lignin dissolved in the organosolv due to physical separation; L_h is the physical mitigation of lignin associated with the hydrolysis of hemicellulose (Liu et al., 2016); and L_c is the condensed lignin. L_p is calculated by the following equation:

$$L_p = k_p T + \delta \quad [4]$$

where k_p is the equilibrium factor describing the maximum capacity of lignin separation without acid; δ is the dissociation factor determined by experiments. L_h is calculated as follows:

$$L_h = X_r \cdot f_X^H \quad [5]$$

where f_X^H is the mass fraction of lignin associated with specific hemicellulose i. It can be determined by pretreatment experiments at mild severity.

Finally, L_c can be calculated by the total dissolved lignin content in the pretreatment liquor times a solvent-specific first-order reaction constant as expressed as follows:

$$L_c = k_c \cdot (L_p + L_h) \cdot e^{k_c \cdot PDF} \quad [6]$$

where k_c is the condensation factor representing the limited non-soluble lignin in the dual solvent system. Notably, Eq. [6] should not be a linear function, and this expression was adopted in this study for its simplicity. The designed experimental conditions allowed only limited condensation, and the lignin condensation behavior under extreme severities is beyond the scope of this work.

4.8.4. Lignin depolymerization

In addition to pretreatment, PDF was also integrated into a lignin depolymerization model to simulate the lignin monomer yield (M_m) based on the following equations:

$$M_m = M_m^{Max} \cdot R_m \quad [7]$$

where M_m^{Max} is the theoretical productivity of the monolignol of the selected biomass, and R_m is the normalized lignin monomer yield (%). The maximum monolignol productivity can be calculated from the total β -O-4 linkage in the lignin, which can be calculated as follows:

$$M_m^{Max} = \left(C_{\beta-O-4}^{PDF=0} \right)^2 \times m_{Lignin} \quad [8]$$

where $C_{\beta-O-4}$ is the abundance of β -O-4 linkage (%); and m_{Lignin} is the total mass of lignin in the biomass, which are both biomass-specific constants and can be determined experimentally. Eq.[7] facilitates the computation of potential monomer yields while attenuating the boundary effect for biomass samples with sufficient β -O-4 linkage (Phongpreecha et al., 2017). Finally, lignin monomer yield over pretreatment conditions can be calculated as follows:

$$R_m = \left[e^{-k_d \times PDF} \right]^2 \times L_e \quad [9]$$

where k_d is the solvent-specific depolymerization constant derived from control experiments. The model and the related parameters were validated based on the results of laboratory experiments.

CHAPTER 5. POLYSACCHARIDE HYDROLYSIS AND LIGNIN DISSOCIATION KINETICS

ABSTRACT

The advancement of whole-biomass-utilized biorefinery technologies offers a promising pathway to valorize all major components of lignocellulosic biomass into high-value products, thereby contributing to carbon neutrality. The pretreatment step is critical for effectively utilizing lignocellulosic biomass, yet the complexity of hemicellulose hydrolysis kinetics remains a key challenge. This chapter systematically investigates the hydrolysis kinetics of hemicelluloses from COS, a highly recalcitrant lignocellulosic biomass rich in hemicellulose and lignin. A comprehensive set of 66 organosolv-acid pretreatment experiments was conducted by varying acid types, concentrations, temperatures, and durations. Advanced SS-NMR spectroscopy, coupled with spectral deconvolution techniques, allowed the quantitative identification of distinct hemicellulose conformations, particularly the two-fold (Xn^{2f}) and three-fold (Xn^{3f}) xylan structures. The experimental results revealed that the hydrolysis kinetics of Xn^{3f} were significantly faster and exhibited a stronger correlation with lignin dissociation compared to other hemicellulose fractions. A process control model, termed the Pretreatment Depolymerization Factor (PDF), was developed to predict lignin yield and structural integrity. This model reduces energy input, preserves lignin integrity during separation, and ultimately facilitates the efficient conversion of whole lignocellulosic biomass. Importantly, the model was further validated across diverse biomass feedstocks, including pine, eucalyptus, and bamboo, consistently achieving R^2 above 0.91. The PDF model effectively simulates hemicellulose degradation behavior, thereby providing mechanistic insights into biomass recalcitrance and guiding the optimization of biorefinery processes.

5.1. Introduction

The advancement of biorefinery technologies converting plant biomass into sustainable products, such as aromatic chemical building blocks and biofuels, for minimizing our dependence on fossil fuels, has played a critical role in achieving carbon neutrality (Lee et al., 2022; Liao et al., 2020). However, the utilization of non-food-grade lignocellulosic biomass remains technically and economically challenging. Plant residues, or lignocellulosic biomass, are highly recalcitrant bio-composites consisting of a wide variety and amounts of cellulose (10-55%), hemicelluloses (10-35%), and lignin (15-38%) (Vassilev et al., 2010). The efficient utilization of each of those components is essential to maximize the value of the natural resource while minimizing the wastes (Liu et al., 2022; Zhang et al., 2023).

The biorefinery process is therefore complex, comprising several critical unit processes. This includes pretreatment, enzymatic hydrolysis, fermentation, lignin depolymerization through the catalytic transfer hydrogenolysis (CTH) process, and the separation and purification of the final products (Alherech et al., 2021; Ponnusamy et al., 2019). Pretreatment is the most essential unit process in the biorefinery (Joshi et al., 2015). It aims to fractionate the lignin aromatics from the polysaccharide-based fibers formed by cellulose and various types of hemicelluloses (Terrett et al., 2019). The role of hemicellulose hydrolysis in lignin release is particularly important yet often overlooked. Effective hydrolysis of hemicelluloses significantly enhances lignin dissociation from the biomass matrix, thus improving lignin extractability and subsequent valorization potential. Advances in analytical methodologies, particularly SS-NMR spectroscopy, have enabled detailed structural analysis of hemicellulose fractions, highlighting their critical influence on lignin separation efficiency. Due to the complex physicochemical properties and varied digestibility of the plant cell wall components, the development of the conventional “sugar-platform” biorefinery has paid significant attention on the utilization of only the cellulose and neglected the values of hemicelluloses and lignin. This strategy has significantly affected the economic potential of the industry (Khan et al., 2023). Lignin residues or the precipitate

redeposited on the pretreated substrates can interfere with the enzymatic hydrolysis process and limit the productivity of fermentable carbohydrates (Chan et al., 2021). The damage to lignin integrity during pretreatment can seriously impact the value and yields of aromatic products in the CTH process (Dong et al., 2019a). As interest grows in the production of high-value lignin monomers, various biorefinery strategies have been developed to enhance lignin productivity and yields (Khan et al., 2022). Organosolv pretreatment has demonstrated notable advantages in lignin valorization. Outstanding performances in biomass fractionation have been reported by using various solvents, such as alcohols, (Pan et al., 2005) diols, (Dong et al., 2019a) tetrahydrofuran (THF), (Cai et al., 2013) and γ -Valerolactone (GVL) (Kim et al., 2018). In those processes, diluted acid is applied to cleave the relatively weaker ether linkages between hemicelluloses and lignin, and organosolv serves as a carrier of the fractionated lignin fragments. However, serious side effects have also been reported in these processes. At high pretreatment severity, acid catalysts can attack the β -O-4 linkage in lignin, an essential structure to produce valuable lignin monomers.

Advanced process control aiming to balance the performance of biomass fractionation and lignin condensation could play a critical role in the biorefinery process. In a conventional biorefinery system, pretreatment severity has been optimized through biomass-specific experiments at various operational conditions, such as the types and doses of organosolv and acid catalysts, biomass-to-solvent ratio, treatment time, and temperature (Islam et al., 2020a). Owing to the inherent recalcitrance and wide diversity of biomass, the construction of a universally applicable process control algorithm for a biorefinery is challenging, especially when the concerned biomass is derived from wastes. The existing process control models, such as the H-factor, (Ramos, 2003) combined severity factor (CSF) (Schell et al., 2003), and combined hydrolysis factor (CHF) (Zhu et al., 2012), have been developed by the pulping industry with preprocessed wood chips.

These models primarily focus on predicting the hydrolysis capacity of pretreated substrate for sugar production, while neglecting the integrity of the fractionated lignin for monolignol production. For example, Zhu et al. constructed the CHF to simulate

the hydrolysis rate of the rapidly and slowly decomposable hemicellulose but did not quantify the integrity of dissolved lignosulfonate (Zhu et al., 2012). Phongpreecha et al. reported the positive relationship between lignin integrity (β -O-4 linkages) and monomer yields, but did not consider the impacts of changing pretreatment severities on lignin structures and hemicellulose removal (Phongpreecha et al., 2017). Meanwhile, the essential information on structural changes of pretreated biomass has been rarely reported, likely due to the limitations and challenges of current analytical tools. A numerical method for accurate simulation and control of the lignin-hemicellulose dual dissociation system has not been established, which has seriously hindered the development of any biorefinery system.

In this study, we investigate the fate of the three key building-block chemicals of lignocellulosic biomass and establish a kinetic model strategy to control the whole-biomass biorefinery processes (**Fig.5-1**). COS was selected for monolignols production due to its significant industrial potential and environmental impacts in Asia (Chaydarreh et al., 2021). Its particularly high lignin and hemicellulose contents also provide us unique opportunities to quantify the hemicellulose hydrolysis kinetics and delignification mechanism over a wider range of pretreatment severities. Solid-state NMR technique was applied to characterize and quantify various hemicellulose contents, including the two-fold and three-fold conformations of xylan (Xn^{2f} and Xn^{3f}) in the solid samples. The model was used to predict the change in lignin yield and integrity by the Xn^{3f} signals, and to upgrade the biorefinery process. The findings of this study provide both mechanistic and numerical insights to utilize whole lignocellulose in a biorefinery.

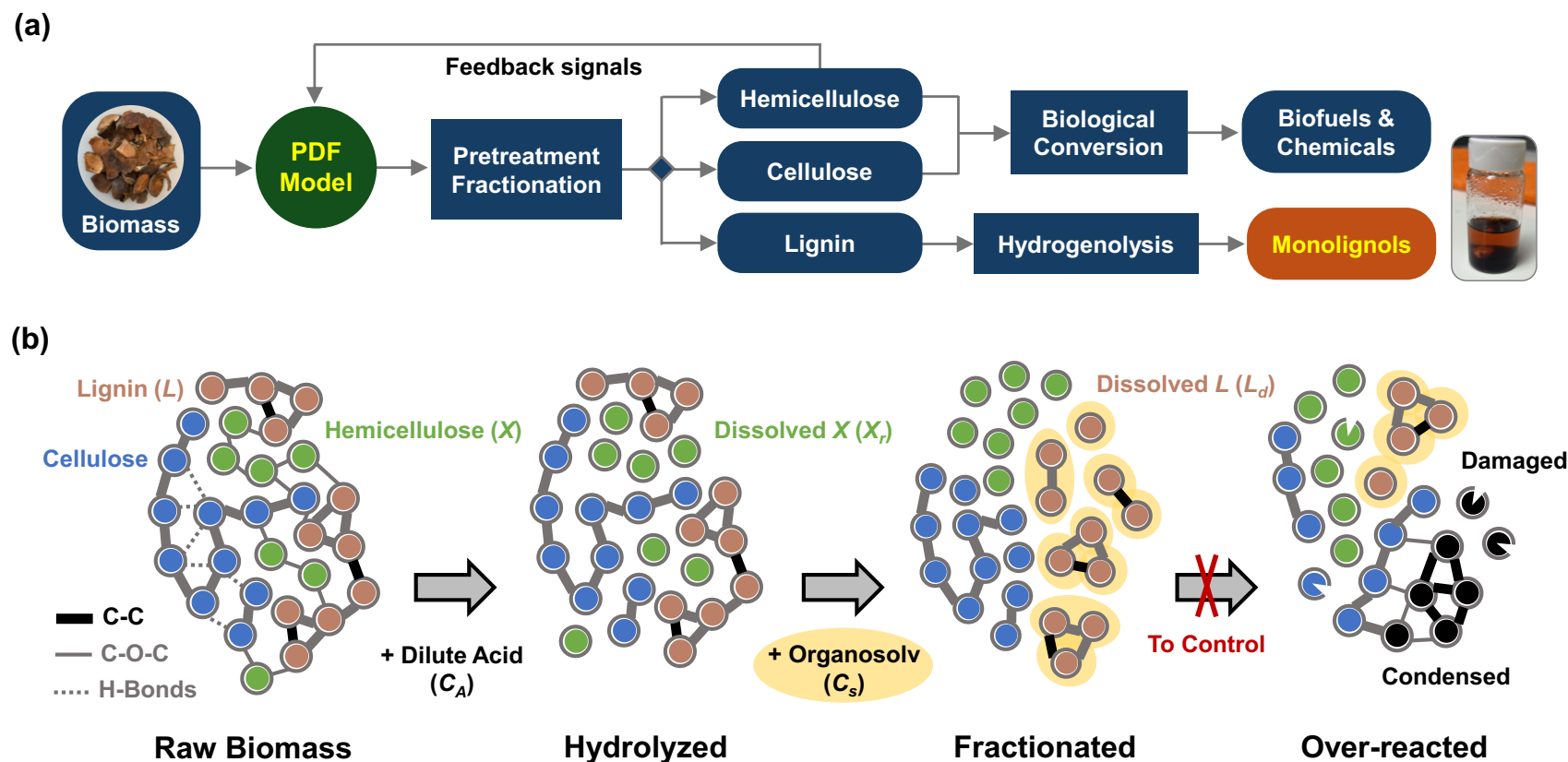


Fig.5-1. Conceptual diagrams of (a) a lignin-first biorefinery based on the newly proposed control algorithm, the pretreatment and depolymerization factor (PDF), for maximizing the monolignol production potential from *Camellia oleifera* shell (COS); (b) structure of the PDF model, with three key building block chemicals, interunit linkages, pretreatment conditions, and reaction mechanisms. Note: cellulose (blue) is formed by glucan connected with glycosidic bonds (solid gray lines) and hydrogen bonds (dashed gray lines). It links with hemicellulose and other related carbohydrates (X, green) and then lignin (L, brown) via ether bonds. The pretreatment process mainly involves two reactions, i.e. carbohydrate dissociation via dilute acid (CA) and lignin dissolution via organosolv (CS). The two reactions have been simulated separately for simplicity although they occur simultaneously in the system. Excessive pretreatment induces lignin C-C condensation and structural degradation (black)

5.2. Experiment Design

5.2.1. Biomass feedstock and chemicals

The *Camellia oleifera shell* (COS) samples were sourced from the tea oil camellia plantation operated by Guangdong Shanma Agricultural and Forestry Development Co., Ltd. (Lianzhou, China). The raw COS biomass consists of approximately 12% cellulose, 31% hemicellulose, and 34% Klason lignin, providing a suitable substrate for investigating the hydrolysis kinetics of hemicelluloses, the structural changes of fractionated lignin, and the yields of monolignols. All reagents and chemicals used in this study were of analytical grade, obtained from Sigma-Aldrich (MO, USA) and J&K Acros Organics (Beijing, China).



Fig.5-2.Photo of the feedstock in this study

5.2.2. Biomass Pretreatment

The COS was dried in a 50°C oven for 72 hours. It was then milled and sieved to obtain dry powder with a maximum particle size of 150 μm . The biomass powder was treated using the Soxhlet setup (ethanol/toluene 1:2, v/v) for 24 h to remove the polar and non-polar extractives. The COS powder was used for further treatment or lignin depolymerization experiments. The extractive-free COS powder was subjected to pretreatment using 1,4-butanediol (1,4-BDO) and different acidic catalysts, i.e., hydrochloric acid (HCl), ferric chloride (FeCl_3), and p-toluene sulfonic acid (p-TsOH). In each experiment, 3-20 g of extractive-free sample was introduced into the biorefinery reactor, along with a solvent mixture of organosolv and water at a ratio of 4:1 (v/v). The

solid-to-solvent ratio was 1:10 and the acid dose was 50-400 mM. The reactor was sealed and heated to the desired temperature (80 and 121 °C) in an autoclave, at a heating rate of 15 °C min⁻¹. Once the target temperature was reached, the reaction was maintained at the temperature for 0.15-15 h. After pretreatment, the reactor was opened when the internal pressure was reduced to 1 atm, and the system was cooled to room temperature in approximately one hour. For clarity and reproducibility, the detailed experimental conditions are summarized in **Tables S5-1** and **S5-2**

Following pretreatment, the biomass was separated into solid and liquid fractions. The solid fraction was thoroughly washed with the same solvent used in pretreatment until the washing solution became colorless, followed by 1-3 rinses with deionized water to remove residual solvent. The cleaned solid was then dried in an oven at 105 °C for at least 24 hours prior to chemical composition analysis. The liquid fraction was combined with 3-10 volumes of antisolvent (deionized water) to induce lignin precipitation. The resulting suspension was centrifuged at 5000 rpm for 1-3 minutes to collect the solid lignin. The supernatant was discarded, and the lignin pellet was washed 1-3 times with deionized water to eliminate residual organosolvent and acid. The purified lignin was subsequently frozen at -20 °C for 24 hours and dried in a freeze dryer for 72 hours to yield dry lignin powder.

5.2.3. Preparation of subtract sample for SS-NMR

The COS samples, subjected to various pretreatment conditions, were separated, dried, and then ground into a fine powder using a mortar and pestle for subsequent SS-NMR analysis. To obtain sufficient ¹³C signal intensity, at least 60 mg of each dried sample was placed into 3.2 mm rotors. To ensure proper packing and minimize static charge, approximately 20% (w/w) of water was incrementally added and thoroughly mixed with the dry COS powder, bringing the total sample weight to around 100 mg. The rehydrated samples were then evenly distributed within the rotors to ensure homogeneity and reduce spectral inconsistencies. The packed rotors were loaded into either a 400 MHz or 600 MHz Bruker solid-state NMR spectrometer. Cross-polarization (CP) and multi-CP techniques were employed to acquire MAS-SS-NMR spectra,

enabling detailed characterization of structural changes in the substrates under different pretreatment conditions.

5.2.4. Preparation of lignin sample

Native COS lignin samples were prepared to gather the structural information of the original lignin in the biomass before pretreatment. Extractive-free COS was ball-milled at 450 rpm, with a 10-minute rest period between each of the 10-min milling cycles, for a total of 8 to 12 hours. The milled powder was solvent-extracted using 96:4 dioxane-water (v/v), at a solvent-to-powder ratio of 10:1 (v/w). The mixture was continuously stirred for 24 hours to ensure thorough extraction of lignin. For the fractionated lignin, the pretreatment spent liquor was filtrated by Advantech 5A filter paper. The reactor and substrate were washed with a solvent mixture three times (10 mL each), and then the solution was mixed with the filtrate of the pretreatment liquor. The solid residue was rinsed with distilled water until the solution was colorless, then oven-dried at 105°C before chemical composition analysis. The lignin-rich solution (10 mL) volume was transferred to 50-mL centrifuge tubes, which were then filled with distilled water. The mixture was stirred and left for 30 minutes for lignin precipitation. The resulting suspension was centrifuged at 5000 rpm for 1-3 minutes to collect the solid lignin. The supernatant was discarded, and the lignin pellet was washed 1-3 times with deionized water to eliminate residual organosolv and acid. The purified lignin was subsequently frozen at -20 °C for 24 hours and dried in a freeze dryer for 72 hours to yield dry lignin powder.

5.3. Results and discussion

5.3.1. Overview of pretreatment results

A total of 66 biorefinery experiments were conducted on the COS biomass to investigate the hydrolysis kinetics of hemicelluloses, structural changes of fractionated lignin, and monolignol yields. The raw COS consists of ~12% cellulose, ~31% hemicellulose, and ~34% Klason lignin. The changes in chemical compositions of the

pretreated COS over different pretreatment conditions are presented in **Fig.5-3**. The experimental matrix covered three acids (i.e., HCl, p-TSOH, FeCl₃), acid concentrations (0-400 mM), temperatures (80-121°C), and time (0-24 h). As a result, all the tested acids can effectively hydrolyzed the hemicellulose and fractionated the lignin with no significant difference observed at the same acid doses. Increasing acid concentrations from 50 to 400 mM reduced the lignin and hemicellulose contents significantly, mostly due to the rapid breakage of the ether linkages (C-O) between lignin and hemicellulose or among hemicelluloses (Islam et al., 2020a; Kang et al., 2019; Woiciechowski et al., 2020).

The kinetic profiles of residual lignin (L_r) and hemicellulose (H_r) at two pretreatment temperatures, i.e., 80°C and 121°C, are shown in the upper right corner of **Fig.5-3**. The pretreatment performance is highly temperature sensitive. Rapid decreases in lignin and hemicellulose contents were observed for all three acids tested. Within the first 3 hours of treatment, the residual lignin content in biomass (over the original content, L_r) reduced from 38.2% to 6.4% and hemicellulose from 31% to 5.2% at 121°C. At 80°C, lignin removal was significantly slower, with a reduction of only 36.8% after 3 hours, leaving 24.2% for L_r . Similarly, H_r decreased from 31% to 22% when the pretreatment temperature decreased from 121°C to 80°C. The rapid separation at higher temperatures increased the reactivity of acids in hydrolyzing lignin-hemicellulose complexes. Notably, considerable differences (~10%) in L_r were observed between 121°C and 80°C under acid-free conditions, which corresponded to the hydrolysis of galactose (Gal). Finally, complete removal of lignin and hemicellulose was difficult even at extended pretreatment time, i.e., L_r was 4.8% (~87.5% removal) and H_r was 3.4% (~89.6% removal) after 12 hours of treatment. COS appears to be significantly more recalcitrant than other types of biomasses, likely due to its significantly higher lignin and hemicellulose contents.

The consistent trends in lignin and hemicellulose removals confirmed the importance of hemicellulose hydrolysis for effective dissociation of lignin fragments from the biomass matrix, as reported by Leu and Zhu (Leu & Zhu, 2013a).

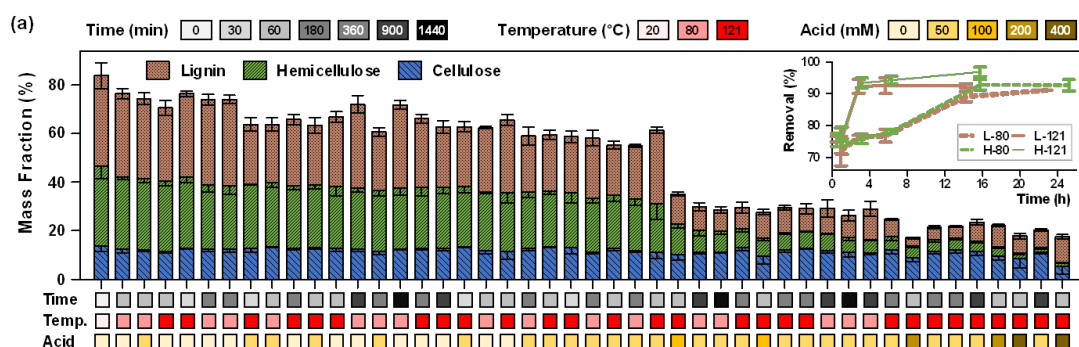


Fig.5-3. Summary of biorefinery experiments. Changes in chemical compositions of biomass at increasing pretreatment severities. (Subfigure upper right) The dynamic changes of lignin and hemicellulose removal at 80 and 121°C.

5.3.2. Characterization of hemicelluloses using solid-state NMR

The solid-state NMR technique associated with the Multi-CP method can serve as an effective tool for direct ^{13}C analysis for both aromatics and non-crystalline structures in the biomass. All the tested samples, including raw COS and pretreated substrates with the highest pretreatment severity (PDF = 15.2) to the lowest (PDF = 0.87) severities, have exhibited well-resolved ^{13}C signals despite not being processed with ^{13}C labels. The ^{13}C signals corresponding to hemicelluloses and lignin aromatics show a high signal-to-noise ratio, allowing us to distinguish (a) the aromatic carbon signals without proton correlation (e.g., C1, C3, C4 in G-lignin unit, and/or C1, C3, C4, C5 in S-lignin unit); and (b) the hemicelluloses with different spatial structures, particularly the Xn^{2f} and Xn^{3f} . Specifically, the aromatic carbons in quantitative ^{13}C spectra are typically located at 110-160 ppm (S3/5 153 ppm G3 149 ppm, G4 145 ppm S1/4 135 ppm), and the hemicelluloses carbon 1 (C1) peaks appear at the range of 95-110 ppm. Normalized by the interior cellulose C4 (i4) peak at 88 ppm, the peak intensities corresponding to lignin and hemicellulose decrease significantly with the progression of increasing PDF. Nevertheless, the peaks of hemicellulose C2-5 still overlap severely with cellulose C2-5 peaks at 60-80 ppm.

To distinguish the peaks, the spectra were deconvoluted based on the chemical shifts determined from the isolated molecules and the integration of the corresponding intensities of carbon atoms in each hemicellulose. Kirui et al. applied solid-state NMR

to quantify intermolecular carbohydrate-aromatic associations, enabling different xylan conformations, a common plant component (Kirui et al., 2022). The carbon chemical shifts of hemicellulose and lignin were obtained by subtracting the two parent spectra after normalization (Table S2).

The chemical shifts of corresponding carbon sugar signals in the quantitative ^{13}C spectra closely match those reported in previous studies, (Kirui et al., 2022; Zhao et al., 2024) indicating minimal differences in hemicelluloses structures between COS and those of other biomass species (e.g., eucalyptus and poplar). In addition, increasing PDF did not result in significant structural changes to the hemicelluloses, confirming the applicability of solid-state NMR and PDF model in different biomass.

The original and deconvoluted ^{13}C peaks are shown in different colors in **Fig.5-4 b-e**. The peak deconvolution methods were applied to quantify the specific hemicellulose contents, including Ara C1 (107.4 ppm), Xn^{2f} C1 (103.8 ppm), Xn^{3f} C1 (101.1 ppm), Gal C1 (100.8 ppm), and Xn^{2f} C4 (82 ppm) under solid-state NMR (**Fig.5-4b**), (Ghassemi et al., 2021) and the less-overlapping peaks were selected for analysis to improve the accuracy of quantification. The simulated spectra showed excellent agreement with the measured spectra (dashed line in **Fig.5-4c**), suggesting the precision of the deconvolution method (**Fig.S5-1**). The chemical shifts of corresponding carbon sugar signals in the quantitative ^{13}C spectra closely match those reported in previous studies, (Kirui et al., 2022; Zhao et al., 2024) indicating minimal differences in hemicelluloses structures between COS and those of other biomass species (e.g., eucalyptus and poplar). In addition, increasing PDF did not result in significant structural changes to the hemicelluloses, confirming the applicability of solid-state NMR and PDF model in different biomass.

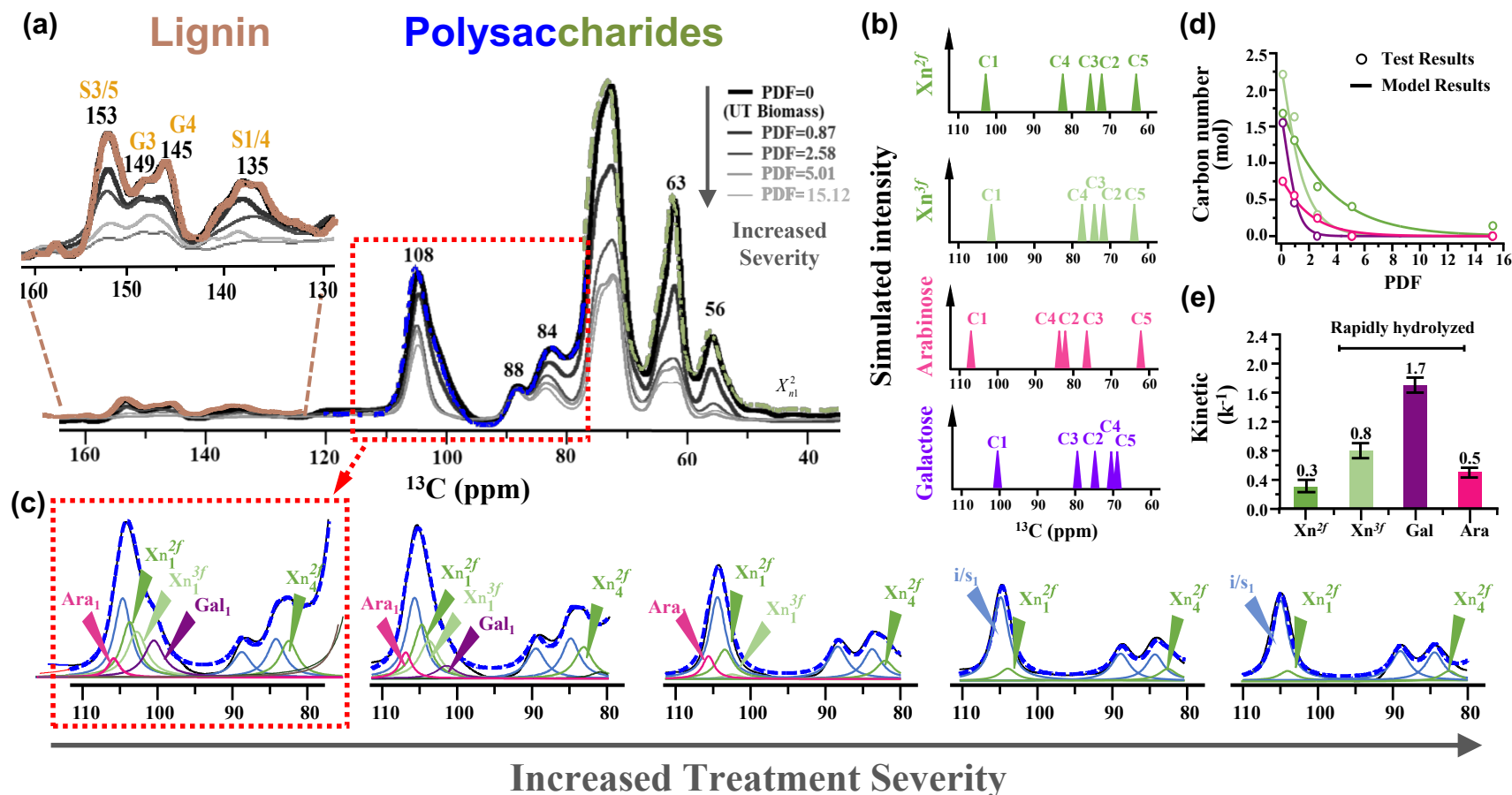


Fig.5-4. Solid-State NMR spectra of biomass with detailed peak assignments and quantitative analysis. (a) 1D ^{13}C Multi-CP spectra of unlabeled COS biomass under increasing pretreatment severity (PDF = 0, 0.87, 2.58, 5.01, and 15.12); (b) peak positions of the main carbohydrates obtained by difference spectra. Notably, the difference of spectra are obtained by subtracting two parent spectra normalized using the standard peak of cellulose at 88ppm; (c) spectral deconvolution of the major carbohydrates carbon peak region (75-115 ppm): Ara C1 (107.4 ppm), Xn ^{2f} C1 (103.8 ppm), Xn ^{3f} C1 (101.1 ppm), Gal C1 (100.8 ppm), and Xn ^{2f} C4 (82 ppm); (details of the spectral deconvolution in Fig.S5-3); (d) simulation results compared with the experimental results; and (e) 1st-order hydrolysis kinetics of concerned carbohydrates.

The original and deconvoluted ^{13}C peaks are shown in different colors in **Fig.5-4 b-e**. The peak deconvolution methods were applied to quantify the specific hemicellulose contents, including Ara C1 (107.4 ppm), Xn^{2f} C1 (103.8 ppm), Xn^{3f} C1 (101.1 ppm), Gal C1 (100.8 ppm), and Xn^{2f} C4 (82 ppm) under solid-state NMR (**Fig.5-4b**),(Ghassemi et al., 2021) and the less-overlapping peaks were selected for analysis to improve the accuracy of quantification. The simulated spectra showed excellent agreement with the measured spectra (dashed line in **Fig.5-4c**), suggesting the precision of the deconvolution method (**Fig.5-5**).

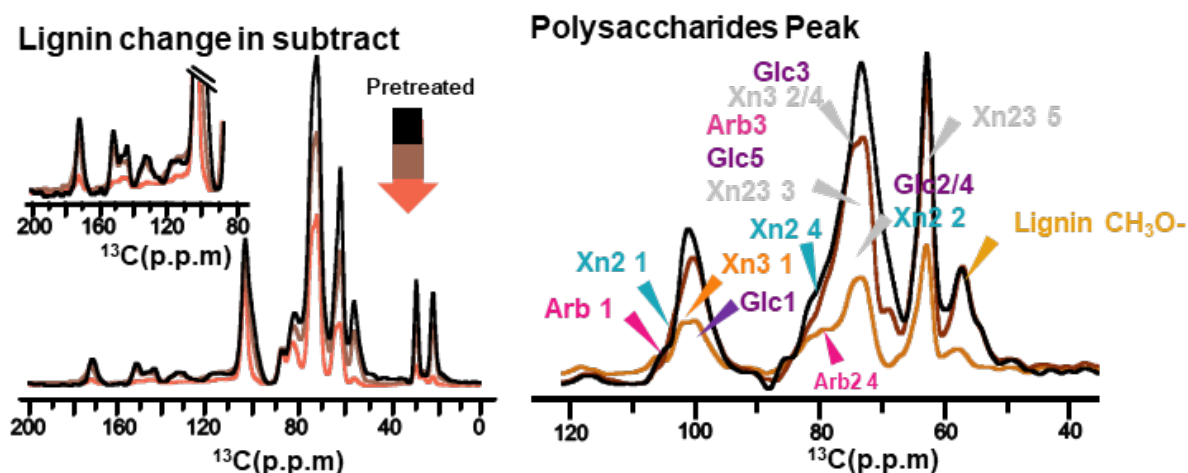


Fig.5-5. Differential spectra with peak deconvolution highlighting characteristic signals of xylan (Xn^{2f} , Xn^{3f}), glucan (Glc), arabinose (Arb), and lignin methoxyl groups, enabling quantitative analysis of hemicellulose conformations.

5.3.3. X-ray diffraction (XRD) for the COS substracts

The XRD patterns of COS samples under different pretreatment conditions were shown in **Fig 5-6**, alongside pinewood as a comparative reference. The untreated COS (pattern a) exhibits a low crystallinity index (CrI) of 23%, significantly lower than that of untreated pine (e, CrI = 54%). This notable difference can be attributed to the inherently higher hemicellulose content and shorter cellulose fibrils in COS, which result in a more pronounced amorphous region and thus reduced overall crystallinity. Upon pretreatment with 1,4-butanediol (1,4-BDO) and 100 mM H_2SO_4 at 121 °C for 1 hour (b), the CrI increased slightly to 26%, indicating the partial removal of amorphous components such as hemicelluloses.

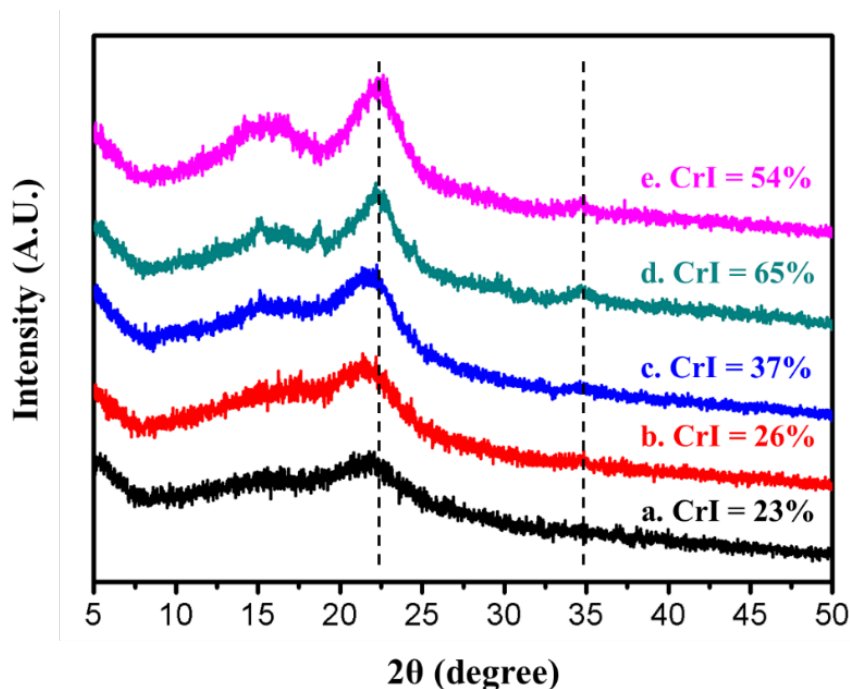


Fig.5-6. XRD of a) untreated; b) 1,4-BDO/100mM H_2SO_4 , 121°C; c) 1,4-BDO/400mM H_2SO_4 , 121°C; d) ethanol/100mM H_2SO_4 , 165°C *Camellia Oleifera* shell; e) untreated pine.

A more substantial increase in CrI was observed in the sample treated under the same conditions with a higher acid concentration (400 mM H_2SO_4), where the CrI reached 37% (c), suggesting more effective removal of amorphous polysaccharides and partial reorganization of the cellulose crystalline regions. In the case of ethanol-based pretreatment at a higher temperature (165 °C, 100 mM H_2SO_4), the CrI increased markedly to 65% (d), indicating the extensive removal of hemicelluloses and disruption of non-crystalline cellulose domains. This sharp increase suggests that severe pretreatment conditions promote the enrichment and reordering of crystalline cellulose.

5.3.4. Morphology of Pretreatment Feedstock (SEM)

The SEM image of untreated *Camellia oleifera* shell (COS) was shown in **Fig5-7**, revealing a loosely packed and poorly organized cell wall structure. The magnified region indicates that the surface is relatively smooth, with minimal textural complexity, suggesting that most of the amorphous components, such as hemicelluloses and extractives, remain intact. This intact structure forms a physical barrier, limiting the

accessibility of chemical reagents to the interior of the biomass during subsequent processing steps.

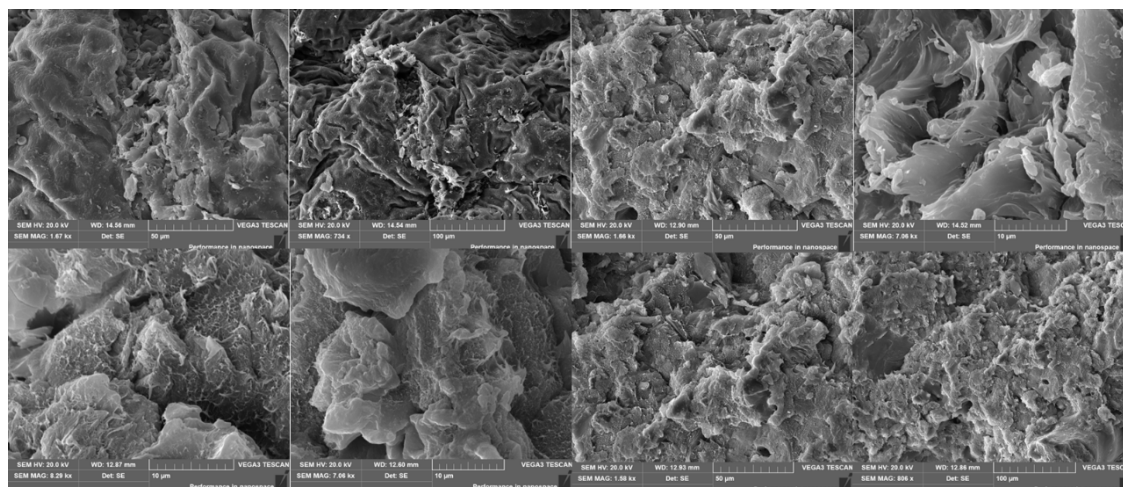


Fig.5-7. SEM of untreated/treated *Camellia Oleifera* shell

The morphology of COS after acid-organosolv pretreatment illustrated in Fig5-9. The originally compact architecture is disrupted, and the surface appears fragmented into irregular, curled clumps with significant roughness. This structural transformation indicates extensive disintegration of the amorphous regions and partial removal of hemicelluloses and lignin, which are more readily solubilized under acidic pretreatment conditions. The resulting fibrillated texture and increased porosity contribute to a significantly enlarged surface area, exposing more cellulose microfibrils. Such an increase in surface roughness and accessibility not only facilitates further enzymatic or chemical conversion of cellulose but also provides more interaction sites when incorporating biomass into polymer matrices. These morphological changes may enhance interfacial adhesion in composite materials, thereby improving mechanical properties and compatibility in bio-composite fabrication.

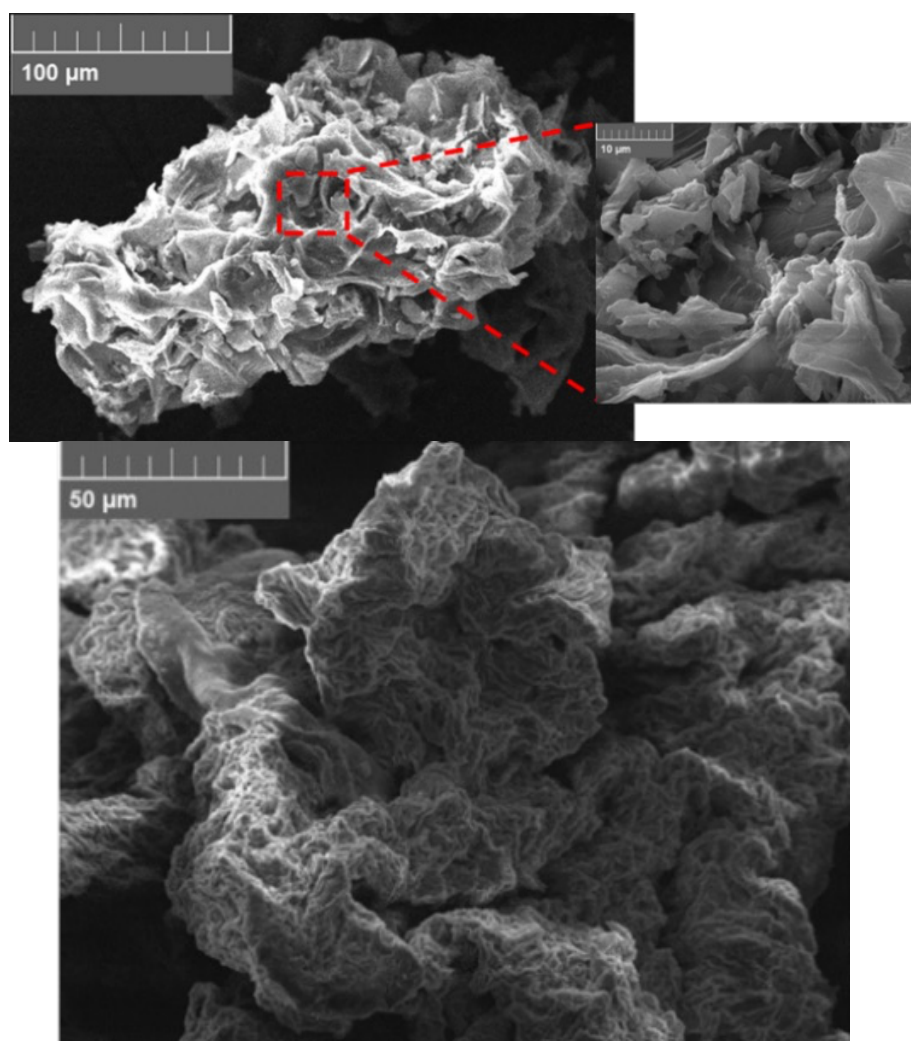


Fig.5-8. SEM of untreated/treated *Camellia Oleifera* shell

5.3.5. Polysaccharide hydrolysis kinetic

The quantified deconvolution confirmed the results of chemical composition analysis conducted by the NREL method (**Table 5-1**). It indicated the outstanding applicability of the solid-state NMR for determining both the spatial and compositional information of plant cell wall structures, without the complex and destructive two-step acid digestion. In this study, solid-state NMR was employed to distinguish and quantify various Polysaccharide components, such as xylan with two-fold and three-fold screw conformations (Xn^{2f} and Xn^{3f}), arabinose (Ara), and galactose (Gal). The spectral deconvolution allowed for precise calculation of the relative mass fractions of these components in the pretreated samples, forming the basis for kinetic analysis. These NMR-derived concentrations were subsequently integrated into a hydrolysis kinetic

framework based on the Pretreatment Depolymerization Factor (PDF), which captures the severity of organosolv-acid pretreatment by aggregating key operational parameters (e.g., time, temperature, acid concentration, and solvent ratio) into a unified quantitative descriptor. To determine the hydrolysis kinetics of each hemicellulose, the rate coefficients (k) were obtained through nonlinear curve fitting. Specifically, the time-resolved experimental data were fitted against simulation outputs derived from the PDF-coupled kinetic model. By adjusting the individual k values iteratively, the simulation results were optimized to align with the experimentally observed degradation extents under varying pretreatment severities. This fitting process enabled the extraction of dynamic hydrolysis rates that reflect the reactivity of each hemicellulose component.

To evaluate the general applicability of the PDF model beyond *Camellia oleifera* shell, we extended the validation to representative softwood (pine), hardwood (eucalyptus), and grass (bamboo) under identical pretreatment conditions. As shown in **Fig.5-9**, both hemicellulose and lignin residues from these diverse taxa closely followed the trajectories predicted by the model. In addition, literature data on eucalyptus (Dong C. et al.) and *Acacia confusa* (Islam M.K. et al.) were incorporated for comparison, and the model reproduced these independent results with high fidelity. A summary of the pretreatment conditions and corresponding PDF values is provided in **Table S5-2**. Together, these results demonstrate that the PDF framework is not restricted to a single feedstock type but can be reliably transferred across woody and herbaceous biomasses, capturing the universal kinetics of hemicellulose hydrolysis and lignin transformation.

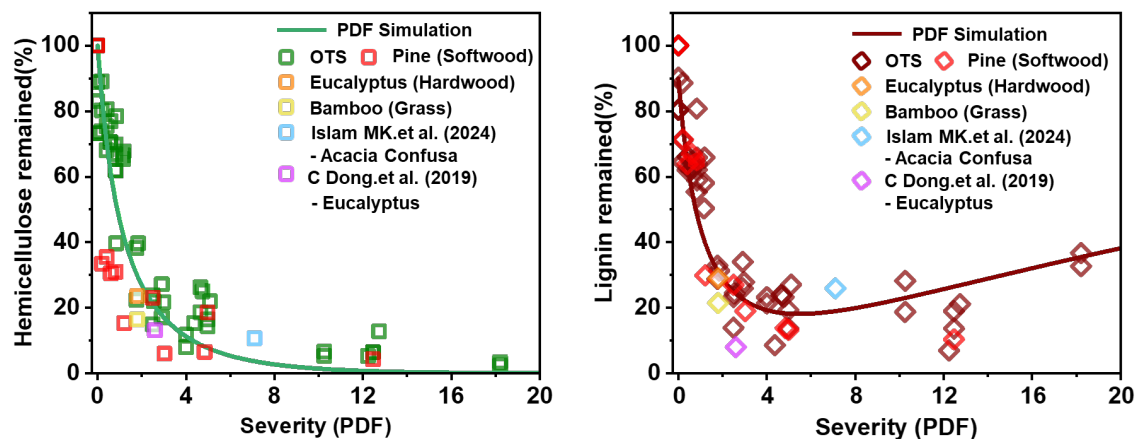


Fig.5-9.Cross-feedstock validation of the Pretreatment Depolymerization Factor (PDF) model. Experimental residues of hemicellulose (left) and lignin (right) from softwood (pine), hardwood (eucalyptus), and bamboo (grass) are well simulated by the PDF model. Literature results for eucalyptus Dong C. et al and Acacia confusa Islam M.K. et al. further support the robustness of the framework. The consistent agreement across different plant types demonstrates that the PDF model can be reliably applied beyond *Camellia oleifera* shell

The results revealed distinct differences in hydrolysis behavior among hemicelluloses. Xn^{3f} and galactose exhibited relatively rapid hydrolysis, with fitted rate constants of 0.81 and 1.71, respectively. In contrast, Xn^{2f} and arabinose showed slower degradation kinetics, with rate constants of 0.32 and 0.49, respectively.

Table 5-1: Comparison of mass fractions measured by NREL and solid-state NMR methods, along with dynamic hydrolysis rates of various hemicelluloses under slow (S) and rapid (R) hydrolysis conditions.

	Xylan ^{2f}	Xylan ^{3f}	Arabinose	Mannose	Galactose
Hydrolysis type*	S	R	S	-	R
Mass fraction θ_i^H (NREL)	0.295	0.295	0.0437	0	0.0579
Mass fraction θ_i^H (solid-state NMR)	0.106	0.136	0.0493	0	0.0833
Hydrolysis rate f_i^H	0.32	0.81	0.49	-	1.71

* θ_i^H Mass fraction of the different polysaccharides i ; S: slowly fraction;
R: rapid fraction

Conclusion

This chapter investigates the structural dissociation and depolymerization dynamics of hemicellulose-lignin complexes in *Camellia oleifera* shell (COS) under organosolv-acid pretreatment. Due to its high contents of hemicelluloses and lignin, COS presents both opportunities and challenges for whole-biomass valorization. The removal of amorphous hemicelluloses, particularly Xylan with Xn^{3f} , was found to be a key factor enabling effective lignin fractionation and crystalline cellulose enrichment. Through quantitative solid-state NMR analysis and XRD characterization, we demonstrated that increased pretreatment severity enhances cellulose crystalline while enabling the selective dissociation of lignin-bound hemicelluloses. To mechanistically link lignin integrity with pretreatment conditions, we developed the Pretreatment Depolymerization Factor (PDF), a kinetic model capable of predicting both hemicellulose hydrolysis and lignin depolymerization. The PDF model successfully captured the differential reactivity of key hemicellulose components and aligned with the experimental yields of lignin and monolignols. This dual-parameter framework offers a powerful tool for optimizing pretreatment strategies in lignin-first biorefineries, balancing carbohydrate recovery and lignin valorization.

CHAPTER 6. LIGNIN CONDENSATION AND DEPOLYMERIZATION

ABSTRACT

Harvesting sustainable aromatic chemicals from lignocellulosic biomass is a critical subject of the bio-based economy. Ideal whole-biomass utilization strategies have been developed for lignin valorization via solvent fractionation followed by catalytic hydrogenolysis. However, the limited understanding on lignin and hemicellulose structure and lack of control strategy on side reactions has seriously hindered the applicability of those novel biorefinery techniques. To overcome this challenge, this study integrated various solution- and solid-state Nuclear Magnetic Resonance (NMR) analyses with 66 biorefinery and hydrogenolysis experiments, aiming to construct a novel Pretreatment Depolymerization Factor (PDF) model for simulating the hydrolysis kinetics of hemicelluloses, including arabinose and xylans in different conformations. The new PDF model successfully simulated the delignification dynamics, lignin integrities, and monolignols yield with 90.9% accuracy, leading to a new discovery of an unexplored C β -C γ lignin bond cleavage initiated by C γ -OH activation. The monolignols yield was increased by 300% through a PDF-modified fractionation strategy without additional solvent, energy, or chemicals. This effective preservation of both carbohydrates and aromatic products highlights the economic and environmental benefits of smart fractionation control in complete lignocellulose utilization.

6.1. Introduction

Global industrialization, predominantly driven by fossil fuels, has significantly contributed to sustained economic growth and infrastructure development (Syvitski et al., 2020). Nonetheless, dependence on petroleum-derived resources continues to raise serious sustainability concerns, particularly regarding greenhouse gas emissions and environmental degradation (Achakulwisut et al., 2023; Welsby et al., 2021). As an alternative, lignocellulosic biomass offers immense potential owing to its abundant availability, rapid regeneration capability, and carbon-neutral lifecycle (Singhvi & Gokhale, 2019).

Efficient utilization of lignocellulosic biomass is essential for maximizing resource value and minimizing waste throughout the transition toward carbon neutrality. However, the lignocellulosic biomass consists of highly recalcitrant bio-composites comprising cellulose, hemicelluloses, and lignin (Dong et al., 2019b). These components are intricately interconnected at the nanoscale, forming complex anisotropic structures within plant secondary cell walls (Kang et al., 2019). The covalent linkages and extensive hydrogen-bonding interactions between lignin and polysaccharides lead to a dense and heterogeneous matrix featuring both hydrophobic and hydrophilic domains, which severely restricts biomass accessibility and impedes enzymatic efficiency in biorefinery processes (Kirui et al., 2022; Liu et al., 2016). The complexity structural of lignocellulosic biomass remains a key barrier to biomass valorization, underscoring the critical role of pretreatment in modern biorefineries. To overcome the biomass recalcitrance, current strategies typically involve a sequence of steps: pretreatment (Islam et al., 2020a), enzymatic hydrolysis (Chan et al., 2021), fermentation (Lin & Tanaka, 2006), lignin depolymerization via catalytic transfer hydrogenolysis (CTH) (Das et al., 2018), and final product recovery (Alherech et al., 2021). Among these, pretreatment is particularly essential, enabling lignin-carbohydrate dissociation and improving substrate accessibility for downstream conversion.

Organosolv pretreatment has demonstrated notable advantages in lignin valorization. Outstanding performances in biomass fractionation have been reported by

using various solvents, such as alcohols,(Pan et al., 2005) diols,(Dong et al., 2019a) tetrahydrofuran (THF),(Cai et al., 2013) and γ -Valerolactone (GVL) (Kim et al., 2018). In those processes, diluted acid is applied to cleave the relatively weaker ether linkages between hemicelluloses and lignin, and organosolv serves as a carrier of the fractionated lignin fragments. However, serious side effects have also been reported in these processes. At high pretreatment severity, acid catalysts can attack the β -O-4 linkage in lignin, an essential structure to produce valuable lignin monomers. This reaction would result in the generation of reactive benzylic carbocation intermediates at the C α position of lignin (Li et al., 2018a). The intermediates could couple with other benzoic rings and form a highly condensed polymer with stable C-C linkage, which severely affects the productivity of both fermentable sugars and monolignols (Islam et al., 2020b). Despite the numerous attempts to address the lignin condensation issues through newly discovered lignin dissociation chemistry, such as the application of formaldehyde to react with lignin side-chain hydroxyl groups,(Shuai et al., 2016b) or the direct induction of C-C formation for bisphenols via selective lignin arylation,(Li et al., 2024a) there have been limited studies to quantitatively evaluate the numerical relationship between delignification efficiencies and monolignol yield from a process control perspective.

In this study, we investigate the fate of the three key building-block chemicals of lignocellulosic biomass and establish a kinetic model strategy to control the lignin-first biorefinery processes. A *Camellia oleifera* shell (COS) biomass was selected for monolignols production due to its significant industrial potential and environmental impacts in Asia. Its particularly high lignin and hemicellulose contents also provide us unique opportunities to quantify the hemicellulose hydrolysis kinetics and delignification mechanism over a wider range of pretreatment severities. Solid-state NMR technique was applied to characterize and quantify various hemicellulose contents, including the two-fold and three-fold conformations of xylan (Xn^{3f} and Xn^{3f} in the solid samples). The changes in lignin structures were then determined using the 2D solution NMR, 31P NMR, and elemental analysis (EA). The results were associated with the sensitive removal of lignin-associated Xn^{3f} and the productivity of monolignols which formed into the basis of a new structural model, the “Pretreatment Depolymerization

Factor (PDF)”. The model was used to predict the change in lignin yield and integrity by the X_n^{3f} signals, and to upgrade the biorefinery process. The findings of this study provide both mechanistic and numerical insights to utilize whole lignocellulose in a biorefinery.

6.2. Experiment Design

6.2.1. Pretreatment lignin separation

To investigate lignin integrity and condensation-induced structural changes under varying pretreatment conditions, extractive-free *Camellia oleifera* shell (COS) powder was prepared by drying the biomass at 50 °C for 72 hours, followed by milling and sieving to obtain particles smaller than 150 μ m. Soxhlet extraction was then performed using an ethanol/toluene mixture (1:2, v/v) for 24 hours to remove polar and non-polar extractives. The extraction-free COS powder was subjected to pretreatment in a solvent system consisting of 1,4-butanediol (1,4-BDO) or ethanol mixed with water (4:1, v/v), under various acid-catalyzed conditions. In each experiment, 3-20 g of extractive-free biomass was loaded into a sealed reactor with the solvent system at a solid-to-liquid ratio of 1:10. Acid concentrations were adjusted between 50 and 400 mM. The reactor was heated to either 80 °C or 121 °C at a rate of 15 °C min⁻¹ and maintained at the target temperature for 0.15-15 hours. Following pretreatment, the liquid fraction was separated and combined with 3-10 volumes of deionized water to induce lignin precipitation. The resulting suspension was centrifuged at 5000 rpm for 1-3 minutes to isolate the lignin pellet, which was then washed thoroughly with deionized water to remove residual solvents and acids. The purified lignin was frozen at -20°C for 24 hours and subsequently freeze-dried for 72 hours to obtain dry lignin powder.

6.2.2. Preparation of lignin samples

The lignin-rich solvent obtained after pretreatment was transferred to 50 mL centrifuge tubes and diluted with distilled water to facilitate lignin precipitation. The mixture was gently stirred and left undisturbed for 30 minutes to allow sufficient

aggregation of lignin molecules. The precipitated lignin was then recovered by centrifugation and subsequently freeze-dried for structural characterization. To obtain native lignin reflecting the original molecular architecture within *Camellia oleifera* shell (COS), a ball milling-enzymatic mild extraction method was applied. Specifically, extractive-free COS powder was subjected to planetary ball milling at 450 rpm with alternating 10-minute milling and rest intervals, for a total duration of 8-12 hours. The resulting ultrafine powder was extracted using a 96:4 (v/v) dioxane-water solution at a solvent-to-powder ratio of 10:1 (v/w). The mixture was stirred continuously for 24 hours to ensure thorough solubilization of the native lignin fraction. The freeze-dried lignin samples were subsequently subjected to structural analyses, including two-dimensional heteronuclear single quantum coherence (2D-HSQC) NMR, Fourier-transform infrared spectroscopy (FTIR), and gel permeation chromatography (GPC) to evaluate inter-unit linkages, functional groups, and molecular weight distributions, respectively. In addition, the isolated lignin samples were used for catalytic transfer hydrogenolysis (CTH) experiments to assess their monomeric depolymerization potential, with the resulting aromatic monomers quantified by gas chromatography (GC).

6.2.3. Catalytic transfer hydrogenolysis (CTH) for lignin depolymerization

The CTH experiments were conducted in a 50-mL stainless steel reactor (Parr Instruments Co). In each experiment, 0.5 g of pre-extracted COS powder or 0.1 g of purified lignin solid was loaded into the Parr reactor, together with 0.1g of Ruthenium/Carbon (Ru/C) catalyst and 20 mL of THF solvent. The reactor was sealed, flushed with 10 bar N₂ three times to remove O₂, and pressurized with H₂ to 40 bar (25°C). The stirring and heating rates were 700 rpm and 10°C/min, respectively, and the reaction was carried out at 230 °C. After the reaction, the Parr reactor was cooled in air and depressurized at 35°C(Shuai et al., 2016a). The detailed reaction conditions are summarized in **Table S6-1** and **Table S6-2**.

6.2.4. PDF-Controlled Semi-Continuous Flow Experiments

Two different pretreatment methods were employed after the establishment of PDF, i.e., a single-batch and a sequencing batch pretreatment. In the single-batch process, 10 g of biomass was treated with 120mL of organosolv containing 75mM of acid catalyst at 121°C for 6 hours. In the sequencing batch process, 10g of biomass was treated with 30 mL of organosolv per sub-batch, with a 1-hour reaction time applied to each of the first three sub-batches. The solid residue was filtered before the next sub-batch of pretreatment experiment. In the final sub-batch, a 3-hour reaction time was applied to ensure that the total solvent volume used in the sequencing batch process was equivalent to that in the single-batch pretreatment.

6.3. Results and discussion

6.3.1. Lignin structural changes at increasing PDF

The fractionated lignin of the COS at different PDFs was characterized using 2D ^{13}C - ^1H HSQC NMR analysis, where the signals of different lignin inter-unit distributions are shown in **Fig.6-1**. Briefly, the radiometric variations of lignin inter-unit linkages include: the lignin β -ether units A (with characteristic β -O-4 aryl ether linkages) (Aa, $\delta\text{C}/\delta\text{H}$ 73.1/7.4 ppm, Ab-G, $\delta\text{C}/\delta\text{H}$ 87.1/6.9 ppm Ab-S, $\delta\text{C}/\text{ph}$. 88.2/6.7 ppm, Ag, $\delta\text{C}/\delta\text{H}$ 62.7/6.2 ppm), with small amounts of resinols B (β - β)(Ba, $\delta\text{C}/\delta\text{H}$ 88.7/7.8 ppm, Bg, $\delta\text{C}/\delta\text{H}$ 64.2/6.4 ppm) and phenyl coumarans C (β -5) (Cg, $\delta\text{C}/\delta\text{H}$ 72.1/6.8 ppm and 72.1/6.6 ppm). Based on the results of 2D-HSQC NMR, the dynamic changes in various inter-unit linkages were quantified in Table 4, and the relative distributions of different lignin inter-unit linkages are presented in the bar charts (bottom row) of **Fig.6-1**. The native COS lignin (PDF = 0) contains 50.7% β -O-4 ether linkages, along with 11.01% β -5 and 13.97% β - β linkages, which is consistent with the findings of other studies (Guerra et al., 2006; Huang et al., 2011). With the increase in PDF (from 0 to 4.8), the percentage of β -O-4 linkage decreased exponentially from 50.7% to 28.9%. However, the contents of β - β and β -5 linkages of these lignin samples remain relatively stable regardless of changes in PDF.

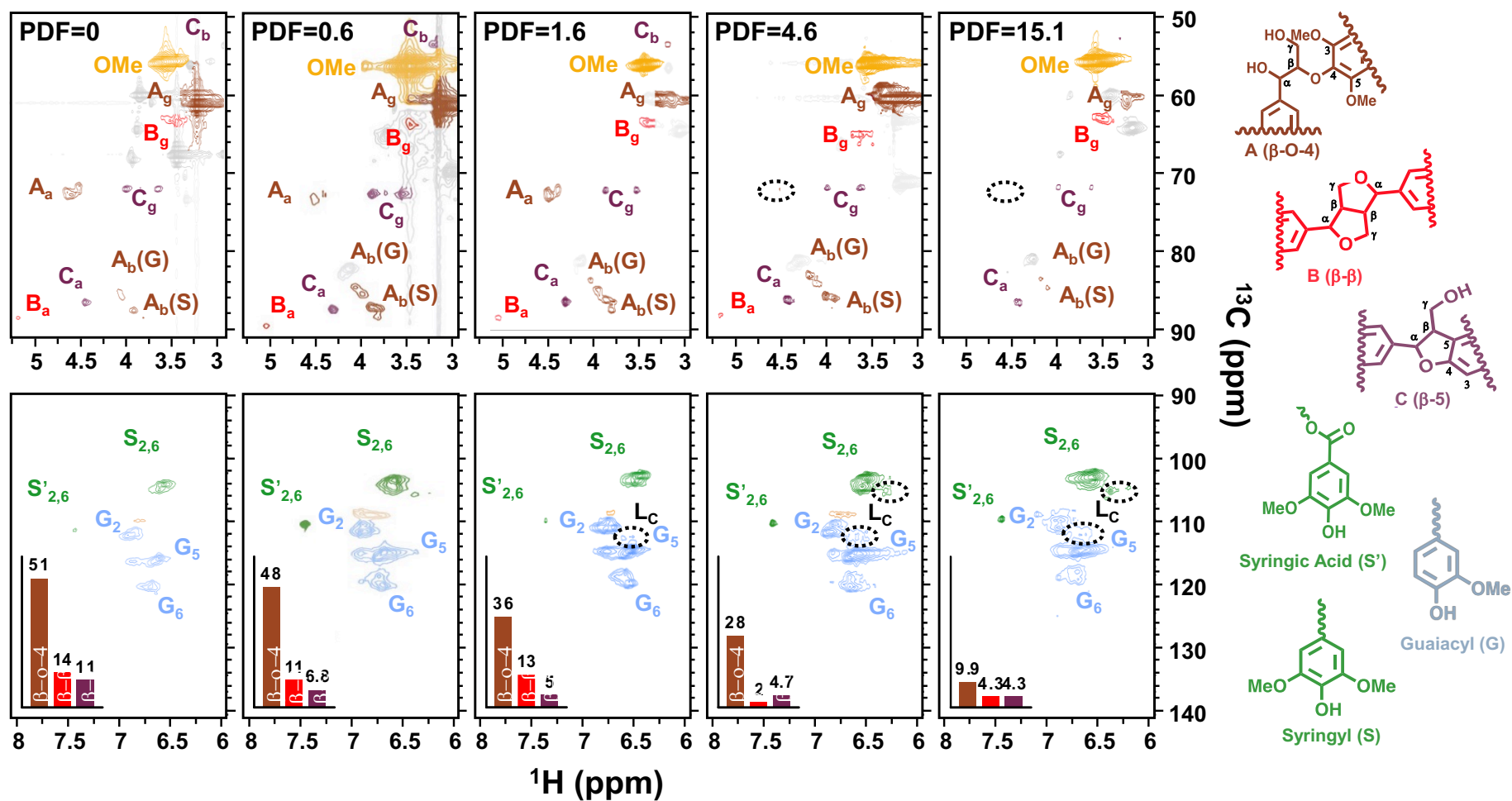


Fig.6-1. ^1H - ^{13}C 2D-HSQC NMR spectra showing lignin interunit linkages (in aliphatic regions, top row) and lignin fraction compositions (e.g., S/G ratios, in aromatic regions, bottom row) over increased pretreatment severities (PDF = 0, 0.46, 1.62, 4.61, and 15.12). Bar charts show the decreasing lignin inter-unit linkages (%) including β -O-4 (brown), β - β (red), and β -5 (purple) over gradually increased PDF.

The relationship between β -O-4 content and PDF was quantified. Although the delignification efficiency increased with increasing PDF, the quantity of β -O-4 linkages decreased with lignin condensation. The condensation products were newly formed stable C-C bonds as determined by NMR. The S/G ratios were quantified based on the 2D-HSQC NMR spectra by integrating the peak volumes of syringyl (S2/6, $\delta C/\delta H$ 104.0/6.68 ppm) and guaiacyl (G2, $\delta C/\delta H$ 111.0/6.96 ppm) units, normalized by the total lignin signals.

Table6-1:Quantitative analysis of lignin characteristics, including S/G ratios (syringyl/guaiacyl ratios) in the solid substrate, dissolved lignin content, and monomer after depolymerization at five gradually increased PDF (0, 1.3, 2.1, 5.4, and 15.1).

PDF	Substrate	Dissolved	Lignin
	lignin	lignin	monomer
S/G Ratio			
0	1.31	0.71	0.83
1.3	0.89	0.86	1.07
2.1	0.57	0.79	1.27
5.4	0.74	0.97	1.11
15.1	0.61	1.22	0.61

The S/G ratios of lignin subunits were lower than 1.0, when the PDF was as low as 0-4.5, suggesting that S-lignin units may be more difficult to remove than G-lignin. This finding confirmed by the results of solid-state NMR analysis, with which a large amount of S-lignin remained in the substrate at low PDF. G-lignin could serve as a linker that forms external sidechains and attached to the main lignin chain, making it more vulnerable to pretreatment. S-lignin may be buried inside the cellulose crystalline or the interlayer of the main lignin polymer chain and is more difficult to remove from the biomass. The vacant position at C5 of the G-lignin could increase the chance of condensation compared to that of S-lignin, in which the C5 has been either occupied by methoxy groups or linked with other lignin units. These factors led to the rapid formation of the free G-lignin fragments and resulted in a low S/G ratio of the lignin

product at low PDF. This finding provides valuable information for determining the optimal pretreatment time and the potential benefits of discharging the dissolved lignin fragments for mitigating the condensation reaction during the pretreatment process.

6.3.2. Lignin Molecular Weight Distribution by GPC

Gel permeation chromatography (GPC) was employed to determine the relative molecular weights of lignin fractions isolated under different Pretreatment Depolymerization Factor (PDF) conditions, enabling a detailed evaluation of lignin structural evolution during pretreatment. To elucidate the relationship between pretreatment severity and lignin depolymerization or condensation behavior, we first compared the molecular weight distributions of lignin and its depolymerized oil products obtained from two contrasting conditions PDF = 0.6 and PDF = 15.1. Under mild pretreatment (PDF = 0.6), the isolated lignin exhibited a relatively narrow and symmetric molecular weight distribution, with a number-average molecular weight (M_n) of 1836 g mol⁻¹ and a weight-average molecular weight (M_w) of 3031 g mol⁻¹ (**Fig.6-2**). These results indicate that the lignin retained much of its native structure. Notably, even at this low severity, lignin extracted from *Camellia oleifera* shell (COS) displayed a relatively short polymer chain length compared to typical hardwood or softwood lignins, whose M_w values often range from 5000 to 10000 g mol⁻¹.

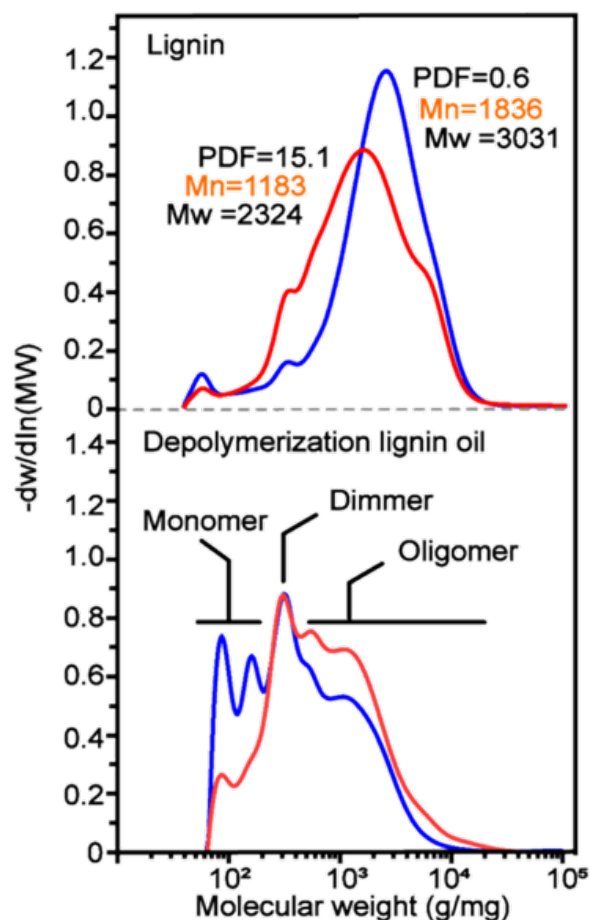


Fig.6-2. Gel permeation chromatograms(GPC) of extracted lignin before and after hydrogenolysis over increased pretreatment severities (PDF = 0.6-blue and 15.12-red). (Top) Lignin before hydrogenolysis. (Bottom) Lignin oil after hydrogenolysis. Mn is number average molecular weight; Mw is weight average molecular weight;

As the pretreatment severity increased, a non-linear trend in molecular weight evolution was observed (**Fig.6-2**). At moderate severities (e.g., PDF = 1.8 and 4.9), Mw increased significantly, reaching approximately 4600 g mol^{-1} . This increase is attributed to condensation reactions among lignin fragments, likely involving the formation of new C-C bonds and repolymerization of reactive intermediates generated upon β -O-4 ether bond cleavage (**Fig.6-4**). These condensation pathways are typically acid-catalyzed and lead to the formation of more recalcitrant, higher-molecular-weight structures. However, under harsh pretreatment conditions (PDF = 15.1), both Mn and Mw decreased sharply to 1163 and 2324 g mol^{-1} , respectively. This suggests that high-severity conditions preferentially cleave nearly all weak inter-unit linkages, such as β -O-4, while leaving

behind condensed structures that are more resistant to further breakdown. Once the labile linkages are fully depolymerized, repolymerization becomes increasingly limited due to the depletion of reactive benzylic phenol intermediates that typically serve as condensation sites.

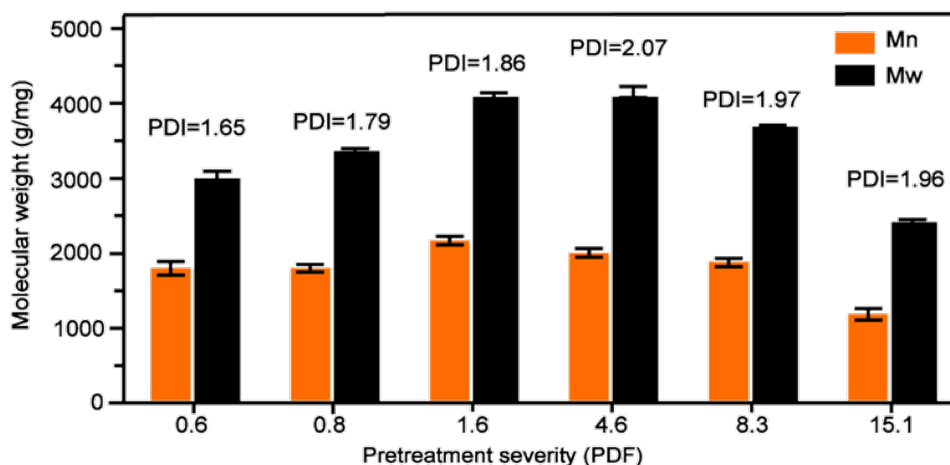


Fig.6-3. GPC results of lignin liquor over increased pretreatment severities (PDF = 0 (MWL), 0.46, 1.62, 4.61, and 15.12). PDI is the polydispersity index, given by M_w/M_n ;

The GPC profiles of the corresponding depolymerized lignin oils (**Fig.6-3**, bottom) further support this interpretation. At PDF = 15.1, the lignin oil exhibited an enhanced yield of dimers, suggesting that while some degree of repolymerization still occurs, the formation of larger polymeric species is thermodynamically or kinetically unfavorable. In contrast, the oil from PDF = 0.6 showed a higher abundance of oligomers, consistent with the preservation of larger lignin fragments rich in cleavable β -O-4 linkages. These results collectively reveal a two-stage molecular response of lignin to pretreatment: (1) an initial condensation-dominated phase under moderate severity that leads to an increase in M_w , and (2) a depolymerization-dominated regime at higher severity, which reduces both M_n and M_w due to extensive bond cleavage and limited repolymerization. This behavior underscores the importance of optimizing PDF to strike a balance between preserving reactive ether linkages and minimizing undesired condensation, thereby maximizing monomer yields in downstream catalytic valorization processes.

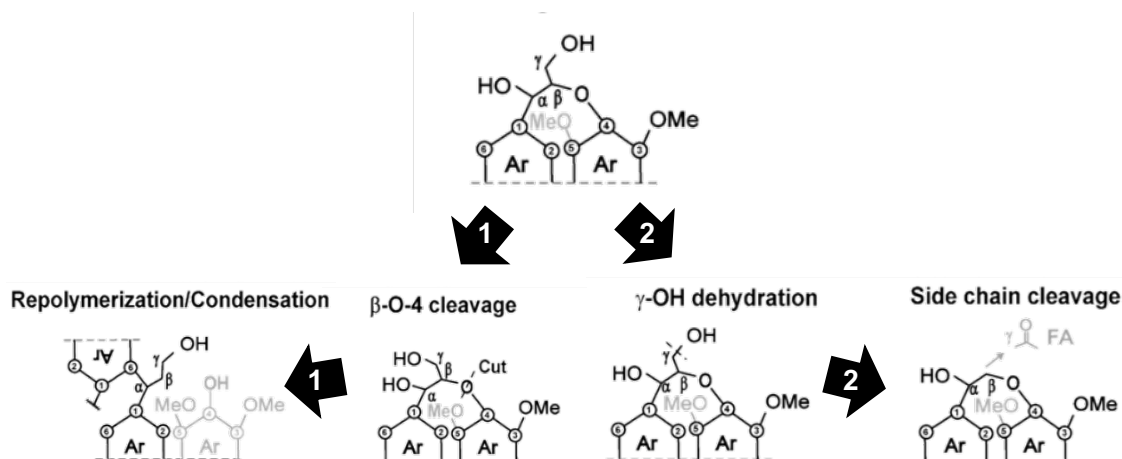


Fig.6-4. Lignin structural changes during pretreatment. Cleavage of β -O-4 linkages reduces lignin molecular weight, while free lignin fragment species promote condensation. Lignin γ -OH dehydration may induce side-chain cleavage and generate formaldehyde (FA)-like structures

6.3.3. Fourier-transform infrared spectroscopy of separated lignin

Fourier-transform infrared spectroscopy (FTIR) was applied to examine the structural features of lignin fractions obtained under different PDF conditions. All samples exhibited comparable spectra, with a strong band at 1029 cm^{-1} , assigned to the deformation of C-C, C=C, and C-OH groups. This peak likely reflected the formation of condensed linkages within the lignin. In addition, distinct aromatic skeletal vibrations were detected at 1546 , 1505 , and 1459 cm^{-1}

A broad absorption at 3302cm^{-1} as attributed to both aromatic and aliphatic hydroxyl groups. The stretching of aliphatic $-\text{CH}_2$ and $-\text{CH}_3$ groups was evident from bands at 2933 , 2885 , and 1450 cm^{-1} . Non-conjugated carbonyl linkages produced a signal at 1694 cm^{-1} , while conjugated carbonyl groups attached to aromatic structures gave rise to peaks at 1630 and 1444 cm^{-1} , likely associated with p-coumaric acid and ferulic acid residues. Syringyl structures were recognized by signals at 1392 and 1122cm^{-1} , corresponding to C-O and C-C stretching vibrations. The presence of p-hydroxyphenyl units was indicated by an absorption at 815cm^{-1} . Furthermore, the overlapping vibrations of C-C, C-O, and C-H were consolidated into a broad band around 1225 - 1029 cm^{-1} .

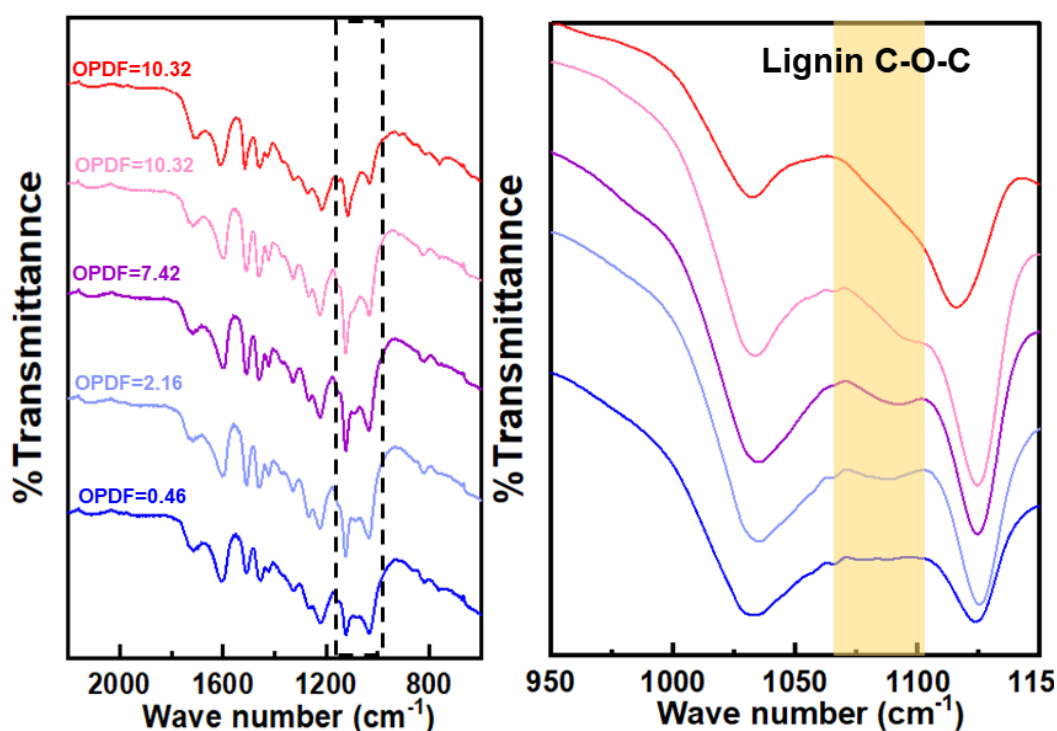


Fig.6-5. FT-IR results in different PDF fractionation condition

6.3.4. Results of catalytic transfer hydrogenolysis (CTH)

CTH cleaves C-O-C linkages in lignin and releases aromatic monomers, offering a direct means to evaluate lignin structural evolution under varying pretreatment severities through monolignol yield and distribution analysis. The chemical formula of six concerned monolignols (M1 to M6) and the corresponding GC-FID signals times

and intensities at three PDFs are presented in **Fig.6-6** and Table S1. The monolignols were selected due to their high market values and potential applicability in future. The yield of the monolignols at different PDFs was shown in **Fig.6-6** and **Table 6-2**.

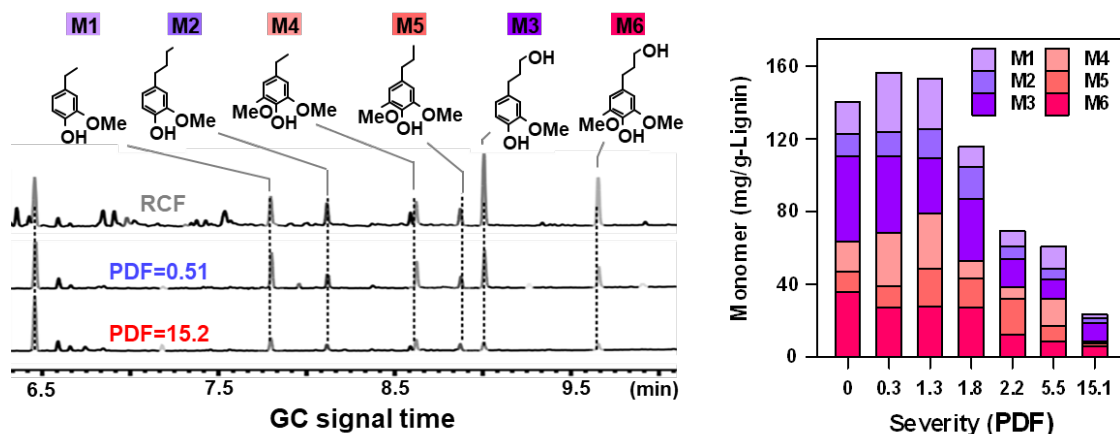


Fig.6-6. (l) GC signals of the six major monolignols produced from biomass at different PDF. (r) Productivities of the concerned monolignols over gradually increased PDFs.

When increasing PDF from 0 to 4.5, the yield of G-type monolignols showed a more significant reduction of 65.8% (from 66.2 to 25.4 mg/g-lignin) than that of the S-type monolignols at 49.8% (from 54.0 to 27.1 mg/g-lignin). The trends and the reduction ranges are consistent with our recent studies, in which fruit-remains and nutshell biomass were depolymerized by the reductive catalytic fractionation (RCF) processes (Khan et al., 2024b). The difference in yield reduction between the G- and S-monolignols can be attributed to the presence of a methoxy group at the aromatic C5 position of S-type monolignols. This methoxy group inhibits lignin re-polymerization by preventing the formation of reactive carbocation intermediates and results in lower yield of S-type monolignols. Notably, unexpected increases in the yields of M1, M2, M4, and M5 monolignols were observed when the PDF increased from 0 to 1.6. These new monolignols lack the γ -OH group or γ -carbon, implying their origin of other lignin monomers. Aliphatic γ -carbon may have undergone dehydration or oxidation during pretreatment, resulting in the cleavage of γ -OH and aliphatic side chains.

Table 6-2: Yields of G-type (M1, M2, M3), S-type (M4, M5, M6) monomers, and S/G ratios at different PDFs.

PDF	G-Type monomer			S-Type monomer			S/G ratio
	(mg/g-Lignin)			(mg/g-Lignin)			
	M1	M2	M3	M4	M5	M6	
0	17.80	12.18	46.80	16.90	11.28	35.75	0.83
0.5	32.61	13.27	42.20	29.32	11.74	27.54	0.78
1.6	11.13	17.49	34.22	9.55	15.96	27.47	0.84
4.6	12.28	5.83	10.74	14.77	8.36	8.98	1.11
15.1	2.30	2.69	9.67	1.47	1.61	5.91	0.61

6.3.5. Analysis of lignin polarity changes during pretreatment

Lignin condensation in acid-organosolv systems is initiated by a proton attack at the C α position of lignin, which facilitates the dimer formation at the side chains and hence replaces the original long-chain structure. The attacked side chains form reactive benzylic carbocation intermediates and promote electrophilic coupling reactions with other monolignol rings, especially the G-lignin units, forming a highly condensed structure. With the increase of molecular weight (Mw), the condensed polymer eventually precipitates (Li et al., 2018a; Li et al., 2024a). However, no increase in Mw of the lignin fragments was observed through Gel Permeation Chromatography (GPC) analysis in our study, even under an extremely high PDF of 15.1. The Mw of the dissociated lignin was even lower than that obtained from the “mild” pretreatment condition with a low PDF of 0.46. The “harsh” pretreatment environment did not cause the increase of Mw of fractionated lignin, the lignin re-precipitation issue shall be caused by other mechanisms rather the growth of lignin particles.

Based on the “like dissolves like” principle, we hypothesized that the polarity of dissociated lignin could change at high PDF due to the loss of hydroxyl groups. This would lead to lignin precipitation in the polar acid-organosolv and decrease delignification efficiency. To clarify our hypothesis, ^{31}P -NMR analysis was performed

to quantify the changes in lignin hydroxyl groups during pretreatment (Argyropoulos, 1994). Quantitative ^{31}P -NMR is the standard method for determining the amounts of two types of hydroxyl groups (-OH), i.e., (i) the total aliphatic-OH groups attached to the $\text{C}\alpha$ and $\text{C}\gamma$ atoms of the lignin aliphatic chains; and (ii) aromatic-OH groups found in S-lignin and G-lignin. However,³¹

P-NMR cannot distinguish the contribution of the aliphatic-OH groups attached to $\text{C}\alpha$ or $\text{C}\gamma$. Therefore, the CTH protocol was employed to specify the aliphatic-OH groups from different regions. As reported by Li and Song, the CTH process can produce monolignols without altering hydroxyl groups at $\text{C}\gamma$ (Li & Song, 2019). The experiments also provide essential information of lignin aliphatic side chains and the yields/distributions of monolignols under different PDFs.

6.3.6. Discovery of new lignin condensation mechanisms

To clarify the self-cleavage mechanism of the C-C linkages, additional pretreatment and CTH experiments were conducted, followed by advanced lignin and monolignol analysis. As 1,4-BDO would introduce new aliphatic-OH groups at $\text{C}\alpha$, which may bias the analysis, ethanol (EtOH) was applied in the pretreatment experiments at five PDFs. Elemental analysis (EA) was conducted to determine the oxygen/carbon (O/C) ratios of the harvested lignin. The detailed information of lignin-OH group contents and O/C ratios was provided in Table S3.

Based on our analysis, a new reaction route elucidating the lignin condensation mechanisms was proposed in **Fig.6-7**. Route-I represents the widely accepted mechanisms where hydroxyl removal is initiated at $\text{C}\alpha$ -OH, driving lignin condensation and β -O-4 cleavage. Route-II is the newly proposed pathway, where dehydration begins at $\text{C}\gamma$ -OH, leading to the cleavage of the lignin aliphatic side chain. Route-III is lignin self-condensation. Together with the proposed reaction routes, the aliphatic-OH, aromatic-OH, carbon, and oxygen contents were also plotted in **Fig.6-6**.

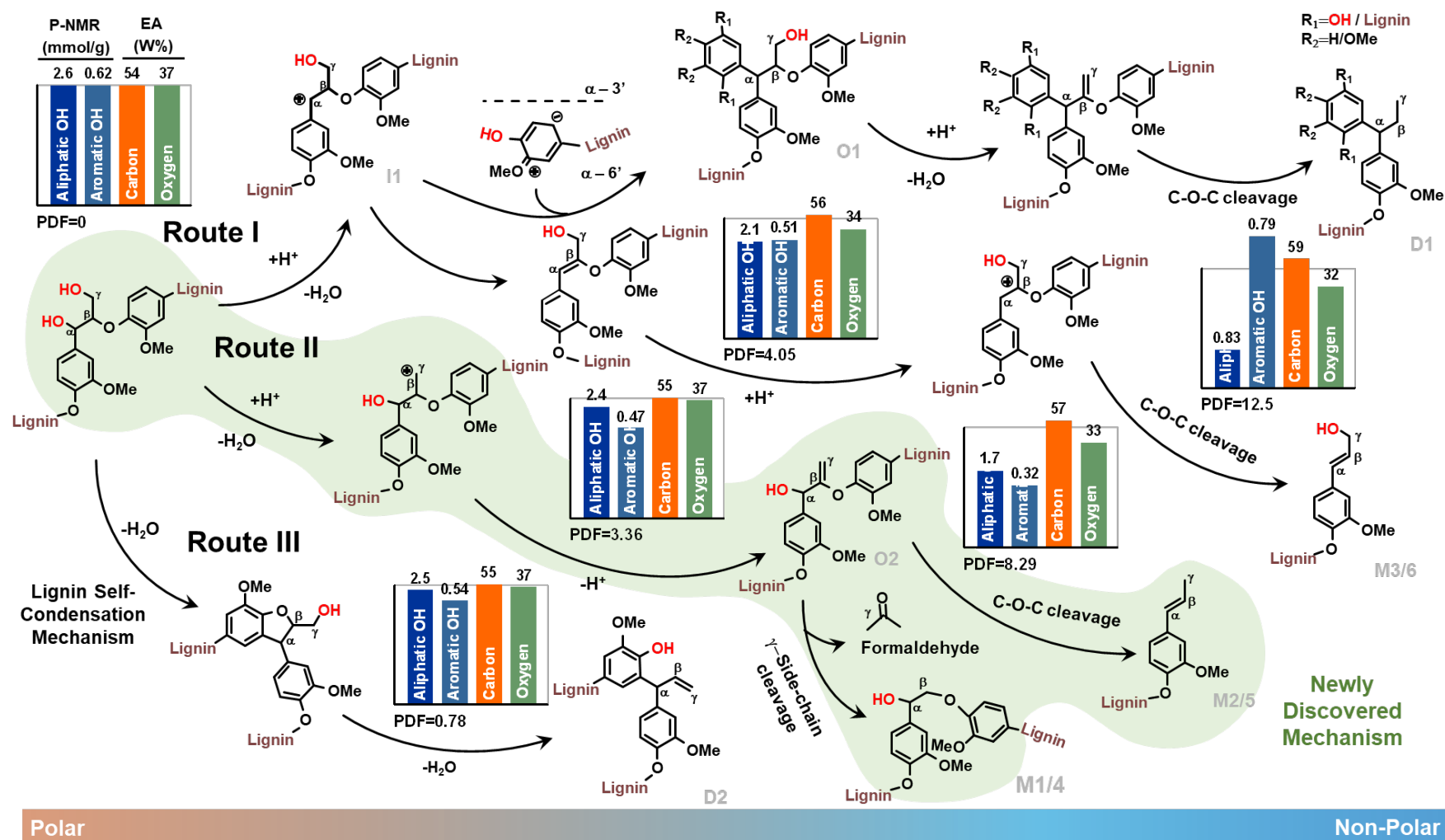


Fig.6-7. Discovery of a new lignin depolymerization and condensation pathway. The reaction route is proposed based on the results of elemental analysis (EA), ^{31}P and 2D-HSQC NMR, associated with the PDF simulation. The bar charts show the changes in lignin characteristics over gradually increasing PDFs, including aliphatic-OH content (dark blue), aromatic-OH content (light blue), carbon mass ratio (orange), and oxygen mass ratio (green).

In Route-I, acid-induced dehydration begins at the α -OH of lignin, resulting in the formation of an α -benzylic carbocation intermediate (I1), which subsequently re-polymerizes with another aromatic fragment and forms the oligomeric intermediate (O1) at the $C\alpha$ position after deprotonation. The O1 intermediate then undergoes further dehydration, leading to the cleavage of β -O-4 linkages and production of dimmer structure (D1) with a monomeric fragment (P1). Alternatively, in another pathway, the proton transfer from $C\beta$ to $C\alpha$ can induce self-dehydration at the α -position, forming an unstable structure with a double bond and ether linkage. This instability leads to ether bond cleavage, yielding a polymer precursor to long side-chain monomers (M3/6). However, our GPC analysis suggested that Route-I may not be the only condensation mechanism in acid-organosolv pretreatment system. A significant decrease in aliphatic-OH content from 2.6 to 0.83 mmol/g ($\sim 68.1\%$) was observed when PDF increased from 0 to 12.54, indicating a reduction in lignin polarity during pretreatment. The reduction of hydroxyl groups at only the α -position seems cannot to fully explain the observed phenomenon.

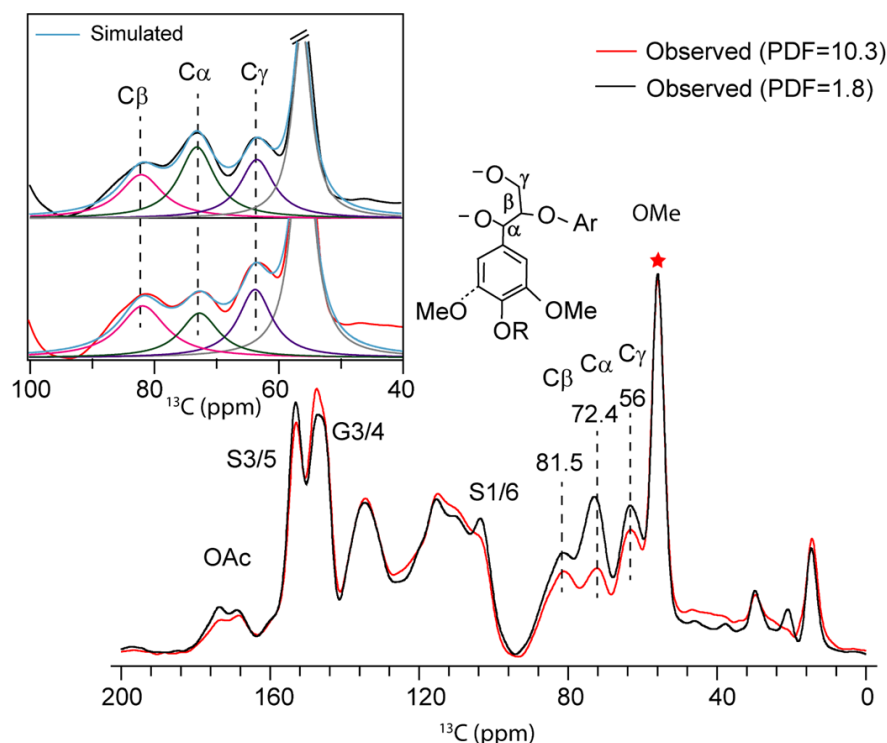


Fig.6-8. ^{13}C multi-CP NMR spectra of lignin at different PDF conditions (PDF=1.8 and 10.3). Distinct decreases in both $C\alpha$ -OH and $C\gamma$ -OH signals are observed, confirming that α -OH loss accompanies γ -OH depletion.

To further substantiate the proposed γ -OH-initiated cleavage pathway, solid-state ^{13}C multi-CP NMR spectra were obtained under different PDF conditions (**Fig.6-8**). In addition to the expected decrease of the $\text{C}\alpha$ -OH resonance, a clear reduction of the $\text{C}\gamma$ -OH signal was also detected in the aliphatic region. This simultaneous attenuation of both α -OH and γ -OH resonances provides direct structural evidence that γ -OH depletion occurs alongside α -OH loss, supporting the mechanistic assignment of the γ -OH-initiated $\text{C}\beta$ - $\text{C}\gamma$ bond cleavage pathway.

Based on the experimental results, we propose a new reaction route (Route-II), where dehydration could first occur at the γ -OH position, forming a γ -benzylic cation. This cation then triggers proton movement, which subsequently leads to the formation of a β -benzylic cation intermediate (O2). With the formation of this intermediate, further dehydration could occur due to the activation of the α -OH structure, potentially causing self-oxidation at the β -position due to the presence of the β -benzylic cation. This sequence of events could ultimately result in the cleavage of the lignin side chain, producing precursors for short-chain lignin monomers. Given that the pretreatment conditions are relatively mild and without a catalyst specifically targeting C-C bond cleavage, significant side-chain cleavage (such as the 45% conversion of M3 to M1) is unlikely. Therefore, we propose that the acid-induced dehydration at the γ -OH position may serve as the major driven force for the self-oxidation reaction at the $\text{C}\beta$ position of lignin, which eventually causes the C-C bond cleavage. This may also explain the loss of hydroxyl groups in the aromatic monomers found in the depolymerization products.

6.3.7. Application of PDF in lignin-first biorefinery

The new findings are incorporated into the PDF model to simulate the lignin condensation kinetics. The simulation results of the delignification status, dynamic changes in β -O-4 linkages, and monolignol yields fit well with the experimental results (**Fig.6-10a**). When the PDF was increased from 0 to 6, the delignification efficiency (blue curve) increased from 11% to 87%, but the delignification efficiency gradually decreased when the PDF exceeded 6 and reached 77% when PDF was 18. Although the highest delignification efficiency was achieved when PDF was 6, only 27% of the initial

cleavable C-O-C linkages ($C_{\beta-O-4}$) remained in the lignin. The low $C_{\beta-O-4}$ leads to the reduction of final monolignol yield.

The PDF functions were optimized for improving the monolignol yield in lab-scale experiments (**Fig.6-10b**). The derivatives of simulated lignin quality (blue), lignin yield (red), and the integrated monolignol yield (black) are presented as a function of PDF. PDF was controlled at as low as 1.6-4.2, which was considered appropriate to prevent a significant drop in lignin quality, although high lignin fractionation efficiency can be maintained until PDF reached 6.3. The function of PDF was demonstrated in the process upgrade of a biorefinery. At a target PDF of 1.6, the pretreatment liquid including dissolved lignin and hemicellulose was discharged and replaced by new solvent. The operational mode of the process was therefore changed from a batch to a sequencing batch reaction (SBR). The profile of applied solvent over pretreatment time is shown in **Fig.6-10c**.

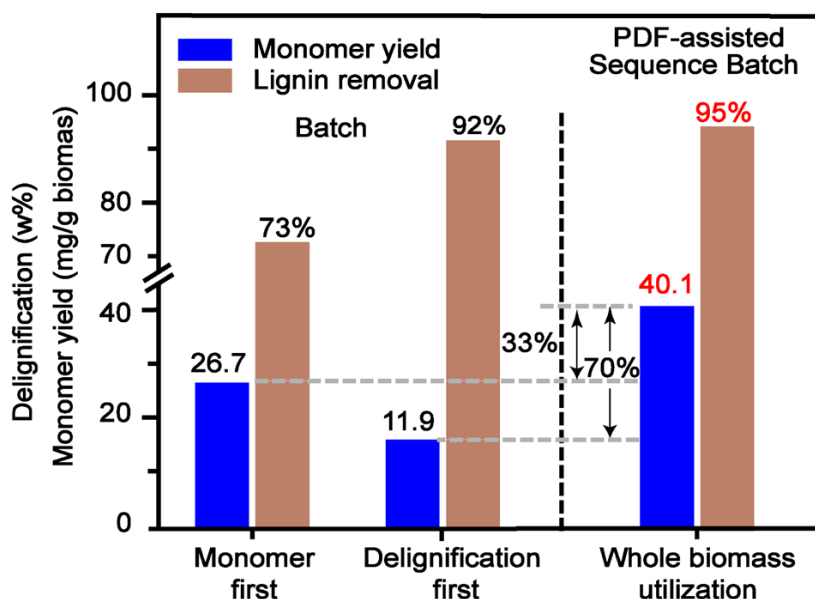


Fig.6-9. Comparison of monomer yields and lignin removal efficiency under different strategies. The monomer-first strategy achieved ~26.7 mg/g-biomass with limited lignin removal. The delignification-first strategy yielded only 11.9 mg/g-biomass. The PDF-assisted sequencing batch strategy achieved the highest yield (40.1 mg/g-biomass) with 95% lignin removal.

In the batch process, all 120mL of solvent was loaded for a total reaction time of

6 hours, while the same amount of solvent was divided into four fractions (30 mL each) to be used in the pretreatment. After each sub-batch, the pretreatment spent liquid fraction was replaced with fresh solvent until the total of 120 mL was used. All the spent liquid was combined and used in the CTH process for monolignols production.

The benefits of PDF-controlled SBR were clearly demonstrated in the experiment. While lignin removal rates were only 4% higher in SBR than the batch (**Fig.6-10d**), the $C_{\beta-O-4}$ in SBR reached 51.2% of the theoretical yield after 6 hours of pretreatment, which was only 13.2% in the batch process (**Fig.6-10e**). The total lignin monomer yield in SBR increased significantly from 2.5% to nearly 10% after process optimization (**Fig.6-10f**), despite using the same total solvent, catalysts, pretreatment temperature, and time.

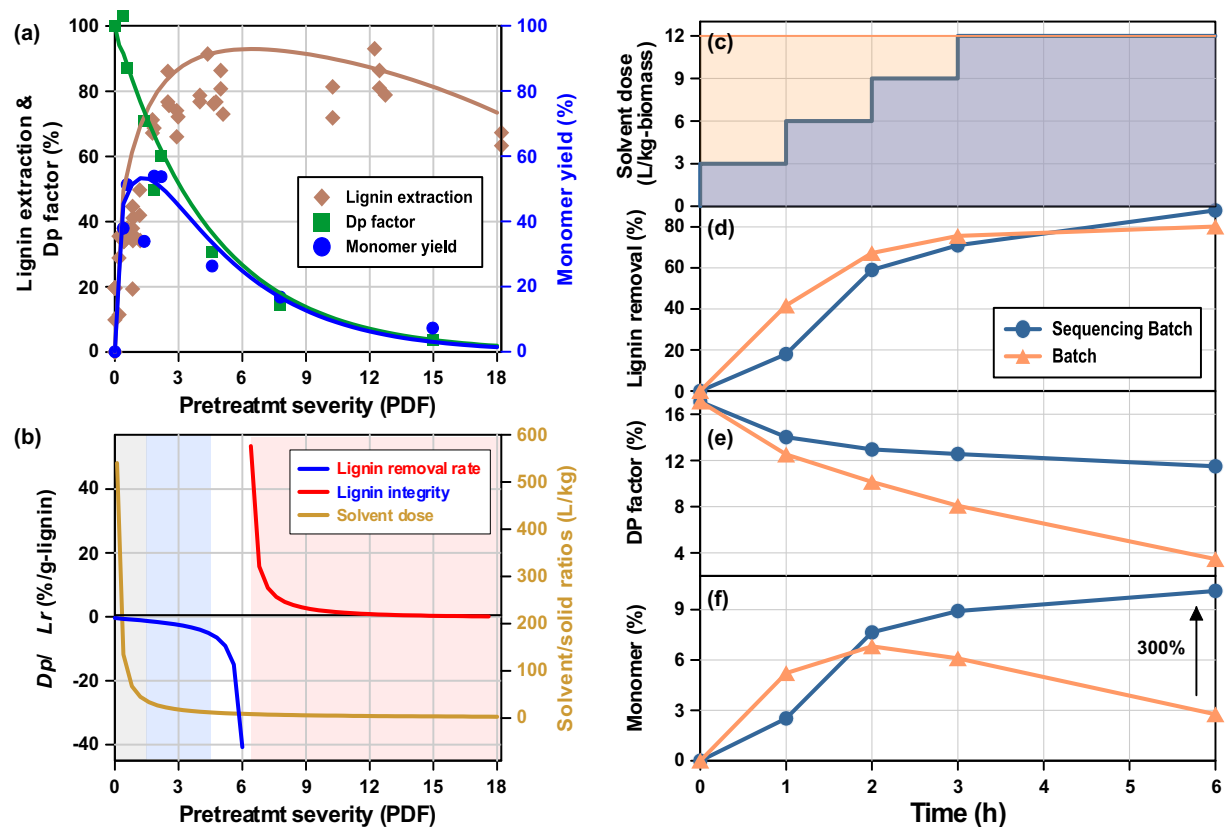


Fig.6-10. Simulation results and application of the PDF model. (a) The brown diamonds show the lignin extraction efficiency; green squares show the depolymerization efficiency calculated based on the β -O-4 content (DP factor = Cb-O-42); blue spheres show the monomer yields. The curve fittings are PDF simulation results, which indicate an optimal depolymerization yield at a PDF of 1.6. (b) Model-based analysis of the decline rate in lignin extraction efficiency per unit shows that a PDF range of 1.6 to 4.2 provides optimal lignin protection without a significant increase in solvent consumption. (c) The operational conditions of PDF simulated pretreatment operations. The light orange lines and symbols show the process design conditions for single batch process and blue figure shows the PDF optimized sequencing batch operation with which organosolv is divided into four parts during pretreatment. The experimental results show a notable increase in depolymerization factor (Dp,%) and lignin monomer yield (%), although lignin removal (%) does not change significantly in the two processes.

Conclusions

The PDF model combined a set of numerical indexes unifying the reaction severities of lignin-first biorefineries, particularly for the accurate simulation of acid-organosolv pretreatment followed by CTH processes. With the standardization of PDFs, advanced instrumental analysis can be applied to quantify the dynamic behaviors of building-block fractionation and condensation, and hence to elucidate the corresponding mechanisms. Based on the PDF-associated solid-state NMR analysis, a new lignin reprecipitation pathway was discovered, highlighting the critical role of hydroxyl group depletion and its subsequent effects on the changes in lignin polarity during delignification. In addition, the newly constructed PDF-based control algorithm successfully upgraded our lab-scale biorefinery system resulting in an overall increase of lignin fractionation efficiency to over 90%, with 300% higher lignin monomers yield. In the future, the PDF could be further expanded to support the design and optimization of the thermal catalytic continuous flow stirred-tank reactor CTH system for an integrated biorefinery.

CHAPTER 7. REVEALING LIGNIN INTERMOLECULAR AND INTRAMOLECULAR INTERACTIONS IN PLANT CELL WALLS BY SOLID-STATE NMR

ABSTRACT

Woody biomass is a major renewable feedstock for biofuels and bioproducts, yet its recalcitrance to deconstruction is governed by the molecular architecture of lignin and its interactions with polysaccharides. In this study, we employed uniformly ^{13}C -labeled *Eucalyptus grandis* cell walls and a combination of quantitative 1D and multidimensional solid-state NMR (SS-NMR) techniques to resolve the nanoscale organization of lignin-carbohydrate complexes (LCCs) and monitor their progressive disruption during stepwise dilute-acid pretreatment. Quantitative ^{13}C direct-polarization spectra revealed rapid early-stage delignification dominated by the removal of ester-linked lignin, accompanied by relative cellulose enrichment. High-resolution 2D spectra resolved native polysaccharide connectivity, including distinct xylan conformers and glucuronic acid substitutions, which disappeared with increasing pretreatment severity. Dipolar-gated ^{13}C - ^{13}C correlation spectroscopy distinguished intramolecular lignin linkages from long-range lignin-lignin and lignin-xylan contacts, enabling the mapping of specific cross peaks to structural motifs such as ferulate- and acetyl-mediated linkages. These long-range contacts diminished sequentially, revealing the preferential cleavage of ester bonds followed by ether- and C-C-linked aromatic networks. The convergence of short- and long-range spectral patterns in the most treated samples indicates the persistence of a condensed, spatially homogeneous lignin domain that likely underlies residual biomass recalcitrance. Our findings provide a molecular framework for understanding lignin distribution, linkage heterogeneity, and interfacial organization in hardwoods, offering new structural targets for genetic and process-based strategies to improve biomass fractionation efficiency.

7.1. Introduction

Terrestrial plants form secondary wall through the formation of cellulose microfibrils coated with hemicellulose and lignin (Zhong & Ye, 2015). Among various linkages, lignin-carbohydrate complexes (LCCs) are formed through covalent bonds, including benzyl ether, benzyl ester, and phenyl glycosidic linkages, and these covalent connections, particularly ester linkages involving ferulate or p-coumarate are thought to play key roles in crosslinking lignin to hemicelluloses (Mnich et al., 2020). While diffraction and imaging methods have revealed aspects of cellulose microfibril packing and lignin distribution, the precise chemical nature and spatial arrangement of lignin-polysaccharide interfaces in native biomass remain elusive. This knowledge gap is due in part to the nanoscale dimensions of the interface, its chemical heterogeneity, and the low abundance of certain interfacial motifs that are easily lost or altered during traditional fractionation.

Recent applications of multidimensional solid-state NMR (SS-NMR) have provided unprecedented molecular-level insights into plant cell walls, revealing how lignin associates with different xylan conformations and how aromatic-aromatic and aromatic-carbohydrate interactions are organized in native structures (Fernando et al., 2023). It has emerged as a powerful, non-destructive tool for probing plant cell wall structure at the molecular level. However, much of this knowledge derives from model plants such as *Arabidopsis* or maize, leaving open questions about whether similar structural principles apply to woody biomass. In hardwoods, for example, the balance between G- and S-type lignin, the occurrence of special lignin substructures such as S' or β -5-linked G units, and the persistence of acetylated lignin residues may all influence both wall architecture and pretreatment response.

Early breakthroughs were made by Wang et al. (2014), who employed SS-NMR to investigate the primary cell wall of *Brachypodium distachyon*, revealing that glucuronoarabinoxylan (GAX) was the dominant hemicellulose in grasses (Wang et al., 2014). That study identified unique GAX-cellulose interactions distinct from those in dicots such as *Arabidopsis thaliana*, offering the first molecular-scale model of 3D wall organization in Poaceae species. Wang further extended this work in 2016 by performing quantitative ^{13}C SS-NMR using direct polarization (DP) and cross

polarization (CP) methods, coupled with 2D INADEQUATE analysis. These methods enabled the differentiation between core and surface cellulose and the mapping of molecular proximities, establishing foundational knowledge on polysaccharide-cellulose assembly. This body of work was expanded by Terrett et al. (2019), who examined the SCW architecture in Norway spruce. Their findings revealed that lignin preferentially surrounds cellulose clusters, with galactoglucomannan (GGM) acting as a key hemicellulose connector. In softwood, xylan was shown to adopt a two-fold screw conformation for cellulose binding, while GGM formed semi-crystalline domains via hydrogen bonding with cellulose microfibrils. This work provided a revised model for softwood walls, refining our understanding of lignin-polysaccharide assembly. Concurrently, Kang et al. (2019) employed ^{13}C -labeling and dynamic nuclear polarization (DNP) to probe Arabidopsis and switchgrass, revealing that lignin forms hydrophobic nanodomains and interacts preferentially with Xn^{2f} or Xn^{3f} (Kang et al., 2019). These observations reinforced the idea that non-covalent electrostatic and hydrophobic forces mediate lignin-xylan interaction, particularly in grass species. Recent progress was furthered by Kirui et al. (2022), who used DNP-enhanced SS-NMR to compare cell wall component distributions across grasses, hardwoods, and softwoods (Kirui et al., 2022). Their study identified a progressive increase in glycan-aromatic interactions, with softwoods showing strong water retention capability due to homogenous lignin-polysaccharide arrangements. Notably, lignin in softwoods primarily associates with Xn^{3f} , with occasional contact at cellulose-xylan junctions, contributing to structural rigidity and recalcitrance. The most recent advancement came from Bennett Addison et al. (2024), who utilized multi-CP SS-NMR in *Populus* to assess the spatial configuration and interaction intensity among cellulose, hemicelluloses, and lignin in SCWs (Addison et al., 2024). By integrating NMR observables into molecular dynamics simulations, this study delivered an atomistic view of the higher-order organization of cell walls, providing a more predictive framework for evaluating biomass digestibility.

In this study, we investigate the native lignin architecture and lignin-polysaccharide interactions in ^{13}C -enriched eucalyptus using a combination of

quantitative 1D and multidimensional solid-state NMR techniques, including INADEQUATE and dipolar-gated 2D ^{13}C - ^{13}C correlation spectroscopy. Stepwise dilute-acid fractionation was used to selectively disrupt ester- and ether-linked LCCs, enabling the progressive isolation of lignin substructures with distinct functional linkages. Through spectral deconvolution, chemical shift assignment, and intra-/intermolecular correlation mapping, we identify structural motifs that are preferentially removed or preserved during pretreatment and construct a detailed molecular map of the eucalyptus lignocellulosic interface. These findings deepen our understanding of lignin distribution, linkage heterogeneity, and packing within woody biomass, providing a structural framework for designing more effective biomass deconstruction strategies.

7.2. Materials and method

7.2.1. Biomass feedstock and chemicals

Eucalyptus chips were provided by the Leizhou Forestry Bureau (Guangdong, China). The chips were manually debarked, Wiley-milled, and screened to collect fractions smaller than 2 mm. Uniformly ^{13}C -enriched stems (97% ^{13}C) of *Eucalyptus grandis* (16 weeks old) were obtained from IsoLife (The Netherlands) for solid-state NMR analysis. The plants were cultivated from seeds under controlled conditions. Upon harvesting, shoots were separated into leaves and stems. The stems were freeze-dried, longitudinally split, sectioned, and carefully debarked by incising the bark and peeling it from the xylem at the cambium interface. Debarked stems were coarsely cut with scissors and gently ground using a mortar and pestle into fragments a few millimeters in size. Ball milling was intentionally avoided to preserve the native ultrastructure of the biomass. All reagents and chemicals were of analytical grade (Sigma-Aldrich, MO, USA; and J&K Acros Organics, Beijing, China).

7.2.2. Lignin dissociation experiment

In a 50-mL glass vial, 0.5 g of freeze-dried, uniformly ^{13}C -enriched eucalyptus stem material (97% ^{13}C ; particle size <1 mm) was mixed with 9 mL of 1,4-dioxane, 420 μL of hydrochloric acid (37%), and 1 mL of deionized (DI) water. The reaction was carried out using a hot plate set to the target temperature (80 $^{\circ}\text{C}$). Temperature was monitored using an infrared thermometer. As the distance between the liquid surface and the hot plate was less than 2 cm, temperature distribution within the vial was uniform. The mixture was stirred at 300 rpm using a magnetic stir bar. At predetermined time points (0, 1, 3, 18, and 36 hours), approximately 500 μL of the solvent phase containing the solid-liquid mixture was withdrawn using a pipette. To allow the passage of solid particles, the pipette tip was trimmed to create an opening >1 mm. The sample was transferred into a 1 mL microcentrifuge tube and centrifuged at 800 rpm for 3 minutes. The supernatant was carefully returned to the original glass vial to continue the reaction. The solid adhering to the centrifuge tube wall was washed with DI water and centrifuged again to remove the solvent and acid. This step was repeated until the supernatant became

colorless. The resulting solid was gently dried at 80 °C for 10-15 minutes to retain slight moisture, thereby minimizing structural changes caused by drying and rehydration. The partially dried sample (~80mg) was then packed into rotors for solid-state NMR analysis. A parallel procedure using non-labeled sample was conducted for comparison, and each condition was repeated in triplicate.

7.2.3. Preparation of lignin samples

To analyze the original lignin structure in eucalyptus sample. Native eucalyptus lignin samples were prepared to gather the structural information of the original lignin in the biomass before pretreatment. The dry eucalyptus sample was Soxhlet-extracted with ethanol/hexane (2:1,v/v) for 48 hours to remove extractives. The extraction-free samples were then dried in an oven at 80 °C for 48 hours prior to further use. Extractive-free eucalyptus was ball-milled at 450 rpm, with a 10-minute rest period between each of the 10-min milling cycles, for a total of 8 to 12 hours. The milled powder was solvent-extracted using 96:4 dioxane-water (v/v), at a solvent-to-powder ratio of 10:1 (v/w). The mixture was continuously stirred for 24 hours to ensure thorough extraction of lignin. For the fractionated lignin, the pretreatment spent liquor was filtrated by Advantech 5A filter paper. The reactor and substrate were washed with a solvent mixture three times (10 mL each), and then the solution was mixed with the filtrate of the pretreatment liquor. The solid residue was rinsed with distilled water until the solution was colorless, then oven-dried at 105 °C before chemical composition analysis. The lignin-rich solution (10 mL) volume was transferred to 50mL centrifuge tubes, which were then filled with distilled water. The mixture was stirred and left for 30 minutes for lignin precipitation. The precipitated lignin was separated from the solution using a centrifuge then freeze-dried before further analysis.

7.2.4. Solid-state NMR experiment

SS-NMR experiments for resonance assignment. Solid-state experiments were conducted on a Bruker Avance 600 MHz (14.1 Tesla) spectrometer and a 400MHz (9.4 Tesla) Bruker Avance spectrometer using 3.2-mm and 4-mm MAS probes respectively. ¹³C chemical shifts were externally referenced to tetramethylsilane (TMS) by calibrating

the adamantane CH₂ resonance to 38.48 ppm, with this spectral reference applied to all plant spectra. One-dimensional ¹³C spectra were obtained using both ¹H-¹³C cross-polarization (CP) and direct polarization (DP) schemes. CP experiments selectively detected rigid molecular components such as lignin and crystalline polysaccharides, while DP spectra acquired with a 35 s recycle delay quantitatively detected all carbons in the cell wall. To further resolve molecular connectivity within cell wall components, two-dimensional CP-based refocused INADEQUATE experiments were performed. This double-quantum-single-quantum (DQ-SQ) approach, mediated by ¹³C-¹³C J-coupling, provided unambiguous assignments of directly bonded carbon pairs in polysaccharides, enabling the identification of inter-unit linkages. Homonuclear ¹³C-¹³C correlation experiments were subsequently conducted using proton-driven spin diffusion (PDSF) with a 1 s mixing period to probe long-range intermolecular contacts within ~10 Å, thereby revealing spatial correlations within the molecular network. Due to the relatively low abundance and weaker signal intensity of aromatic carbons, gated-DARR experiments with a 100 ms mixing period were employed to selectively suppress protonated carbohydrate signals while preserving deprotonated aromatic carbons. This approach enhanced the detection of lignin aromatic motifs and enabled detailed investigation of lignin-polysaccharide spatial interactions in the eucalyptus cell wall.

7.3. Results and discussion

7.3.1. Structural resolution of eucalyptus cell walls by ¹³C SS-NMR

To investigate the molecular architecture of eucalyptus cell walls, quantitative ¹³C DP spectra with a 35 s recycle delay were acquired. The ¹³C DP-35 s spectra exhibited high spectral resolution, particularly for aromatic and mobile molecular domains, enabling clear separation of minor resonances from the main polysaccharide peaks. For example, the GlcC2 resonance is clearly resolved from the adjacent C5 and xylan signals, and the GlcC1 anomeric carbon peak is distinctly visible despite its proximity to overlapping polysaccharide resonances. Such spectral clarity provides a robust basis for differential spectroscopy and peak deconvolution analyses.

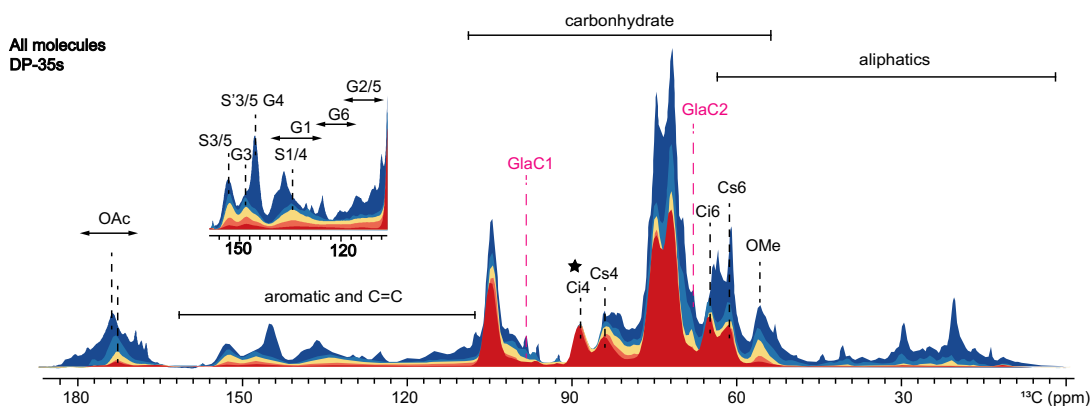


Fig.7-1. Structural heterogeneity of eucalyptus cell walls revealed by quantitative ^{13}C DP-35s spectra. From top to bottom: untreated (blue), 1 h (light blue), 3 h (yellow), 18 h (orange), and 36 h (red) treatments. Signals above 170 ppm arise from lipid carbonyls; 100-160 ppm corresponds to aromatic carbons (mainly lignin, expanded in inset); 100-60 ppm contains polysaccharide carbons including cellulose, hemicellulose, and lignin C1-C5/6; the lignin methoxy group resonates at 56.5 ppm; signals below 50 ppm originate from alkyl carbons in aromatic side chains. Progressive spectral changes reflect compositional and structural evolution during treatment. All spectra were collected from untreated and pretreated eucalyptus on Bruker Avance 600 MHz (14.1 T) with 3.2 mm MAS probes, spinning at 14 kHz and 98 K.

In the carbonyl region, strong lipid carbonyl resonances were observed between 170-175 ppm, as expected, but eucalyptus samples also displayed several additional minor peaks below 170 ppm. These peaks may arise from carbonyl carbons influenced by conjugation with aromatic units, reflecting possible ester-aromatic interactions in the cell wall matrix. The aromatic region (100-160 ppm) revealed distinct separation of S-type lignin resonances, whereas the G-type lignin C1/2 and C6 signals appeared as broader, less resolved peaks. This broadening suggests greater structural heterogeneity in the side chains of G-type lignin compared to the more uniform S-type units. Peaks attributed to methoxy carbons were detected at 56.5 ppm, consistent with lignin $-\text{OCH}_3$ groups. Signals below 50 ppm corresponded to alkyl carbons in aromatic side chains and other nonpolar moieties. As pretreatment progressed, a gradual attenuation of aromatic ^{13}C signals was observed, accompanied by the complete disappearance of the prominent peak at ~ 145 ppm, which is typically assigned to non-protonated aromatic carbons in S-

and G-type lignin bearing either or conjugated linkages.

Based on the quantitative ^{13}C DP-35 s spectra, delignification and cellulose retention during pretreatment were determined by dividing the spectra into two integration regions. Lignin content was quantified by integrating the aromatic region (105-160 ppm) together with the methoxy signal at ~ 56.5 ppm ($-\text{OCH}_3$), while cellulose content was determined from deconvoluted $\text{C}_{\text{i}4}$ and $\text{C}_{\text{s}4}$ peaks. The calculated values are summarized in **Table 7-1**. The results show a progressive reduction in lignin content from 0.239 (untreated) to 0.018 (36 h), corresponding to a degree of 92.4%. Significant lignin removal occurred within the first 3 h, reaching 65.4% delignification, with the most rapid change observed between 1 h and 3 h. In contrast, cellulose content increased steadily from 0.278 in the untreated sample to 0.775 after 36 h, reflecting the relative enrichment of cellulose as lignin and hemicellulose were removed. Notably, cellulose content increased sharply between 0 h and 1 h (0.278 to 0.467), indicating that early-stage pretreatment rapidly enriched cellulose in the substrate. These quantitative changes are consistent with the spectral evolution in **Fig.7-1**, where attenuation of aromatic signals and enhancement of cellulose (C_{i} and C_{s}) peaks reflect highly efficient lignin dissociation. To minimize lignin condensation during this process, the experiments were conducted under very mild reaction temperatures and with a high solvent-to-solid ratio.

Table7-1: Quantitative analysis of lignin content, delignification degree, and cellulose content in eucalyptus substrates during pretreatment, calculated from ^{13}C DP-35 s spectra. Lignin content was determined by integrating the aromatic region and the methoxy peak, while cellulose content was obtained from deconvoluted Ci4 and Cs4 peaks.

Reaction time(h)	Cellulose(C%)	Lignin(C%)	Delignification
0	27.8%	23.9%	0.0%
1	46.7%	23.7%	42.0%
3	41.7%	19.3%	51.4%
18	65.7%	22.2%	65.4%
36	73.6%	17.1%	75.1%

Two-dimensional ^{13}C - ^{13}C refocused INADEQUATE spectroscopy was employed to resolve carbon-carbon connectivities in ^{13}C labeled eucalyptus cell walls and to elucidate structural changes in polysaccharides during pretreatment. This correlation technique provides direct through-bond information mediated by ^{13}C - ^{13}C J-coupling, enabling the unambiguous assignment of carbon pairs in polysaccharide backbones and side chains. This approach enables atomic-level assignment of carbon linkages and conformations in cellulose and hemicelluloses, including glucuronoxylan, arabinoxylan, galactoglucomannan, and trace xyloglucan. Overlaid CP- (gray) and DP-based (orange) INADEQUATE spectra distinguish rigid domains such as cellulose and lignin from mobile domains including hemicellulose side chains. In the untreated sample, strong correlations from hemicellulose components, glucuronic acid (GlaC1, GlaC2), arabinose (A), and multiple xylan conformers (Xn , Xn^{2f} , Xn^{3f}) are clearly resolved, with dashed lines tracking carbon connectivities along xylan backbones. The narrow ^{13}C linewidths of 0.5-0.9 ppm highlights the high spectral resolution achieved, enabling detailed inspection of polysaccharide polymorphism and unambiguous connectivity assignments. After 36 h of pretreatment, these hemicellulose-derived correlations are largely absent, while cellulose cross peaks from both interior and exterior surfaces remain well resolved.

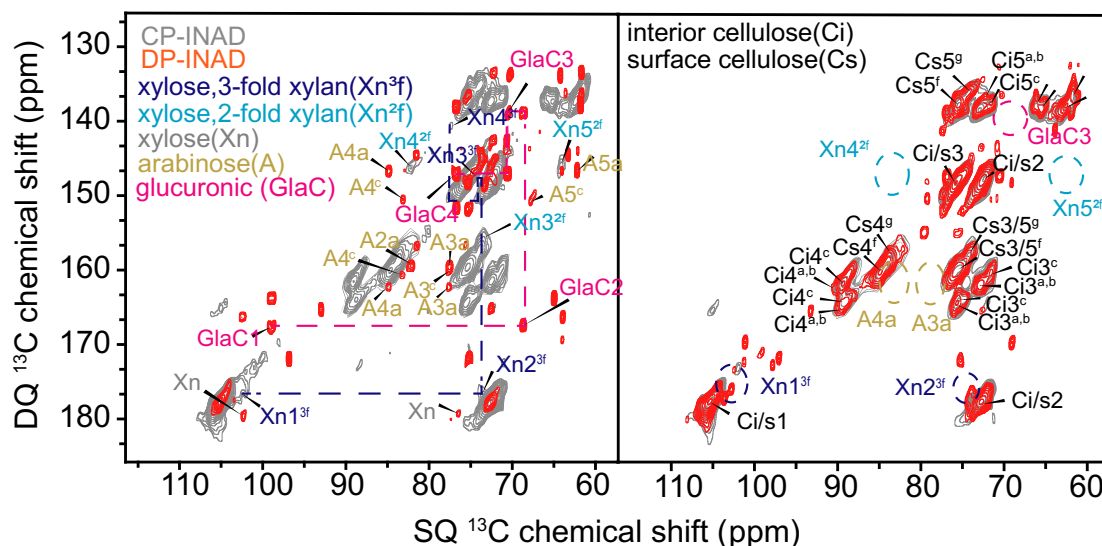


Fig.7-2. Representative 2D ^{13}C refocused INADEQUATE spectra of eucalyptus cell walls acquired with cross polarization (CP, gray) selecting rigid domains and direct polarization (DP, orange) enhancing mobile domains. Spectra were recorded on a 600 MHz spectrometer with 14 kHz MAS and ^1H TPPM decoupling at 83 kHz. The left panel shows untreated eucalyptus, highlighting carbon connectivity of non-rigid molecules assigned by DP-INADEQUATE, including glucuronic acid (GlaC), arabinose (A), and xylan (Xn) conformers; dashed lines trace the carbon connectivity within xylan. The right panel shows material after 36 h pretreatment, where residual cellulose dominates; both interior and exterior cellulose signals are clearly resolved in the INADEQUATE spectra, consistent with whole-cell wall analysis Sequential cleavage of ester- and ether-linked lignin structures

Although 2D ^{13}C - ^{13}C INADEQUATE spectroscopy can effectively resolve C-C intramolecular connectivity, the relatively low abundance, weaker signal intensity, and lower structural repetitiveness of lignin make it challenging to directly determine lignin intramolecular linkages from INADEQUATE spectra. To elucidate the native lignin architecture in biomass, a stepwise dilute-acid fractionation strategy was applied to the ^{13}C -labeled eucalyptus sample, exploiting the differential acid stability of ester and ether linkages in lignin. This approach enabled the progressive separation of lignin substructures with distinct functional linkages, consistent with the pronounced differences observed in the aromatic-region difference spectra obtained at different

fractionation stages.

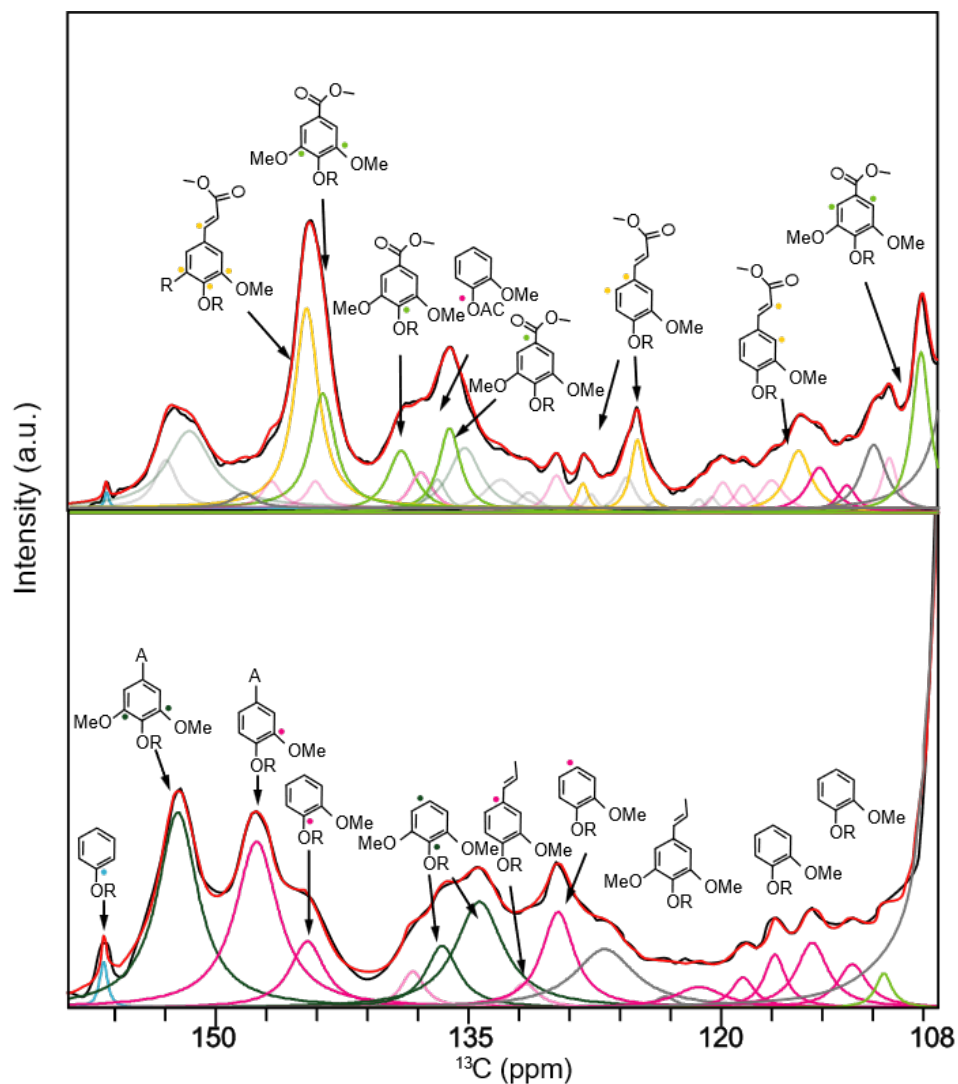


Fig.7-3. Peak deconvolution of aromatic-region difference spectra obtained from quantitative ^{13}C -DP(35s) NMR under different pretreatment conditions. Difference spectra were generated by subtracting spectra between two time points: (top) untreated sample and 1h pretreatment; (bottom) 1h and 3h pretreatment. Black lines represent experimental spectra, and red lines represent simulated fits from peak deconvolution. Colored dots indicate assignments of adjacent carbons for specific lignin structures, with each color corresponding to the peak color in the deconvoluted fit.

The upper panel of **Fig. 7-3 (top)** presents the difference spectrum in the aromatic region for the fraction enriched in ester-linked lignin structures. A prominent resonance at ~145 ppm rises well above the surrounding baseline, indicating substantial peak

overlap in this region. Based on literature reports and further supported by gated-DARR analysis, this resonance is primarily assigned to the C3-C5 carbons of S' units and the C3, C4, C5, and C7 carbons of ferulic acid (FA) moieties. Peak deconvolution resolved several additional FA-associated resonances, lending strong support to this structural assignment. These features indicate that the initial stages of fractionation predominantly remove lignin substructures with labile ester linkages, particularly S' and FA units, which are likely covalently linked to hemicelluloses through ester bonds. The preferential removal of these motifs is consistent with their lower acid stability and their tendency to occur at lignin-carbohydrate junctions, facilitating early-stage delignification.

In contrast to the upper panel, where ester-linked resonances dominate, the remaining aromatic signals in the lower panel are largely derived from S- and G-type lignin units. The most intense peak corresponds to S3/5 (~153.3 ppm), whose high intensity arises from contributions of two carbons at these positions. Comparison of the S3/5 and G3 (~147.0 ppm) peak intensities suggests that G-type lignin undergoes faster removal than S-type lignin at this stage. Moreover, G units exhibit a greater number of well-resolved peaks in the C1 (~126.5 ppm), C2 (~111.2 ppm), and C6 (~118.8 ppm) regions, each with distinct chemical shifts, indicating the presence of multiple structural conformations. These features imply that the separated G-type lignin at this stage may include a diversity of interunit linkages, and that β - β or β -5 motifs are likely associated with G-type units in the residual lignin fraction. The persistence of these peaks in the lower panel suggests that the lignin remaining after ester-linked removal is enriched in more condensed and ether-linked structures, requiring harsher conditions for cleavage. This progressive shift from ester-linked to ether- and C-C-linked motifs illustrate the sequential breakdown of lignin networks during stepwise fractionation.

7.3.2. Mapping Lignin-Carbohydrate Interactions and Intramolecular Linkages

To verify the structural features of the isolated lignin and its interactions with polysaccharides, ^{13}C -enriched eucalyptus cell wall samples were analyzed using 2D dipolar-gated ^{13}C - ^{13}C correlation experiments under short- and long-range mixing conditions. Short-range correlations primarily represent intramolecular C-C connectivities within lignin and between lignin side-chain carbons, while long-range

correlations capture intermolecular interactions, including lignin-lignin and lignin-polysaccharide contacts. By comparing spectra obtained under different pretreatment durations, we dynamically tracked the remaining polysaccharide and lignin signals to identify structural motifs with strong correlation patterns. These observations enable the assignment of specific lignin-polysaccharide linkages and lignin-lignin crosslinks that are progressively disrupted or preserved during pretreatment. Notably, the ester carbonyl correlations near ~170 ppm gradually diminished, accompanied by the disappearance of the associated ferulic acid-derived lignin structures, indicating the cleavage of esterified linkages between lignin and polysaccharides.

In the short-range spectra (0.1s mixing), prominent intramolecular lignin-lignin correlations such as S2/6-S3/5 (~104-152 ppm) and G2-G6 (~111-119 ppm) were retained in the early pretreatment stages but weakened after prolonged treatment, suggesting progressive depolymerization of β -O-4-linked units. Side-chain correlations such as C α -C β (~72-85 ppm) and C β -C γ (~85-62 ppm) remained detectable in the early stages but diminished markedly after 4 h, reflecting cleavage of β -aryl ether linkages. In the long-range spectra (1.0 s mixing), several strong lignin-polysaccharide contacts were identified, including S2/6 (~104 ppm) with xylan C4 (~82 ppm) and G2 (~111 ppm) with xylan C2/C3 (~72-75 ppm), indicating close spatial proximity between aromatic rings and hemicellulose backbones. These cross peaks were substantially reduced after extended pretreatment, consistent with the solubilization of hemicelluloses and disruption of lignin-xylan associations. Notably, correlations involving ester carbonyls (~170 ppm) with xylan backbone carbons (~75-82 ppm) were almost completely absent after 4 h, in parallel with the loss of ferulate-derived aromatic carbons (~127-148 ppm), indicating the cleavage of ferulate-mediated crosslinks.

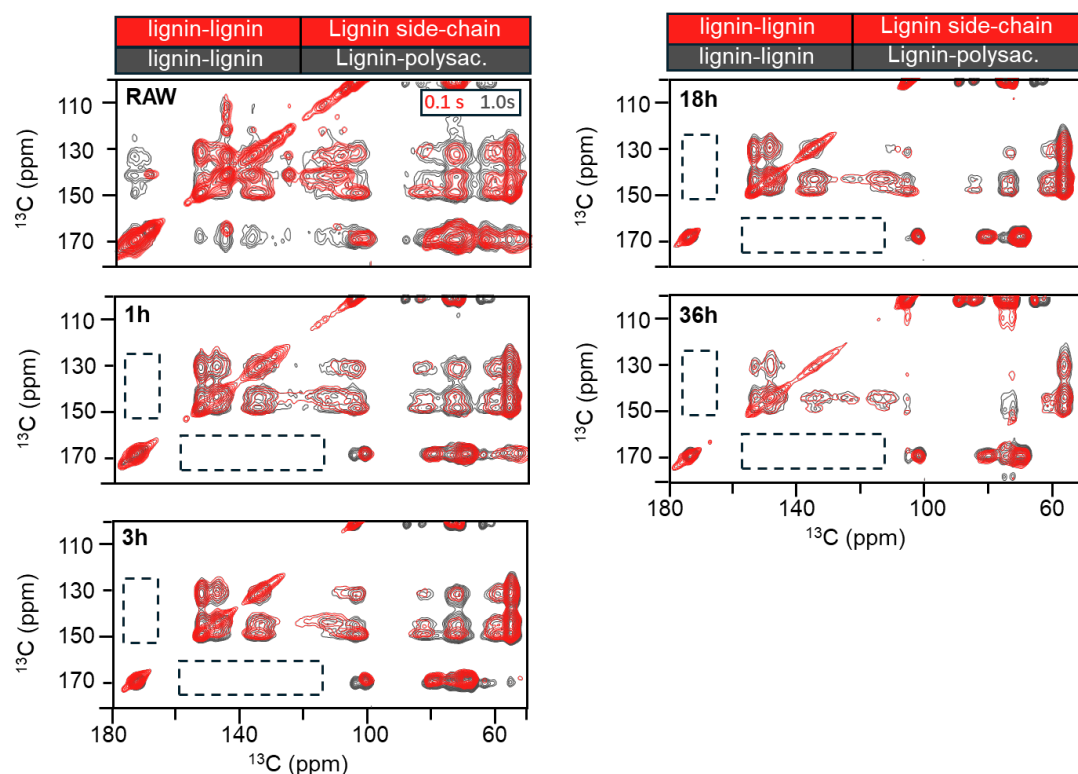


Fig.7-4. Overlay of short-range (0.1 s mixing; red) and long-range (1.0s mixing; grey) dipolar-gated ^{13}C - ^{13}C correlation spectra of ^{13}C -labeled eucalyptus cell walls under different pretreatment durations: untreated, 1h, 3h, 18h, and 36h. Short-range correlations primarily represent intramolecular C-C connectivity within lignin and between lignin side-chain carbons, while long-range correlations capture intermolecular interactions, including lignin-lignin and lignin-polysaccharide contacts. A pronounced loss of ester carbonyl correlations near ~ 170 ppm is observed with pretreatment, accompanied by the disappearance of associated ferulic acid-derived lignin structures.

To unambiguously assign solid-state NMR signals, all distinct intramolecular cross peaks in the dipolar-gated 2D ^{13}C - ^{13}C correlation spectra were annotated. Each cross peak represents an intramolecular contact between two carbons within the same molecule. Across pretreatment conditions, well-resolved patterns of lignin intramolecular correlations were observed. As pretreatment proceeded, the short- and long-mixing spectra became increasingly similar, indicating that the remaining lignin and polysaccharide domains adopt tighter and more homogeneous spatial arrangements, a feature likely contributing to the high recalcitrance of the residual biomass.

Intramolecular correlations within aromatic units were used to identify dominant lignin substructures in eucalyptus cell walls. The methoxyl carbon in lignin aromatic rings (S/G/FA) correlated with S3/5 at (56.5, 153) ppm and with G3 at (56.5, 148) ppm, enabling the mapping of carbon arrangements within the same monomer. Particularly, the FA9 carbon at 169 ppm, with its characteristic downfield shift, showed a short-range correlation with xylan C3 (Xn3), revealing a close spatial association between the ferulic acid moiety and the xylan backbone. These short-range correlations are critical for pinpointing side-chain connectivity in specific aromatic units.

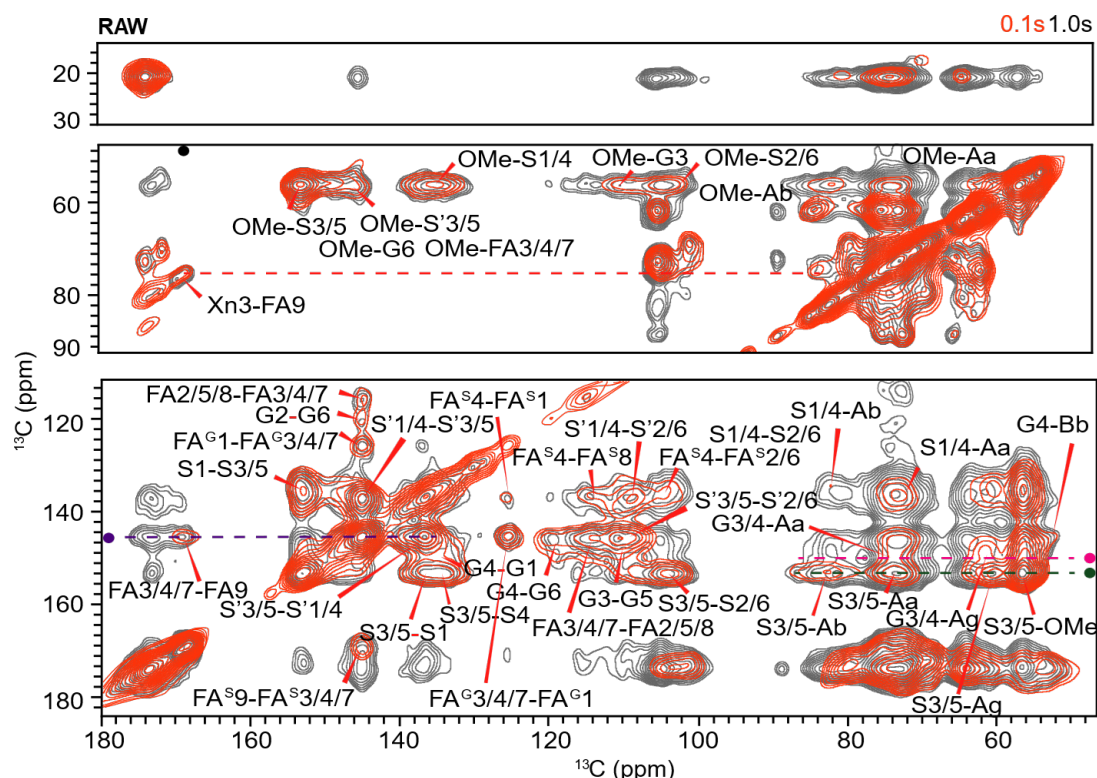


Fig.7-5. Representative dipolar-gated 2D ^{13}C - ^{13}C correlation spectra measured with short (0.1s, orange) and long (1.0s, black) mixing times for untreated eucalyptus cell walls. Compared with the short mixing time, the 1.0s spectrum reveals a greater number of long-range intermolecular cross peaks, which are not detected at 0.1s. Notably, the FA9 carbon at 169 ppm exhibits a characteristic chemical shift and shows a short-range correlation with Xn3, indicating a close spatial association between these two structural features. Only intramolecular cross peaks are labeled.

Long-range spectra further positioned these aromatic motifs within the lignin-

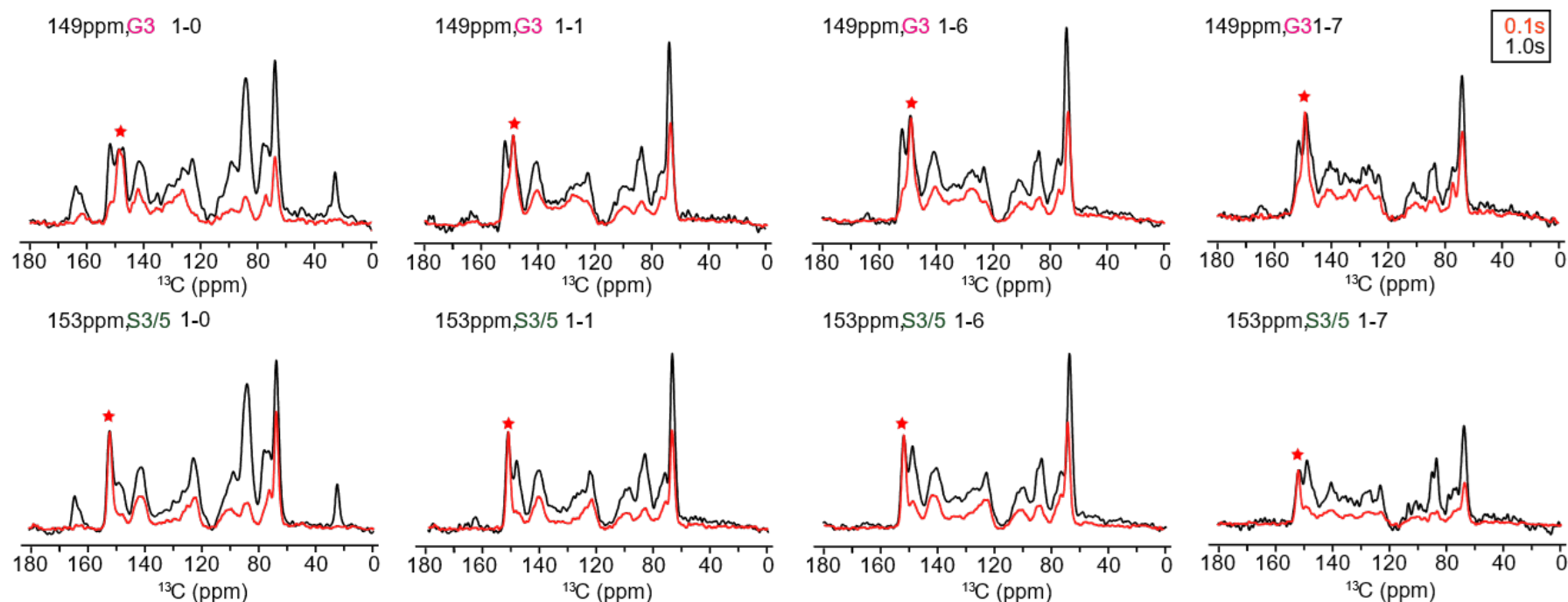
carbohydrate matrix. Characteristic correlations included (i) acetyl-lignin, (ii) xylan-lignin, (iii) acetyl-cellulose, (iv) acetyl-lignin methyl, (v) lignin methoxyl-cellulose, and (vi) between distinct lignin residues. Such interaction patterns place ferulate- and acetyl-bearing structures within the 3D lignin-carbohydrate network and highlight the sequential loss of ester-mediated and carbohydrate-associated contacts during pretreatment.

Fig.7-6. Representative dipolar-gated 2D ^{13}C - ^{13}C correlation spectra measured with short (0.1 s, orange) and long (1.0 s, black) mixing times for untreated eucalyptus cell walls. The long mixing time spectrum reveals more intermolecular cross peaks than the short mixing time spectrum, highlighting additional lignin-lignin and lignin-polysaccharide interactions. Notable examples include the G4-S2/6 and FA3/4/7-GlcA1 correlations. Only intermolecular cross peaks are labeled.

aromatic packing in the native wall structure. Together, these spectral features map the arrangement of lignin and carbohydrate domains in untreated biomass.

To better visualize these interactions, 1D slices were extracted from selected cross-peak positions in both short- and long-mixing-time spectra, enabling direct comparison of signal intensity and confirming the enhancement of intermolecular contacts at longer mixing times. From the 2D ^{13}C - ^{13}C correlation spectra slices, distinct and well-resolved cross peaks can be observed, enabling the separation of intramolecular and intermolecular contacts with high specificity. In both the single-slice and stacked-slice presentations, the 149 ppm (G3) and 153 ppm (S3/5) positions, corresponding to characteristic aromatic carbons in G- and S-type lignin display clear spectral features whose relative intensities vary with mixing time. Short mixing time spectra (0.1 s, orange) predominantly reflect intramolecular correlations within the aromatic units, whereas long mixing time spectra (1.0 s, black) capture both intra- and intermolecular contacts, revealing additional lignin-lignin and lignin-polysaccharide interactions. In the separated slices, peaks present exclusively or more strongly in the long-range spectra can be attributed to intermolecular associations. The stacked slice overlays, ordered from untreated samples at the bottom to the 36h pretreated sample at the top, show a progressive attenuation of long-range-only peaks as pretreatment severity increases. This diminishing gap between short- and long-range spectra indicates that the surviving lignin-carbohydrate and lignin-lignin contacts become more homogeneous and tightly packed, likely representing condensed, recalcitrant domains resistant to further cleavage. Such convergence in spectral patterns aligns with the hypothesis that extensive removal of labile lignin leaves behind a structurally uniform, highly crosslinked aromatic network. The use of star-marked reference peaks at G4 and S3/5 ensures accurate normalization and allows changes in cross-peak profiles to be confidently attributed to structural reorganization rather than spectral artifacts. Together, these 1D slice analyses deliver both spatial resolution-disentangling intra/intermolecular contacts and temporal resolution tracking the progressive remodeling of lignin domain architecture during stepwise fractionation. The resulting picture is one in which the accessible, ester-linked lignin-carbohydrate associations give way to a smaller set of tightly integrated ether- and C-C-linked motifs that persist through the most severe pretreatment conditions.

1



2

3 **Fig.7-7.** One-dimensional slices extracted at 149 ppm (top row, G3 of G-lignin) and 153 ppm (bottom row, S3/5 of S-lignin)
 4 from dipolar-gated 2D ^{13}C - ^{13}C correlation spectra of untreated eucalyptus (1-0), 1 h (1-1), 18 h (1-6), and 36 h (1-7) pretreated samples.
 5 Orange traces correspond to short mixing time (0.1 s) spectra, highlighting predominantly intramolecular correlations, while black traces
 6 correspond to long mixing time (1.0 s) spectra, which capture both intra- and intermolecular correlations. Comparison between short-
 7 and long-range slices allows identification of peaks arising from intermolecular contacts versus those from intramolecular connectivities
 8 within the same lignin unit. Peaks marked with a star indicate G4 or S3/5 aromatic positions, which were used as reference points for
 9 normalization across samples.

7.4. Perspective

This study provides a molecular-level description of the lignin-carbohydrate interface in ^{13}C -labeled Eucalyptus secondary cell walls, revealing lignin-carbohydrate covalent linkages are progressively disrupted under controlled fractionation. Quantitative DP and multidimensional ^{13}C - ^{13}C spectroscopy using in this study, overlapping resonances were disentangled, ester- and ether-linked lignin-carbohydrate complexes were identified, and spatial proximities between aromatic units and polysaccharide backbones were mapped. These findings clarify the heterogeneous distribution and connectivity of guaiacyl and syringyl lignin in woody biomass and highlight the structural role of ferulate- and glucuronate-mediated linkages in constructing secondary wall architecture. Moreover, the methodological framework established here can be extended to other lignocellulosic feedstocks to evaluate species-specific wall architectures and their implications for industrial processing. Understanding such structural bases of cell wall recalcitrance provides inspiration for bioinspired strategies. Importantly, the principles uncovered, such as the persistence of condensed ether-linked aromatic networks and the spatial organization of residual lignin domains offer guidance for engineering approaches aimed at complete lignin removal or enhanced deconstruction efficiency. Future integration of in situ isotopic labeling with advanced sensitivity-enhancement techniques, including dynamic nuclear polarization (DNP), will enable probing of low-abundance interfacial motifs and monitoring of structural dynamics under real processing conditions. Together, these advances accelerate the translation of fundamental knowledge of lignocellulose architecture into scalable strategies for sustainable bioenergy and bioproduct manufacturing.

CHAPTER 8. Conclusion

Understanding the molecular basis of lignocellulosic biomass recalcitrance remains a central challenge in the development of efficient and selective biorefinery processes. This thesis addresses this challenge by systematically linking feedstock compositional diversity to the molecular architecture of secondary plant cell walls. Comparative analysis demonstrates that fractionation resistance is governed primarily by interactions among cellulose, hemicellulose, and lignin rather than bulk composition alone. From an application perspective, hardwoods enriched in syringyl units and β -O-4 linkages exhibit clear advantages for lignin-first biorefineries, whereas C-lignin-rich seed coats are better suited for processes targeting uniform aromatic structures. In contrast, softwoods and seed shells dominated by guaiacyl lignin are less favorable for lignin depolymerization but more compatible with sugar-platform biorefineries, while bark rich in suberin requires alternative pretreatment strategies.

To optimize pretreatment conditions, a dynamic predictive model (Pretreatment Depolymerization Factor) was established. The PDF model describes the coupled kinetics of hemicellulose hydrolysis and lignin dissociation. It illustrates how xylan conformation affects lignin extraction and incorporates lignin condensation and reprecipitation into the framework. A previously unknown $C\beta$ - $C\gamma$ cleavage pathway initiated by γ -OH activation was identified. The final integrated kinetic model enables a transition from empirical pretreatment adjustment to mechanism-based control. Application of the refined model to semi-continuous pretreatment achieved 95% delignification with an approximately 30% improvement in monolignol yield. Moreover, Validation across different biomass types highlights the general applicability of this framework

At the molecular scale, the lignin-carbohydrate interface in ^{13}C -labeled Eucalyptus secondary walls was resolved. Solid-state NMR identified ester- and ether-linked complexes and mapped the spatial relationships between aromatic units and polysaccharide backbones. The results revealed heterogeneous distributions of guaiacyl

and syringyl lignin's. Ferulate- and glucuronate-mediated linkages were shown to play key structural roles in wall organization. The persistence of guaiacyl structure was identified as a fundamental origin of recalcitrance. This analytical framework can be also extended to other feedstocks to assess species-specific wall architectures.

Overall, this thesis advances molecular-level understanding of lignocellulosic biomass deconstruction and provides a rational basis for the design of severity-controlled and whole biomass utilization-first biorefinery processes.

Appendix

Table S1-1: Chemical composition, β -O-4 / benzodioxane contents, and RCF monomer yield of different LCB types. Data for the Graphical abstract are referred from this table.

Biomass	Species	Glu ^a	Xyl ^a	Ara ^a	Gal ^a	Man ^a	KL ^a	β -O-4 / Benzo- dioxane ^h	Theoretical monomer yield ^k	RCF monomer yield ⁱ	Reference
Hardwood	Maple	41.8	18.4	0.3	0.3	2.0	24.9	55.6	77.0	27.0	(Pepper & Hibbert, 1948; Zhang et al., 2013)
	Aspen	44.0	19.0	0.5	0.9	2.4	19.0	63.9	77.6	50.0	(Li et al., 2018a; Pepper & Fleming, 1978; Shao et al., 2020; Wang et al., 2018)
	Birch	39.0	24.3	0.5	0.4	2.5	20.8	63.5	83.9	49.0	(Wang et al., 2018)
	Poplar	45.7	14.6	0.0	0.0	4.6	22.8	64.04	93.5	42.0	(Chen et al., 2017a; Chen et al., 2017b; Wang et al., 2019b)
	Eucalyptus	44.7	16.0	0.4	1.4	1.1	22.9	59.6	81.3	44.0	(Guo et al., 2018; Li et al., 2018a; Renders et al., 2018)
	Beech	37.1	15.3	2.1	2.2	1.9	22.8	62	133.1	48.3	(Shuai et al., 2016a)
	F5H-poplar ^d	42.9	18.1	0.0	0.0	0.0	14.2	88.8	112.0	76.7	(Shuai et al., 2016a)
	Black locust	36.2	13.4	0.8	0.7	0.2	23.0	N/A	N/A	34.8	(Vangeel et al., 2019a)
Summary		37.1-45.7	14.6-24.3	0.0-0.5	0.0-1.43	1.13-4.6	14.2-24.9	55.6-65.0	77-131	27.0-50.0	
Average	Hardwood	41.4	17.4	0.6	0.7	1.8	21.3	65.3	94.1	46.5	
Softwood	Pine	48.4	6.2	0.0	1.8	12.1	24.7	49.16	24.0	22.0	(Van Aelst et al., 2020)
	Spruce	42.0	5.5	1.1	1.6	12.3	28.2	38.1	40.9	21.0	(Wang et al., 2018)
	Cedar	43.5	6.1	0.0	4.1	7.5	36.6	N/A	N/A	10.0	(Wang et al., 2020a; Yamashita et al., 2010)
	Summary	42.0-48.4	5.5-6.2	0.0-1.1	1.6-4.13	7.49-12.2	24.7-36.6	38.1-49.2	24-41	10.0-22.0	
Average	Softwood	44.6	5.9	0.4	2.5	10.5	29.8	43.6	32.5	17.7	
Grass	Miscanthus	49.0	21.2	3.1	1.7	2.4	21.4	63.0	84.9	26.4	(Warren-Walker et al., 2024)
	Bamboo	41.9	22.9	0.0	0.0	0.0	27.1	56.0	85.0	19.1	(Jiang et al., 2016)
	Corn stalk	42.2	22.3 ^g				19.5	51.0	50.7	24.0	(Sun et al., 2022)
	Sorghum	31.0	20.5	0.0	0.0	0.0	23.7	N/A	N/A	N/A	(Joy & Krishnan, 2022)
	Wheat straw	35.7	18.0	2.8	0.6	0.0	19.8	43	117.7	25.0	(Brienza et al., 2022; Zeng et al., 2013; Zheng et al., 2018)
	Summary	31.0-52.5	18.0-22.9	0.0-3.1	0.0-1.7	0.0-2.4	18.5-27.1	51.0-77.1	50.7-118	19.1-26.4	
Average	Grass	40.0	21.0	1.5	0.6	0.6	22.3	53.3	84.6	23.6	

Barks	Black locust	14.0	2.5	4.3	2.0	1.0	36.0	N/A	N/A	9.1	(Vangeel et al., 2019a)
	Cedar-outer	58.50e					33.1	N/A	N/A	N/A	(Usia & Kara, 1997)
	Endemic oak	20.27f	10.91f	1.83f	0.71f	Tracef	30.8	N/A	N/A	N/A	(Balaban Ucar & Uçar, 2001)
	Norway spruce	38.5	3.8	6.6	3.6	Trace	40.0	32	41.0	8.1	(Samec et al., 2024)
	Summary	14.0-58.5	2.5-10.9	4.3-1.8	0.7-3.6	Trace	30.8-40.0	32	41	8.1-9.1	
Average	Barks	32.8	5.7	4.2	2.1	1.0	35.0	32.0	41.0	8.6	
Seed coats ^b (C-lig)	Candlenut	15.0	10.5 ^g	N/A	N/A	N/A	48.0	N/A	N/A	N/A	(Khan et al., 2024a)
	Candlenut	19.8	15.2 ^g	N/A	N/A	N/A	51.0	63 ^l	202.4	N/A	(Yin et al., 2024)
	Vanilla	12.1	1.8	1.9	1.7	0.2	61.0/13 ^m	100 ^m	130.0	N/A	(Stone et al., 2018)
	Vanilla	14.2	0.3	0.4	1.3	Trace	65.4/13 ^m	100 ^m	130.0	84 ^j	(Li et al., 2018c)
	Castor	15.6	9.5	0.0	0.0	0.0	59.3/12.9 ^m	86 ^m	95.4	82 ^j	(Wang et al., 2021)
	Castor	13.2	9.43 ^g	N/A	N/A	N/A	60.6/12.9 ^m	99 ^m	126.4	N/A	(Xia et al., 2025)
	Chinese tallow	16.4	10.6 ^g	N/A	N/A	N/A	62.0/15.4 ^m	96 ^m	141.9	83 ^j	(Su et al., 2023a)
	Jatropha	N/A	N/A	N/A	N/A	N/A	45.5	92 ^m	/	71 ^j	(Su et al., 2021)
	Summary	12.1-19.6	0.3-10.6	0.0-1.9	0.0-1.7	0.0-0.2	45.5-65.4	70-100	95-202	71-84 ^j	
Average	Seed coats ^b	15.2	8.2	0.8	1.0	0.1	56.6	90.9	137.7	15.2	
Seed shell	Rice husks	34.4	15.6	1.8	1.9	0.3	19.4	N/A	N/A	33.0	(Ahn et al., 2024)
	Cameilla	13.7	34.3	0.0	0.0	0.0	30.2	55.4	114.2	25.3	(Cheng et al., 2023)
	Walnut	20.0	14.9	0.0	1.8	0.0	30.0	40.1	48.2	25.8	(Khan et al., 2024a)
	Peanut	25.2	8.4	0.5	1.5	0.0	28.0	58.1	94.5	9.8	(Khan et al., 2024a)
	Pistachio	25.0	26.5	0.0	0.0	0.0	18.0	61.4	67.9	32.9	(Khan et al., 2024a)
	ginkgo	21.3	19.9	0.0	1.9	0.0	42.4	40	67.8	16	(Jiang et al., 2018)
	Cherry	19.7	14.9	0.4	0.9	0.0	31.0	46.0	65.6	38.0	(Khan et al., 2024a)
	Peach	18.1	15.9	0.0	1.7	0.0	38.0	39.1	58.1	15.9	(Khan et al., 2024a)
	Summary	13.7-34.4	8.4-34.3	0.0-1.8	0.0-1.9	0.0-0.3	18.0-38.0	39.1-61.4		9.8-38.0	
Average	Seed shell	22.2	18.8	0.6	1.1	0.0	29.6	48.6	73.8	24.6	

^a All values are in weight% of biomass dry weight. Glu = glucose; Xyl = xylose; Ara = arabinose; Gal = galactose; Man = mannose; KL = Klason lignin. ^b Seed coats which contain natural C-lignin ^c Seed coats which contain no natural C-lignin ^d F5H-poplar = genetically engineered poplar to have high S-lignin units ^e Holocellulose, which consists of both cellulose and hemicellulose ^f Measured after ethanol/benzene extraction and 1% NaOH extraction ^g Hemicellulose amount ^h

Amount of β -O-4 after lignin isolation; and amount of benzodioxane after C-lignin isolation (for seed coats with C-lignin). The values of benzodioxane are italicized. ⁱ RCF monomer yield = weight% based on mass of lignin input into RCF reactor (RCF reaction)

^j CTH monomer yield = mole% based on caffeyl alcohol in isolated C-lignin input into CTH reactor (PT-CTH reaction) ^k Theoretical monomer yield (mg/g) = KL weight% * (β -O-4 content/Benzo-dioxane)² ^l the total β -O-4/ benzo-dioxane of total Aromatic, Candlenut including G/S and C type lignin ^m C-type lignin mass fraction and benzo-dioxane content in C-type aromatics. Other KL in seed coats contains less β -O-4. Theoretical monomer yield only consider C-lignin part

Table S3-1. Parameters and efficiency of different organosolv pretreatments (PT), and monomer yield of the subsequent hydrogenolysis (CTH)

Biomass	Species	Solvent	SL ^a	Temp (°C)	Time (h)	Acid	Composition ^a			Lignin ^d				Ref.		
						Conc.	Type	CR	HY	LY	β-O-4 Ar% ^d	β-β Ar%	β-5 Ar%		Monomer yield	
Softwood	Acacia Confusa	Pentanol	14.29%	170	1	45 mM	H ₂ SO ₄	89.46%	81.18%	65%	26.2	10.4	3.2	6.86% ^c	(Islam et al., 2021a)	
	Acacia Confusa	Ethanol	14.29%	170	1	45 mM	H ₂ SO ₄	85.41%	84.71%	71.43%	7.20	9.10	0	0.52% ^c	(Islam et al., 2021a)	
	Beech	Dioxane	10.00%	80	5	100 mM	H ₂ SO ₄			76.85%	26.65 ^c			7.10%	(Shuai et al., 2016a)	
	Populus	Ethanol	5.00%	180	1	1w% ^b	H ₂ SO ₄				4.80	9.50	0	0.23% ^c	(Yang et al., 2020)	
	Birch	MA ^a	6.67%	120	0.5					60.57%	26.70	8.90	0	7.13% ^c	(Cai et al., 2023)	
	Acacia Confusa	NMP ^a	14.29%	170	1	50 mM	H ₂ SO ₄	83.09%	69.90%	74.69%	16.36			2.68% ^c	(Islam et al., 2024)	
	Acacia Confusa	Ethanol	14.29%	170	1	50 mM	H ₂ SO ₄	87.44%	89.32%	74.07%	11.40			1.30% ^c	(Islam et al., 2024)	
	Populus	Ethanol	8.00%	180	0.33	1.25w% ^b	H ₂ SO ₄			74.00%	1.40	0.80	0	0.02% ^c	(Yoo et al., 2017)	
	Eucalyptus	Ethanol	14.29%	170	1	20 mM	H ₂ SO ₄	86.40%	98.30%	82.00%					(Dong et al., 2019b)	
	Hardwood	Eucalyptus	1,4-BDO ^a	14.29%	170	1	20 mM	H ₂ SO ₄	86.90%	97.90%	92.00%					(Dong et al., 2019b)
		Eucalyptus	EG ^a	14.29%	120	1	200 mM	H ₂ SO ₄	92.40%	89.03%	90.11%	23.40	9.60	0	7%	(Dong et al., 2022)
		Poplar	<i>p</i> -TsOH ^a	5.00%	90	1	2.5 M	<i>p</i> -TsOH	82.70%	85.90%	70.81%	10.00	3.10	0	1.00% ^c	(Wang et al., 2019a)
		Birch	Dioxane	10.00%	80	5	100mM	HCL			86.90%	36.88 ^c			13.60%	(Lan et al., 2018a)
		Oak	Methanol	5.00%	160	2	0.05 mmol/g ^b	H ₂ SO ₄	95%	73%	70%	42.19 ^c			17.80%	(Ouyang et al., 2018)
		Oak	Methanol	5.00%	160	2	0.05 mmol/g ^b	AL(OTf) ₃	79%	72%	61%	43.01 ^c			18.50%	(Ouyang et al., 2018)
	Oak	Methanol	5.00%	160	2	0.05 mmol/g ^b	<i>p</i> -TsOH	76%	54%	64%	51.19 ^c			26.20%	(Ouyang et al., 2018)	
	Oak	Methanol	5.00%	160	2	0.05 mmol/g ^b	HCL	74%	37%	47%	52.35 ^c			27.40%	(Ouyang et al., 2018)	
	Oak	Methanol	5.00%	160	2	0.05 mmol/g ^b	H ₃ PO ₄	89%	17%	44%	50.00 ^c			25.00%	(Ouyang et al., 2018)	
	Hybrid poplar	Ethanol	12.50%	160	1	0.05w% ^b	H ₂ SO ₄			70.30%	38.73 ^c			15%	(Bär et al., 2018)	
Summary				90-170	0.33-5			83-95%	17-98%	47-92%	1.4-52	0.8-10	0-3.2	0.2-27.4%		
Average	Hardwood		9.59%	149	1.6			85%	73%	71%	28	7.3	0.5	10.4%		

Table S3-2. Parameters and efficiency of different organosolv pretreatments (PT), and monomer yield of the subsequent hydrogenolysis (CTH)

Biomass	Species	Solvent	SL ^a	Temp (°C)	Time (h)	Acid		Composition ^a			Lignin ^d				Ref.
						Conc.	Type	CR	HY	LY	β-O-4 Ar%	β-β Ar%	β-5 Ar%	Monomer yield	
Softwood	Masson pine	GVL ^a	10.00%	170	1	25mM	H ₂ SO ₄	82.98%	96.96%	92.33%	3.50	8.27	8.88	0.12% ^c	(Huang et al., 2022)
	Masson pine	THF ^a	10.00%	170	1	25mM	H ₂ SO ₄	87.34%	96.86%	69.67%	9.36	6.15	10.21	0.88% ^c	(Xu et al., 2023)
	Japanese cypress	Ethanol	18.76%	170	0.75	0.4w% ^b	HCL	65.00%	~100%	67.00%					(Hiden et al., 2013)
	Norway spruce	Ethanol	25.00%	200	0.5	1w% ^b	H ₂ SO ₄	97.64%	49.25%	71.71%					(Matsakas et al., 2019)
	Norway spruce	Ethanol	25.00%	200	1	1w% ^b	H ₂ SO ₄	94.51%	70.90%	61.94%					(Matsakas et al., 2019)
	Loblolly pine	Ethanol	14.29%	170	1	1w% ^b	H ₂ SO ₄	79.30%	96.10%	61.20%					(Sannigrahi et al., 2010)
	Lodgepole pine	Ethanol	10.00%	170	1	1.1w% ^b	H ₂ SO ₄	71.50%	89.00%	73.80%					(Wang et al., 2010)
	Masson pine	Ethanol	10.00%	170	1	25mM	H ₂ SO ₄	89.36%	73.91%	53.33%	19.41	7.65	10.63	3.77% ^c	(Xu et al., 2023)
	Masson pine	2-MeTHF ^a	10.00%	170	1	25mM	H ₂ SO ₄	74.47%	96.52%	26.67%	26.41	11.20	11.23	6.97% ^c	(Xu et al., 2023)
Summary				170-200	0.5-1			74-98%	70-100%	26-92%	3.5-26.4	6.2-11.2	8.9-11.2	0.1-7%	
Average		Softwood	14.78%	177	0.92			82%	84%	64%	15	8.3	10	3.68%	
Grass	Corn stover	Ethanol	12.50%	160	0.75	1.2w% ^b	H ₂ SO ₄			68.20%	13.57	4.22	0	1.84% ^c	(Xu et al., 2022)
	Corn stover	Ethylene-diamine	12.50%	120	0.75	0.8ml/g	Ethylene-diamine	93.18%	35.30%	66.81%	20.86	2.93	0	4.35% ^c	(Xu et al., 2022)
	Barley straw	Ethanol	14.29%	170	1	1.6w% ^b	FeCl ₃	83.00%		76.00%					(Kim et al., 2010)
	Miscanthus	Ethanol	12.50%	170	1	0.5w% ^b	H ₂ SO ₄	95.00%	88.20%	70.30%					(Obama et al., 2012)
	Miscanthus	FA+AA ^a	4.00%	107	3			85.10%	78.80%	79.60%					(Vanderghem et al., 2012)
	Corn stover	FA ^a	7.69%	80	3			89.50%	79.80%	65.60%					(Yu et al., 2013)
	Wheat straw	Ethanol	10.00%	190	1	0.29w% ^b	H ₂ SO ₄	91.10%	81.40%	59.10%					(Wildschut et al., 2013)
	Napier grass	Ethanol	10.00%	190	1	1w% ^b	H ₂ SO ₄	83.08%	49.00%	62.10%					(Jomnonkhaow et al., 2022)
	Bamboo	BDO	10.00%	140	1	1w% ^b	H ₂ SO ₄	91.24%	82.50%	48.64%	48.64	18.16	1.40	3.30% ^c	(Huang et al., 2022)
Summary				80-190	0.75-3			83-95%	35-88%	49-80%	13.6-49	3-18.1	0-1.4	1.84-3.3%	
Average		Grass	10.39%	147	1.4			89%	71%	66%	28	8.4	0.5	3.2%	

Table S3-3. Parameters and efficiency of different organosolv pretreatments (PT), and monomer yield of the subsequent hydrogenolysis (CTH)

Biomass	Species	Solvent	SL ^a	Temp (°C)	Time (h)	Acid		Composition ^a			Lignin ^d			

						Conc.	Type	CR	HY	LY	β -O-4 Ar%	β - β Ar%	β -5 Ar%	Monomer yield	Ref.
Bark	Pine	Ethanol	7.69%	165	1	1w% ^b	H ₂ SO ₄	67.2% carbohydrate		24.80%	11.00	2.00	5.00	1.21% ^c	(Liu et al., 2020)
	Oak	Ethanol	3.85%	180	1	1w% ^b	H ₂ SO ₄	75.6% carbohydrate		55.77%	12.00	3.00	3.00	1.44% ^c	(Liu et al., 2020)
	Norway Spruce	Ethanol	3.85%	160	1.25 ^f	1.5w% ^b	H ₂ SO ₄	99%	97.90%	63%	37.00			9.00% ^c	(Rietzler et al., 2022)
	Eucalyptus	Ethanol	12.50%	200	1			89.04%	69.03%	49.45%	50.55				(Romaní et al., 2019)
	Spruce	Ethanol	10.00%	160	1	1w% ^b	H ₂ SO ₄	54.00%	69.50%	40.60%	59.42				(Hrůzová et al., 2021)
	Spruce	Ethanol	10.00%	200	1	1w% ^b	H ₂ SO ₄	39.70%	92.30%	53.91%	46.12				(Hrůzová et al., 2021)
	Aspen	Ethanol	10.00%	80	5	1.65ml/100ml	HCL	91.4% ^g		23.52% ^g					(Pals et al., 2022)
	Aspen	Butanol	10.00%	80	5	1.65ml/100ml	HCL	104.7% ^g	95.70% ^g	57.93% ^g					(Pals et al., 2022)
	Aspen	Dioxane	10.00%	80	5	1.65ml/100ml	HCL	107.9% ^g	77.80% ^g	65.53% ^g					(Pals et al., 2022)
Summary				80-200	1-5			40-108%	4.3-98%	24-66%	11-59	2-3	3-5	1.2-9.0%	
Average	Bark		8.65%	145	2.4			84% ^g	84% ^g	48.27%	36.0	2.5	4.0	3.9%	
Seed	Castor	Ethanol	10.00%	190	4	1.2%w	H ₂ SO ₄			16.80%				3.21% ^c	(Wang et al., 2020b)
	Castor	Dioxane	10.00%	85	3	pH=2	HCL							1.92% ^e	(Wang et al., 2021)
	Jatropha	Dioxane	10.00%	85	3									3.61% ^e	(Su et al., 2021)
	Candlenut	2-MeTHF	5.00%	160	2	1mmol/g	HCL	54.03%	45.30%	21.00%				10.52% ^e	(Yin et al., 2024)
	Walnut	EtOH	10.00%	120	6	5–10%v HCl	HCL			55%					
	Cameilla	AA	10.00%	90	2	0.1v%	HCL	83.00%	83.16%	77.5%	33.42			11.22% ^e	(Cheng et al., 2023)
	Walnut	EtOH	10.00%	85	5	0.24M	HCL			11%	36.00				
	Pistachios	EtOH	10.00%	85	5	0.24M	HCL			23.5%	31.00				
Summary				85-190	2-4			54%	45%	17-21%				2-10.5% ^c	
Average	Seed		9.38%	112.5	3.8			68.5%	64.2%	34.1%	33.47			6.08%	
Protection way pretreat-ment	Beech	Dioxane	10.00%	80	5	100	H ₂ SO ₄	89.46%	88%	78.96%	67.13 ^c			45.06%	(Shuai et al., 2016a)
	Birch	MA	6.67%	120	0.25	60%		88%		55.45%	53.20	13.60	0.80	28.30%	(Cai et al., 2023)
	Poplar	<i>p</i> -TsOH	5.00%	90	1	2.5M	<i>p</i> -TsOH		79.85%	65.69%	47.60	4.50	2.50	22.66%	(Wang et al., 2019a)
	Birch	Dioxane	10.00%	80	5	100	HCL	93.50%	80.00%	89.00%	64.50 ^c			41.60%	(Lan et al., 2018c)

Summary	80-120	0.25-5	88-94%	80-88%	55-89%	48-67.1	4.5-14	0.8-2.5	23-45%
---------	--------	--------	--------	--------	--------	---------	--------	---------	--------

^a SL=solid loading, Values are in unit of g-initial-dry-weight-of-biomass/volume of solvent;

CR = cellulose remain; HY = hemicellulose separation yield; LY = lignin yield. **All values are calculated based on the initial composition of biomass.**

GVL= γ -Valerolactone; 2-MeTHF =2-Methyltetrahydrofuran; THF = Tetrahydrofuran; FA=Formic acid; AA=Acetic acid;

MA=Maleic acid; NMP = 1-Methyl-2-pyrrolidinone; 1-4BDO= 1,4-Butanediol; EG=Ethylene glycol; *p*-TsOH=*p*-Toluenesulfonic; AL(OTf)₃=acid aluminum triflate;

^b Values are based on total dry weight of the biomass

^c The relationship between β -O-4 content and monomer yield was estimated based on the study by T. Phongpreecha(Phongpreecha et al., 2017) in 2017, following the equation: **Monomer yield = (β -O-4 content)²**

^d The internal linkages of aromatic units were quantified based on the number of aromatic rings

^e Monomer yield calculated based on the total dry weight of the biomass.

^f Extraction conducted at 100 bars for 15 cycles, with each cycle lasting 5 minutes.

^g The NREL compositional analysis method may introduce errors when applied to biomass rich in extractives, such as bark.

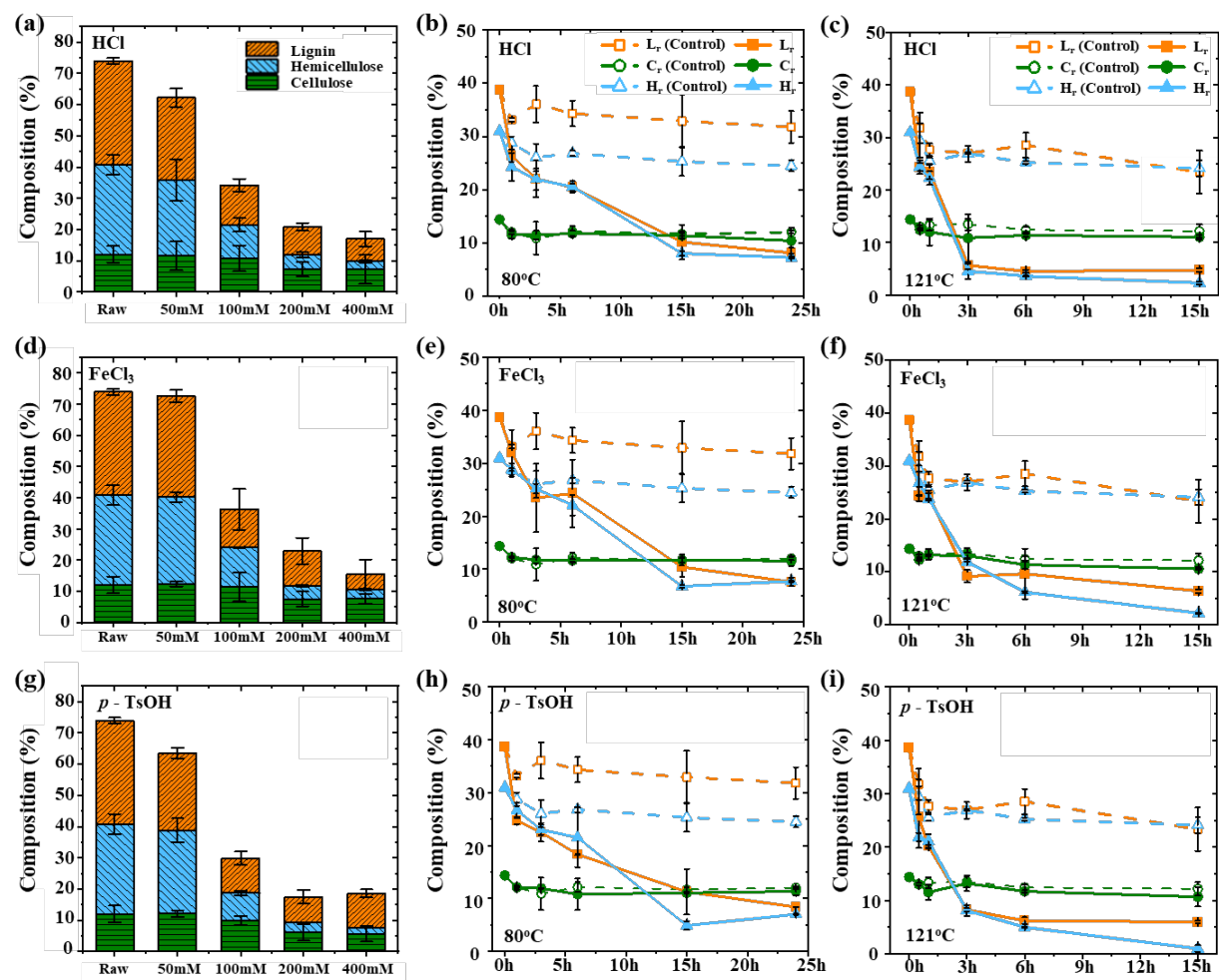


Fig S5-1. Changes in chemical compositions of the lignocellulosic biomass during pretreatment. The experiments evaluated the effects of increased acid doses (left column) of three different catalysts, i.e., HCl, FeCl₃, and *p*-TsOH (top to bottom rows); and the dynamic changes of the three building block components at 80 and 121°C (middle and right columns, respectively).

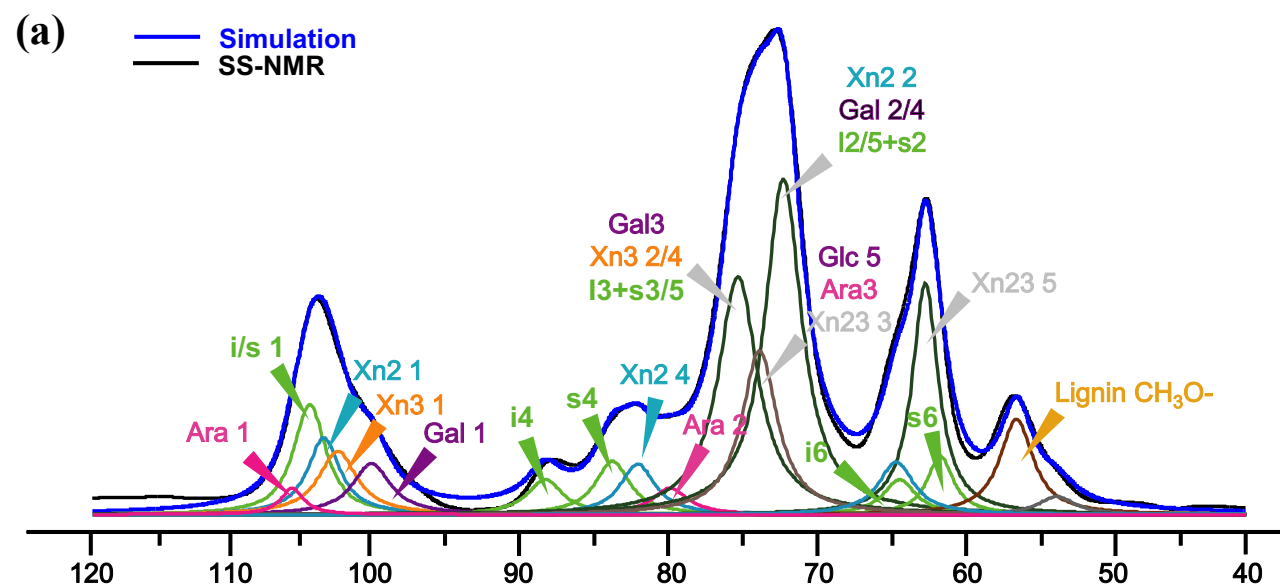


Fig S5-2. Spectral deconvolution of quantitative ^{13}C spectra for compositional analysis of COS polysaccharides. The experimentally measured spectra (black line) are overlaid with the simulated spectra (blue line), with the deconvoluted peaks underlying the spectral envelope;

Table S5-2: Experimental conditions and results of organosolv-acid pretreatments of *Camellia oleifera* shell (OTS). Parameters include solvent system (BDO), solvent-to-lignin ratio (S:L), temperature, reaction time, and acid type/concentration. Results are reported as Pretreatment Depolymerization Factor (PDF), hemicellulose hydrolysis ratio (HR), lignin removal (LR), and cellulose retention (CR).

Entry	Feedstock	Solvent ^c	S:L ^c	Temp(°C)	Time(h)	Acid		PDF	HR	LR	CR
						Type	Conc.				
P-1	OTS	BDO	1:10	121	1	HCl	0mM	0.0	73.24%	80.34%	83.70%
P-2	OTS	BDO	1:10	121	1	HCl	50mM	0.8	61.81%	80.69%	80.09%
P-3	OTS	BDO	1:10	121	1	HCl	100mM	1.8	38.19%	32.82%	75.22%
P-4	OTS	BDO	1:10	121	1	HCl	200mM	4.0	11.74%	23.19%	70.30%
P-5	OTS	BDO	1:10	121	1	HCl	400mM	10.3	6.61%	18.67%	62.92%
P-6	OTS	BDO	1:10	80	1	HCl	50mM	0.2	73.81%	71.15%	85.13%
P-7	OTS	BDO	1:10	80	3	HCl	50mM	0.6	70.46%	62.06%	84.40%
P-8	OTS	BDO	1:10	80	6	HCl	50mM	1.2	65.27%	58.04%	85.86%
P-9	OTS	BDO	1:10	80	15	HCl	50mM	2.9	27.15%	25.95%	81.49%
P-10	OTS	BDO	1:10	80	24	HCl	50mM	4.7	26.23%	23.81%	80.76%
P-11	OTS	BDO	1:10	121	0.5	HCl	50mM	0.4	75.34%	66.07%	97.50%
P-12	OTS	BDO	1:10	121	1	HCl	50mM	0.8	69.85%	62.06%	97.50%
P-13	OTS	BDO	1:10	121	3	HCl	50mM	2.5	14.95%	13.91%	82.95%
P-14	OTS	BDO	1:10	121	6	HCl	50mM	5.0	14.34%	13.64%	88.04%
P-15	OTS	BDO	1:10	121	15	HCl	50mM	12.5	6.41%	13.64%	89.49%
P-16	OTS	BDO	1:10	121	1	HCl	500mM	15.0	4.88%	24.77%	60.22%
P-17	OTS	BDO	1:10	165	3	HCl	100mM	18.1	3.72%	28.92%	57.37%
P-18	OTS	BDO	1:10	121	1	FeCl ₃	0mM	0.0	73.24%	80.34%	83.70%
P-19	OTS	BDO	1:10	121	1	FeCl ₃	50mM	0.8	61.81%	63.90%	67.77%
P-20	OTS	BDO	1:10	121	1	FeCl ₃	100mM	1.8	22.35%	28.79%	69.22%
P-21	OTS	BDO	1:10	121	1	FeCl ₃	200mM	4.0	7.92%	21.28%	42.75%
P-22	OTS	BDO	1:10	121	1	FeCl ₃	400mM	10.3	5.27%	28.15%	39.21%
P-23	OTS	BDO	1:10	80	1	FeCl ₃	50mM	0.2	80.22%	64.46%	97.50%
P-24	OTS	BDO	1:10	80	3	FeCl ₃	50mM	0.6	70.46%	62.06%	92.40%
P-25	OTS	BDO	1:10	80	6	FeCl ₃	50mM	1.2	66.49%	50.29%	81.49%

P-26	OTS	BDO	1:10	80	15	FeCl ₃	50mM	2.9	17.08%	33.97%	77.85%
P-27	OTS	BDO	1:10	80	24	FeCl ₃	50mM	4.7	18.61%	23.81%	81.49%
P-28	OTS	BDO	1:10	121	0.5	FeCl ₃	50mM	0.4	68.02%	62.06%	96.04%
P-29	OTS	BDO	1:10	121	1	FeCl ₃	50mM	0.8	66.80%	55.37%	92.40%
P-30	OTS	BDO	1:10	121	3	FeCl ₃	50mM	2.5	23.79%	23.27%	95.31%
P-31	OTS	BDO	1:10	121	6	FeCl ₃	50mM	5.0	16.47%	19.26%	88.77%
P-32	OTS	BDO	1:10	121	15	FeCl ₃	50mM	12.5	6.41%	18.99%	80.76%
P-33	OTS	BDO	1:10	165	3	FeCl ₃	100mM	18.1	5.77%	24.47%	47.32%
P-34	OTS	BDO	1:10	121	1	<i>p</i> -TsOH	0mM	0.0	88.76%	80.34%	83.70%
P-35	OTS	BDO	1:10	121	1	<i>p</i> -TsOH	50mM	0.8	39.65%	58.85%	81.49%
P-36	OTS	BDO	1:10	121	1	<i>p</i> -TsOH	100mM	1.8	39.65%	31.30%	77.85%
P-37	OTS	BDO	1:10	121	1	<i>p</i> -TsOH	200mM	4.0	15.25%	8.56%	64.76%
P-38	OTS	BDO	1:10	121	1	<i>p</i> -TsOH	400mM	10.3	5.27%	6.95%	62.57%
P-39	OTS	BDO	1:10	80	1	<i>p</i> -TsOH	50mM	0.2	89.06%	88.54%	88.04%
P-40	OTS	BDO	1:10	80	3	<i>p</i> -TsOH	50mM	0.6	76.86%	64.46%	82.95%
P-41	OTS	BDO	1:10	80	6	<i>p</i> -TsOH	50mM	1.2	67.71%	65.80%	82.22%
P-42	OTS	BDO	1:10	80	15	<i>p</i> -TsOH	50mM	2.9	21.66%	27.82%	85.13%
P-43	OTS	BDO	1:10	80	24	<i>p</i> -TsOH	50mM	4.7	25.01%	23.27%	85.86%
P-44	OTS	BDO	1:10	121	0.5	<i>p</i> -TsOH	50mM	0.4	80.52%	66.07%	92.40%
P-45	OTS	BDO	1:10	121	1	<i>p</i> -TsOH	50mM	0.8	78.39%	65.80%	95.31%
P-46	OTS	BDO	1:10	121	3	<i>p</i> -TsOH	50mM	2.5	21.96%	24.34%	93.13%
P-47	OTS	BDO	1:10	121	6	<i>p</i> -TsOH	50mM	5.0	21.96%	27.02%	88.04%
P-48	OTS	BDO	1:10	121	15	<i>p</i> -TsOH	50mM	12.5	12.81%	21.13%	82.95%
P-49	OTS	BDO	1:10	80	1	-	-	C	90.59%	91.75%	90.22%
P-50	OTS	BDO	1:10	80	3	-	-	C	80.52%	94.69%	89.49%
P-51	OTS	BDO	1:10	80	6	-	-	C	81.13%	93.35%	90.22%
P-52	OTS	BDO	1:10	80	15	-	-	C	76.86%	92.02%	90.95%
P-53	OTS	BDO	1:10	80	24	-	-	C	76.56%	91.48%	90.95%
P-54	OTS	BDO	1:10	121	0.5	-	-	C	88.76%	92.02%	93.86%
P-55	OTS	BDO	1:10	121	1	-	-	C	77.17%	76.77%	93.13%
P-56	OTS	BDO	1:10	121	3	-	-	C	78.39%	72.76%	93.86%

P-57	OTS	BDO	1:10	121	6	-	-	C	76.56%	75.70%	93.13%
P-58	OTS	BDO	1:10	121	15	-	-	C	76.25%	66.34%	93.86%

Table S5-3: Experimental pretreatment conditions and corresponding PDF values for diverse lignocellulosic feedstocks (softwood, hardwood, and grass), together with hemicellulose residue (HR) and lignin residue (LR). The conditions include solvent type, solid-to-liquid ratio, temperature, residence time, and acid dosage. Literature cases (eucalyptus and Acacia confusa) are also included for comparison, further supporting the wide applicability of the PDF model.

Species	Type ^b	Pretreatment Condition						PDF	HR	LR
		Solvent ^c	S:L ^c	Temp (°C)	Time (h)	Acid Type	Acid Conc. ^d			
Pine ^a	SW	BDO	1:10	80	0	HCl	50mM	0.00	100.0%	100.0%
Pine ^a	SW	BDO	1:10	80	1	HCl	50mM	0.20	35.4%	63.8%
Pine ^a	SW	BDO	1:10	80	3	HCl	50mM	0.61	31.0%	64.6%
Pine ^a	SW	BDO	1:10	80	6	HCl	50mM	1.21	23.1%	27.0%
Pine ^a	SW	BDO	1:10	80	15	HCl	50mM	3.03	18.6%	13.2%
Pine ^a	SW	BDO	1:10	80	24	HCl	50mM	4.85	4.3%	10.3%
Pine ^a	SW	BDO	1:10	121	0.5	HCl	50mM	0.20	33.4%	71.3%
Pine ^a	SW	BDO	1:10	121	1	HCl	50mM	0.61	30.5%	66.9%
Pine ^a	SW	BDO	1:10	121	3	HCl	50mM	1.21	15.3%	29.8%
Pine ^a	SW	BDO	1:10	121	6	HCl	50mM	3.03	6.0%	19.0%
Pine ^a	SW	BDO	1:10	121	15	HCl	50mM	4.85	6.5%	13.8%
Eucalyptus ^a	HW	BDO	1:10	121	1	HCl	100mM	1.76	23.60%	29.00%
Bamboo ^a	GS	BDO	1:10	121	1	HCl	100mM	1.76	16.40%	21.38%
Eucalyptus(Dong et al., 2019b)	HW	BDO	1:7	170	1	H ₂ SO ₄	20mM	2.6	13.10%	8.00%
Acacia Confusa (Islam et al., 2024)	HW	EtOH	1:7	170	1	H ₂ SO ₄	50mM	7.1	10.68%	25.93%

Table S5-4:Locations of major polysaccharide peaks in solid-state NMR

		Kirui, Alex, et al.2022(Kirui et al., 2022)	Gao et al. 2020(Gao et al., 2020)	This work
Xn ^{2f}	C1	104.8	105.1	103.8
	C2	72.1	72.3	72.22
	C3	74	75.3	74.52
	C4	82.2	82.1	82
	C5	64	64.1	63.67
Xn ^{3f}	C1	102.5	102.5	101.7
	C2	73.4	73.7	72.22
	C3	74.1	74.7	74.52
	C4	77.4	77.3	77.45
	C5	63.3	63.8	63.67
Wang et al. 2014(Wang et al., 2014)				
Aa	C1	107.8	107.1	107.39
	C2	81.8	89.8	82
	C3	77.2	76.8	77.45
	C4	84.6	78.6	83.82
	C5	61.8	70.7	62.08
Duan et al. 2015(Duan et al., 2021)				
Gal	C1	101.5	99	100.8
	C2	71.3	72	-
	C3	79.8	-	-
	C4	69.9	-	-
	C5	76.6	-	-
ACCO		174	174	173.05
ACMe		21	21.4	20.91

Table S6-1: Retention time in GC/MS and effective carbon number of main monomer calculated (Shuai et al., 2016a)

Monomer		Retention Time	ECN
M1	4-ethyl-2-methoxyphenol	7.79	7
M2	2-methoxy-4-propylphenol	8.12	8
M3	4-(3-hydroxypropyl)-2-methoxyphenol	9.01	7.4
M4	4-ethyl-2,6-dimethoxyphenol	8.63	7
M5	2,6-dimethoxy-4-propylphenol	8.87	8
M6	4-(3-hydroxypropyl)-2,6-dimethoxyphenol	9.65	7.4

Table S6-2: Reaction conditions and lignin monomer yields from catalytic transfer hydrogenolysis of different pretreated samples. Reported values include the Pretreatment Depolymerization Factor (PDF), individual monomer yields (M1-M6), total yield, and S/G ratio.

Entry	Feedstock	Hydrogenolysis	Catalyst	Solvent	S:L	PDF	Yields of lignin monomers							
							M1	M2	M3	M4	M5	M6	Total	S/G
H-1	OTS	230°C H ₂ 40Bar	Ru/C	THF	1:40	0.0	17.80	12.18	46.80	16.90	11.28	35.75	140.71	0.83
H-2	P-6(L)	230°C H ₂ 40Bar	Ru/C	THF	1:200	0.3	32.61	13.27	42.20	29.32	11.74	27.54	156.68	0.78
H-3	P-8(L)	230°C H ₂ 40Bar	Ru/C	THF	1:200	1.3	27.77	16.34	29.92	30.86	20.87	27.70	153.47	1.07
H-4	P-3(L)	230°C H ₂ 40Bar	Ru/C	THF	1:200	1.8	11.13	17.49	34.22	9.55	15.96	27.47	115.82	0.84
H-5	P-13(L)	230°C H ₂ 40Bar	Ru/C	THF	1:200	2.2	8.36	6.83	15.42	6.61	19.79	12.35	69.37	1.27
H-6	P-14(L)	230°C H ₂ 40Bar	Ru/C	THF	1:200	5.5	12.28	5.83	10.74	14.77	8.36	8.98	60.96	1.11
H-7	P-16(L)	230°C H ₂ 40Bar	Ru/C	THF	1:200	15.1	2.30	2.69	9.67	1.47	1.61	5.91	23.64	0.61

Table S6-3: Quantitative hydroxyl content analysis of lignin from ethanol and 1,4-butanediol pretreatment experiments using ³¹P-NMR

	Ethanol						1-4 Butanediol
PDF	0	0.78	3.36	4.05	8.29	12.50	1.60
Aliphatic OH	2.59	2.49	2.42	2.10	1.66	0.83	4.50
Aromatic OH	0.62	0.54	0.47	0.32	0.51	0.79	0.51
Total -OH	3.21	3.03	2.89	2.42	2.17	1.62	5.01

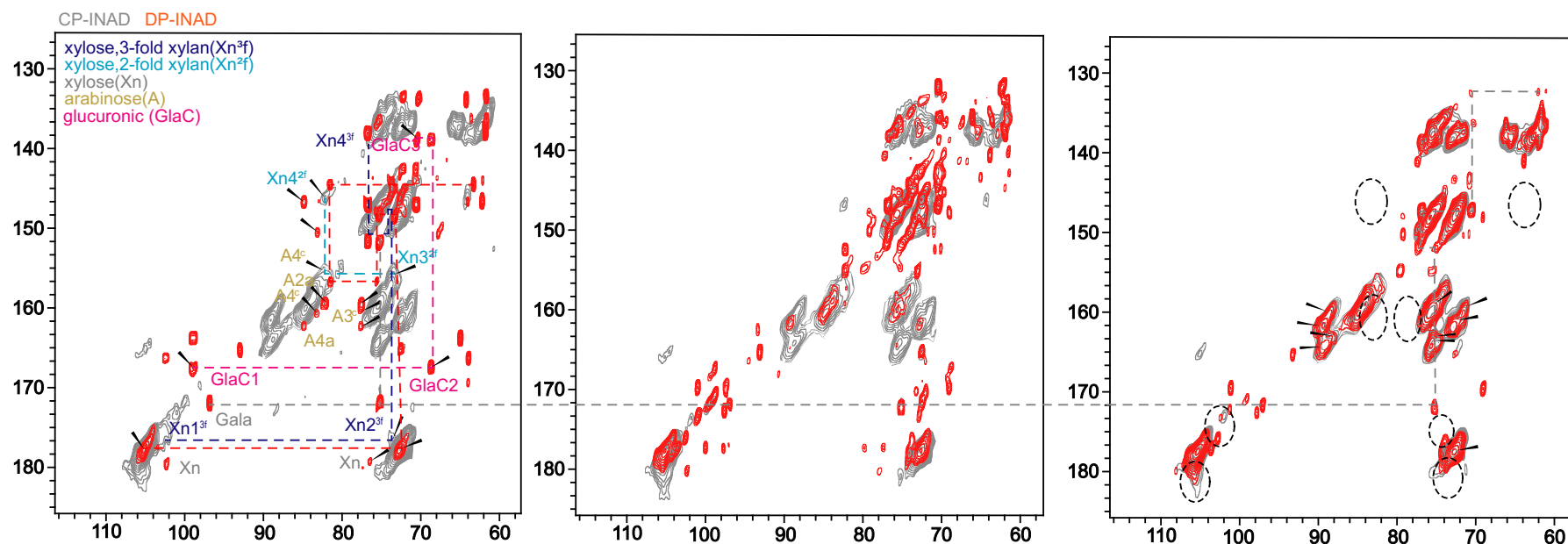


Fig S7-1 Overlay of 2D ^{13}C refocused INADEQUATE spectra acquired with cross polarization (CP, gray) and direct polarization (DP, red) for untreated eucalyptus (left), 1 h pretreated (middle), and 36 h pretreated (right) cell walls. Spectra were recorded on a 600 MHz spectrometer with 14 kHz MAS and ^1H TPPM decoupling at 83 kHz. Peak assignments follow the abbreviations: xylan (Xn), arabinose (A), and glucuronic acid (GlaC), with dashed lines indicating carbon connectivities. The untreated and 1 h spectra show correlations from hemicellulose components (xylan, arabinose, glucuronic acid), whereas the 36 h spectrum is dominated by cellulose signals.

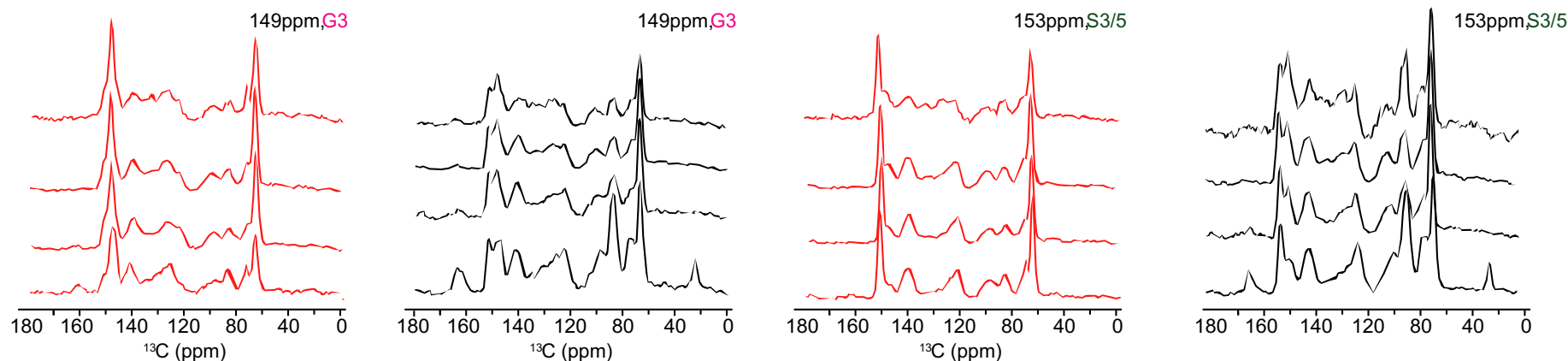


Fig S7-2. Stacked one-dimensional slices extracted at 149 ppm (G3 of G-lignin, left) and 153 ppm (S3/5 of S-lignin, right) from dipolar-gated 2D ^{13}C - ^{13}C correlation spectra of eucalyptus cell walls under increasing pretreatment severity (bottom to top: untreated, 1 h, 18 h, and 36 h). Orange traces correspond to short mixing time (0.1s) spectra, representing predominantly intramolecular correlations, while black traces correspond to long mixing time (1.0s) spectra, capturing both intra- and intermolecular correlations. Comparison of the two mixing times within each treatment condition allows identification of peaks associated with intermolecular contacts versus those originating from within the same lignin aromatic unit. Progressive changes in peak intensities and line shapes with increasing pretreatment highlight alterations in lignin-lignin and lignin-carbohydrate spatial associations, with the star-marked resonances serving as internal reference positions for normalization and alignment across all conditions.

Reference

- Abolore, R.S., Jaiswal, S., Jaiswal, A.K. 2024. Green and sustainable pretreatment methods for cellulose extraction from lignocellulosic biomass and its applications: A review. *Carbohydrate Polymer Technologies and Applications*, **7**, 100396.
- Abu-Omar, M.M., Barta, K., Beckham, G.T., Luterbacher, J.S., Ralph, J., Rinaldi, R., Román-Leshkov, Y., Samec, J.S., Sels, B.F., Wang, F. 2021. Guidelines for performing lignin-first biorefining. *Energy & Environmental Science*, **14**(1), 262–292.
- Achakulwisut, P., Erickson, P., Guivarch, C., Schaeffer, R., Brutschin, E., Pye, S. 2023. Global fossil fuel reduction pathways under different climate mitigation strategies and ambitions. *Nature communications*, **14**(1), 5425.
- Addison, B., Bu, L., Bharadwaj, V., Crowley, M.F., Harman-Ware, A.E., Crowley, M.F., Bomble, Y.J., Ciesielski, P.N. 2024. Atomistic, macromolecular model of the Populus secondary cell wall informed by solid-state NMR. *Science Advances*, **10**(1), eadi7965.
- Ahn, H.G., Lee, J.E., Kim, H., Jung, H.J., Oh, K.K., Heo, S.H., Kim, J.S. 2024. Optimized Furfural Production Using the Acid Catalytic Conversion of Xylan Liquor from Organosolv-Fractionated Rice Husk. *Polysaccharides*, **5**(4), 552–566.
- Alherech, M., Omolabake, S., Holland, C.M., Klinger, G.E., Hegg, E.L., Stahl, S.S. 2021. From lignin to valuable aromatic chemicals: lignin depolymerization and monomer separation via centrifugal partition chromatography. *ACS central science*, **7**(11), 1831–1837.
- Argyropoulos, D.S. 1994. Quantitative phosphorus-31 NMR analysis of lignins, a new tool for the lignin chemist. *Journal of wood chemistry and technology*, **14**(1), 45–63.
- Bai, Y., Zhang, X.-F., Wang, Z., Zheng, T., Yao, J. 2022. Deep eutectic solvent with bifunctional Brønsted-Lewis acids for highly efficient lignocellulose fractionation. *Bioresource Technology*, **347**, 126723.
- Balaban Ucar, M., Uçar, G. 2001. Extractives and Structural Components in Wood and Bark of Endemic Oak Quercus vulcanica Boiss. *Holzforschung*, **55**, 478–486.
- Bär, J., Phongpreecha, T., Singh, S.K., Kral Yilmaz, M., Foster, C.E., Crowe, J.D., Hodge, D.B. 2018. Deconstruction of hybrid poplar to monomeric sugars and aromatics using ethanol organosolv fractionation. *Biomass Conversion and Biorefinery*, **8**, 813–824.
- Boerjan, W., Ralph, J., Baucher, M. 2003. Lignin biosynthesis. *Annual review of plant biology*, **54**(1), 519–546.
- Brienza, F., Van Aelst, K., Devred, F., Magnin, D., Sels, B.F., Gerin, P., Cybulska, I., Debecker, D.P. 2023. Toward a Hydrogen-Free Reductive Catalytic Fractionation of Wheat Straw Biomass. *ChemSusChem*, **16**(13), e202300103.
- Brienza, F., Van Aelst, K., Devred, F., Magnin, D., Sels, B.F., Gerin, P.A., Cybulska, I., Debecker, D.P. 2022. Reductive Catalytic Fractionation of Wheat Straw Biomass. *ACS Sustainable Chemistry & Engineering*, **10**(34), 11130–11142.
- Bugli, F., Baldelli, A., Thomas, S., Sgarzi, M., Gigli, M., Crestini, C., Cavani, F., Tabanelli, T. 2024. Improved Reductive Catalytic Fractionation of Lignocellulosic Biomass through the Application of a Recyclable Magnetic Catalyst. *ACS Sustainable Chemistry & Engineering*, **12**(45), 16638–16651.

- Cai, C., Li, N., Liu, H., Zhang, J., Zhu, J., Wang, F. 2023. Extracting high β -O-4 content lignin and by-producing substrate susceptible to enzymatic hydrolysis by a green flow through process. *Chemical Engineering Journal*, **453**, 139730.
- Cai, C.M., Zhang, T., Kumar, R., Wyman, C.E. 2013. THF co-solvent enhances hydrocarbon fuel precursor yields from lignocellulosic biomass. *Green Chemistry*, **15**(11), 3140–3145.
- Cassidy, E.S., West, P.C., Gerber, J.S., Foley, J.A. 2013. Redefining agricultural yields: from tonnes to people nourished per hectare. *Environmental research letters*, **8**(3), 034015.
- Chan, K.L., Ko, C.H., Chang, K.L., Leu, S.Y. 2021. Construction of a structural enzyme adsorption/kinetics model to elucidate additives associated lignin–cellulase interactions in complex bioconversion system. *Biotechnology and bioengineering*, **118**(10), 4065–4075.
- Chaydarreh, K.C., Lin, X., Guan, L., Yun, H., Gu, J., Hu, C. 2021. Utilization of tea oil camellia (*Camellia oleifera* Abel.) shells as alternative raw materials for manufacturing particleboard. *Industrial crops and products*, **161**, 113221.
- Chen, H. 2014. Chemical composition and structure of natural lignocellulose. in: *Biotechnology of lignocellulose: Theory and practice*, Springer, pp. 25–71.
- Chen, L., Dou, J., Ma, Q., Li, N., Wu, R., Bian, H., Yelle, D.J., Vuorinen, T., Fu, S., Pan, X., Zhu, J. 2017a. Rapid and near-complete dissolution of wood lignin at $\leq 80^{\circ}\text{C}$ by a recyclable acid hydrotrope. *Science Advances*, **3**(9), e1701735.
- Chen, T.-Y., Wang, B., Shen, X.-J., Li, H.-Y., Wu, Y.-Y., Wen, J.-L., Liu, Q.-Y., Sun, R.-C. 2017b. Assessment of structural characteristics of regenerated cellulolytic enzyme lignin based on a mild DMSO/[Emim] OAc dissolution system from triploid of *Populus tomentosa* Carr. *RSC Advances*, **7**(6), 3376–3387.
- Cheng, X., Ning, R., Zhang, F., Ji, L., Wang, K., Jiang, J. 2023. Structural conversion and characterization of *Camellia oleifera* shell lignin based on mechanochemical-assisted pretreatment. *Biomass and Bioenergy*, **170**, 106708.
- Das, L., Li, M., Stevens, J., Li, W., Pu, Y., Ragauskas, A.J., Shi, J. 2018. Characterization and catalytic transfer hydrogenolysis of deep eutectic solvent extracted sorghum lignin to phenolic compounds. *ACS Sustainable Chemistry & Engineering*, **6**(8), 10408–10420.
- Delgado-Mellado, N., Larriba, M., Navarro, P., Rigual, V., Ayuso, M., García, J., Rodríguez, F. 2018. Thermal stability of choline chloride deep eutectic solvents by TGA/FTIR-ATR analysis. *Journal of Molecular Liquids*, **260**, 37–43.
- Dong, C., Meng, X., Leu, S.-Y., Xu, L., Wu, Z., Cravotto, G., Fang, Z. 2022. Enhancing α -etherification of lignin in Eucalyptus diol pretreatment to improve lignin monomer production. *Industrial Crops and Products*, **185**, 115130.
- Dong, C., Meng, X., Yeung, C.S., Ho-Yin, T., Ragauskas, A.J., Leu, S.-Y. 2019a. Diol pretreatment to fractionate a reactive lignin in lignocellulosic biomass biorefineries. *Green Chemistry*, **21**(10), 2788–2800.
- Dong, C., Meng, X., Yeung, C.S., Tse, H.-Y., Ragauskas, A.J., Leu, S.-Y. 2019b. Diol pretreatment to fractionate a reactive lignin in lignocellulosic biomass biorefineries. *Green Chemistry*, **21**(10), 2788–2800.
- Duan, P., Kaser, S.J., Lyczakowski, J.J., Phyto, P., Tryfona, T., Dupree, P., Hong, M. 2021. Xylan structure and dynamics in native *Brachypodium* grass cell walls investigated by solid-state NMR spectroscopy. *Acs Omega*, **6**(23), 15460–15471.

- Fasahati, P., Dickson, R., Saffron, C., Woo, H.C., Liu, J.J. 2022. Seaweeds as a sustainable source of bioenergy: Techno-economic and life cycle analyses of its biochemical conversion pathways. *Renewable and Sustainable Energy Reviews*, **157**, 112011.
- Fernández-Rodríguez, J., Erdocia, X., Hernández-Ramos, F., Alriols, M.G., Labidi, J. 2019. Chapter 7 - Lignin Separation and Fractionation by Ultrafiltration. in: *Separation of Functional Molecules in Food by Membrane Technology*, (Ed.) C.M. Galanakis, Academic Press, pp. 229–265.
- Fernando, L.D., Zhao, W., Gautam, I., Ankur, A., Wang, T. 2023. Polysaccharide assemblies in fungal and plant cell walls explored by solid-state NMR. *Structure*, **31**(11), 1375–1385.
- Gao, Y., Lipton, A.S., Wittmer, Y., Murray, D.T., Mortimer, J.C. 2020. A grass-specific cellulose–xylan interaction dominates in sorghum secondary cell walls. *Nature communications*, **11**(1), 6081.
- Ghassemi, N., Poulhazan, A., Deligey, F., Mentink-Vigier, F., Marcotte, I., Wang, T. 2021. Solid-state NMR investigations of extracellular matrixes and cell walls of algae, bacteria, fungi, and plants. *Chemical reviews*, **122**(10), 10036–10086.
- Gierer, J. 1980. Chemical aspects of kraft pulping. *Wood Science and Technology*, **14**(4), 241–266.
- Guan, J., Khan, A., Zhang, Y., Zhou, Y., Li, M.M.-J., Patria, R.D., Leu, S.-Y. 2025. Strategic design principles for greener biorefineries: a substrate–process matrix emphasizing complete lignocellulose utilization from various biomass feedstocks. *Green Chemistry*, **27**(38), 11581–11606.
- Guerra, A., Filpponen, I., Lucia, L.A., Argyropoulos, D.S. 2006. Comparative evaluation of three lignin isolation protocols for various wood species. *Journal of Agricultural and Food Chemistry*, **54**(26), 9696–9705.
- Guo, H., Miles-Barrett, D.M., Neal, A.R., Zhang, T., Li, C., Westwood, N.J. 2018. Unravelling the enigma of lignin OX: can the oxidation of lignin be controlled? *Chemical Science*, **9**(3), 702–711.
- Han, S., Wang, R., Wang, K., Jiang, J., Xu, J. 2022. Low-condensed lignin and high-purity cellulose production from poplar by synergistic deep eutectic solvent-hydrogenolysis pretreatment. *Bioresource Technology*, **363**, 127905.
- Hansen, B.B., Spittle, S., Chen, B., Poe, D., Zhang, Y., Klein, J.M., Horton, A., Adhikari, L., Zelovich, T., Doherty, B.W., Gurkan, B., Maginn, E.J., Ragauskas, A., Dadmun, M., Zawodzinski, T.A., Baker, G.A., Tuckerman, M.E., Savinell, R.F., Sangoro, J.R. 2021. Deep Eutectic Solvents: A Review of Fundamentals and Applications. *Chemical Reviews*, **121**(3), 1232–1285.
- Hideno, A., Kawashima, A., Endo, T., Honda, K., Morita, M. 2013. Ethanol-based organosolv treatment with trace hydrochloric acid improves the enzymatic digestibility of Japanese cypress (*Chamaecyparis obtusa*) by exposing nanofibers on the surface. *Bioresource Technology*, **132**, 64–70.
- Hrůzová, K., Matsakas, L., Rova, U., Christakopoulos, P. 2021. Organosolv fractionation of spruce bark using ethanol–water mixtures: Towards a novel bio-refinery concept. *Bioresource Technology*, **341**, 125855.

- Huang, C., Zhan, Y., Wang, J., Cheng, J., Meng, X., Liang, L., Liang, F., Deng, Y., Fang, G., Ragauskas, A.J. 2022. Valorization of bamboo biomass using combinatorial pretreatments. *Green Chemistry*, **24**(9), 3736–3749.
- Huang, F., Singh, P.M., Ragauskas, A.J. 2011. Characterization of milled wood lignin (MWL) in loblolly pine stem wood, residue, and bark. *Journal of agricultural and food chemistry*, **59**(24), 12910–12916.
- Huang, H., Zhang, X., Ma, L., Liao, Y. Reductive Catalytic Fractionation of Lignocellulose Toward Propyl- or Propenyl-Substituted Monomers and Mechanistic Understanding. *Angewandte Chemie International Edition*, **n/a**(n/a), e202502545.
- Islam, M.K., Guan, J., Rehman, S., Patria, R.D., Hu, C., Guan, L., Leu, S.-Y., Vuppaladadiyam, A.K. 2024. N-Methyl-2-pyrrolidone pre-treatment of lignocellulose for high lignin yield and cellulose digestibility. *Biomass Conversion and Biorefinery*, **14**(4), 5435–5446.
- Islam, M.K., Rehman, S., Guan, J., Lau, C.-Y., Tse, H.-Y., Yeung, C.S., Leu, S.-Y. 2021a. Biphasic pretreatment for energy and carbon efficient conversion of lignocellulose into bioenergy and reactive lignin. *Applied Energy*, **303**, 117653.
- Islam, M.K., Thaemngoan, A., Lau, C.Y., Guan, J., Yeung, C.S., Chaiprapat, S., Leu, S.-Y. 2021b. Staged organosolv pretreatment to increase net energy and reactive lignin yield in whole oil palm tree biorefinery. *Bioresource Technology*, **326**, 124766.
- Islam, M.K., Wang, H., Rehman, S., Dong, C., Hsu, H.-Y., Lin, C.S.K., Leu, S.-Y. 2020a. Sustainability metrics of pretreatment processes in a waste derived lignocellulosic biomass biorefinery. *Bioresource technology*, **298**, 122558.
- Islam, M.K., Wang, H., Rehman, S., Dong, C., Hsu, H.Y., Lin, C.S.K., Leu, S.Y. 2020b. Sustainability metrics of pretreatment processes in a waste derived lignocellulosic biomass biorefinery. *Bioresour Technol*, **298**, 122558.
- Jiang, B., Zhang, Y., Guo, T., Zhao, H., Jin, Y. 2018. Structural characterization of lignin and lignin-carbohydrate complex (LCC) from ginkgo shells (*Ginkgo biloba* L.) by comprehensive NMR spectroscopy. *Polymers*, **10**(7), 736.
- Jiang, Z., Fei, B., Li, Z. 2016. Pretreatment of bamboo by ultra-high pressure explosion with a high-pressure homogenizer for enzymatic hydrolysis and ethanol fermentation. *Bioresource Technology*, **214**, 876–880.
- Jindal, M., Uniyal, P., Thallada, B. 2023. Reductive catalytic fractionation as a novel pretreatment/lignin-first approach for lignocellulosic biomass valorization: A review. *Bioresource Technology*, **385**, 129396.
- Johnson, R.L., Schmidt-Rohr, K. 2014. Quantitative solid-state ¹³C NMR with signal enhancement by multiple cross polarization. *Journal of Magnetic Resonance*, **239**, 44–49.
- Jomnonkhaow, U., Sittijunda, S., Reungsang, A. 2022. Assessment of organosolv, hydrothermal, and combined organosolv and hydrothermal with enzymatic pretreatment to increase the production of biogas from Napier grass and Napier silage. *Renewable Energy*, **181**, 1237–1249.
- Joshi, O., Grebner, D.L., Khanal, P.N. 2015. Status of urban wood-waste and their potential use for sustainable bioenergy in Mississippi. *Resources, conservation and recycling*, **102**, 20–26.
- Joy, S.P., Krishnan, C. 2022. Modified organosolv pretreatment for improved cellulosic ethanol production from sorghum biomass. *Industrial Crops and Products*, **177**, 114409.

- Kang, X., Kirui, A., Dickwella Widanage, M.C., Mentink-Vigier, F., Cosgrove, D.J., Wang, T. 2019. Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nature communications*, **10**(1), 347.
- Khan, R., Jianyu, G., Lau, C., Zhuang, H., Rehman, S., Leu, S.-Y. 2024a. Monolignol Potential and Insights into Direct Depolymerization of Fruit and Nutshell Remains for High Value Sustainable Aromatics. *ChemSusChem*, **17**, e202301306.
- Khan, R.J., Guan, J., Lau, C.Y., Zhuang, H., Rehman, S., Leu, S.-Y. 2023. Monolignol Potential and Insights into Direct Depolymerization of Fruit and Nutshell Remains for High Value Sustainable Aromatics. *ChemSusChem*, e202301306.
- Khan, R.J., Guan, J., Lau, C.Y., Zhuang, H., Rehman, S., Leu, S.Y. 2024b. Monolignol Potential and Insights into Direct Depolymerization of Fruit and Nutshell Remains for High Value Sustainable Aromatics. *ChemSusChem*, **17**(7), e202301306.
- Khan, R.J., Lau, C.Y., Guan, J., Lam, C.H., Zhao, J., Ji, Y., Wang, H., Xu, J., Lee, D.-J., Leu, S.-Y. 2022. Recent advances of lignin valorization techniques toward sustainable aromatics and potential benchmarks to fossil refinery products. *Bioresource technology*, **346**, 126419.
- Kim, J.S., Lee, Y.Y., Kim, T.H. 2016. A review on alkaline pretreatment technology for bioconversion of lignocellulosic biomass. *Bioresource Technology*, **199**, 42–48.
- Kim, K.H., Dutta, T., Sun, J., Simmons, B., Singh, S. 2018. Biomass pretreatment using deep eutectic solvents from lignin derived phenols. *Green chemistry*, **20**(4), 809–815.
- Kim, Y., Yu, A., Han, M., Choi, G.W., Chung, B. 2010. Ethanosolv pretreatment of barley straw with iron (III) chloride for enzymatic saccharification. *Journal of Chemical Technology & Biotechnology*, **85**(11), 1494–1498.
- Kirui, A., Zhao, W., Deligey, F., Yang, H., Kang, X., Mentink-Vigier, F., Wang, T. 2022. Carbohydrate-aromatic interface and molecular architecture of lignocellulose. *Nature communications*, **13**(1), 538.
- Klein, I., Saha, B., Abu-Omar, M.M. 2015. Lignin depolymerization over Ni/C catalyst in methanol, a continuation: effect of substrate and catalyst loading. *Catalysis Science & Technology*, **5**(6), 3242–3245.
- Kuglarz, M., Alvarado-Morales, M., Dąbkowska, K., Angelidaki, I. 2018. Integrated production of cellulosic bioethanol and succinic acid from rapeseed straw after dilute-acid pretreatment. *Bioresource Technology*, **265**, 191–199.
- Lan, W., Amiri, M.T., Hunston, C.M., Luterbacher, J.S. 2018a. Protection group effects during α , γ -diol lignin stabilization promote high-selectivity monomer production. *Angewandte Chemie International Edition*, **57**(5), 1356–1360.
- Lan, W., Amiri, M.T., Hunston, C.M., Luterbacher, J.S. 2018b. Protection Group Effects During α , γ -Diol Lignin Stabilization Promote High-Selectivity Monomer Production. *Angewandte Chemie International Edition*, **57**(5), 1356–1360.
- Lan, W., Amiri, M.T., Hunston, C.M., Luterbacher, J.S. 2018c. Rücktitelbild: Protection Group Effects During α , γ -Diol Lignin Stabilization Promote High-Selectivity Monomer Production (Angew. Chem. 5/2018). *Angewandte Chemie*, **130**(5), 1434–1434.
- Lee, K., Jing, Y., Wang, Y., Yan, N. 2022. A unified view on catalytic conversion of biomass and waste plastics. *Nature Reviews Chemistry*, **6**(9), 635–652.

- Leu, S.-Y., Zhu, J. 2013a. Substrate-related factors affecting enzymatic saccharification of lignocelluloses: our recent understanding. *BioEnergy Research*, **6**(2), 405–415.
- Leu, S.-Y., Zhu, J., Gleisner, R., Sessions, J., Marrs, G. 2013. Robust enzymatic saccharification of a Douglas-fir forest harvest residue by SPORL. *biomass and bioenergy*, **59**, 393–401.
- Leu, S.-Y., Zhu, J.Y. 2013b. Substrate-Related Factors Affecting Enzymatic Saccharification of Lignocelluloses: Our Recent Understanding. *BioEnergy Research*, **6**(2), 405–415.
- Li, H., Bunrit, A., Li, N., Wang, F. 2020. Heteroatom-participated lignin cleavage to functionalized aromatics. *Chemical Society Reviews*, **49**(12), 3748–3763.
- Li, H., Song, G. 2019. Ru-catalyzed hydrogenolysis of lignin: base-dependent tunability of monomeric phenols and mechanistic study. *ACS Catalysis*, **9**(5), 4054–4064.
- Li, N., Li, Y., Yoo, C.G., Yang, X., Lin, X., Ralph, J., Pan, X. 2018a. An uncondensed lignin depolymerized in the solid state and isolated from lignocellulosic biomass: a mechanistic study. *Green Chemistry*, **20**(18), 4224–4235.
- Li, N., Yan, K., Rukkijakan, T., Liang, J., Liu, Y., Wang, Z., Nie, H., Muangmeesri, S., Castiella-Ona, G., Pan, X. 2024a. Selective lignin arylation for biomass fractionation and benign bisphenols. *Nature*, 1–6.
- Li, N., Yan, K., Rukkijakan, T., Liang, J., Liu, Y., Wang, Z., Nie, H., Muangmeesri, S., Castiella-Ona, G., Pan, X. 2024b. Selective lignin arylation for biomass fractionation and benign bisphenols. *Nature*, **630**(8016), 381–386.
- Li, Y., Shuai, L., Kim, H., Motagamwala, A.H., Mobley, J.K., Yue, F., Tobimatsu, Y., Havkin-Frenkel, D., Chen, F., Dixon, R.A. 2018b. An “ideal lignin” facilitates full biomass utilization. *Science advances*, **4**(9), eaau2968.
- Li, Y., Shuai, L., Kim, H., Motagamwala, A.H., Mobley, J.K., Yue, F., Tobimatsu, Y., Havkin-Frenkel, D., Chen, F., Dixon, R.A., Luterbacher, J.S., Dumesic, J.A., Ralph, J. 2018c. An “ideal lignin” facilitates full biomass utilization. *Science Advances*, **4**(9), eaau2968.
- Liao, Y., Koelewijn, S.-F., Van den Bossche, G., Van Aelst, J., Van den Bosch, S., Renders, T., Navare, K., Nicolai, T., Van Aelst, K., Maesen, M. 2020. A sustainable wood biorefinery for low-carbon footprint chemicals production. *Science*, **367**(6484), 1385–1390.
- Lin, Y., Tanaka, S. 2006. Ethanol fermentation from biomass resources: current state and prospects. *Applied microbiology and biotechnology*, **69**(6), 627–642.
- Liu, H., Sun, J., Leu, S.Y., Chen, S. 2016. Toward a fundamental understanding of cellulase-lignin interactions in the whole slurry enzymatic saccharification process. *Biofuels, Bioproducts and Biorefining*, **10**(5), 648–663.
- Liu, L.-Y., Patankar, S.C., Chandra, R.P., Sathitsuksanoh, N., Saddler, J.N., Renneckar, S. 2020. Valorization of bark using ethanol–water organosolv treatment: isolation and characterization of crude lignin. *ACS Sustainable Chemistry & Engineering*, **8**(12), 4745–4754.
- Liu, Y., Deak, N., Wang, Z., Yu, H., Hameleers, L., Jurak, E., Deuss, P.J., Barta, K. 2021. Tunable and functional deep eutectic solvents for lignocellulose valorization. *Nature Communications*, **12**(1), 5424.
- Liu, Z., Li, H., Gao, X., Guo, X., Wang, S., Fang, Y., Song, G. 2022. Rational highly dispersed ruthenium for reductive catalytic fractionation of lignocellulose. *Nature Communications*, **13**(1), 4716.

- Marchel, M., Cieśliński, H., Boczkaj, G. 2022. Thermal Instability of Choline Chloride-Based Deep Eutectic Solvents and Its Influence on Their Toxicity—Important Limitations of DESs as Sustainable Materials. *Industrial & Engineering Chemistry Research*, **61**(30), 11288–11300.
- Martínez, C.M., Cantero, D.A., Cocero, M.J. 2018. Production of saccharides from sugar beet pulp by ultrafast hydrolysis in supercritical water. *Journal of Cleaner Production*, **204**, 888–895.
- Massiot, D., Fayon, F., Capron, M., King, I., Le Calvé, S., Alonso, B., Durand, J.O., Bujoli, B., Gan, Z., Hoatson, G. 2002. Modelling one-and two-dimensional solid-state NMR spectra. *Magnetic resonance in chemistry*, **40**(1), 70–76.
- Matsakas, L., Raghavendran, V., Yakimenko, O., Persson, G., Olsson, E., Rova, U., Olsson, L., Christakopoulos, P. 2019. Lignin-first biomass fractionation using a hybrid organosolv–steam explosion pretreatment technology improves the saccharification and fermentability of spruce biomass. *Bioresource technology*, **273**, 521–528.
- Meng, X., Pu, Y., Li, M., Ragauskas, A.J. 2020. A biomass pretreatment using cellulose-derived solvent Cyrene. *Green Chemistry*, **22**(9), 2862–2872.
- Mnich, E., Bjarnholt, N., Eudes, A., Harholt, J., Holland, C., Jørgensen, B., Larsen, F.H., Liu, M., Manat, R., Meyer, A.S. 2020. Phenolic cross-links: building and de-constructing the plant cell wall. *Natural product reports*, **37**(7), 919–961.
- Nicolai, T., Arts, W., Calderon-Ardila, S., Smets, R., Van Der Borght, M., Thevelein, J.M., Sels, B.F. 2024. Uncovering Challenges for Complete Carbohydrate to Bioethanol Utilization in a Reductive Catalytic Fractionation Biorefinery. *ACS Sustainable Chemistry & Engineering*, **12**(22), 8353–8365.
- Obama, P., Ricochon, G., Muniglia, L., Brosse, N. 2012. Combination of enzymatic hydrolysis and ethanol organosolv pretreatments: Effect on lignin structures, delignification yields and cellulose-to-glucose conversion. *Bioresource technology*, **112**, 156–163.
- Ouyang, X., Huang, X., Hendriks, B.M., Boot, M.D., Hensen, E.J. 2018. Coupling organosolv fractionation and reductive depolymerization of woody biomass in a two-step catalytic process. *Green Chemistry*, **20**(10), 2308–2319.
- Pals, M., Lauberts, M., Zijlstra, D.S., Ponomarenko, J., Arshanitsa, A., Deuss, P.J. 2022. Mild organosolv delignification of residual aspen bark after extractives isolation as a step in biorefinery processing schemes. *Molecules*, **27**(10), 3185.
- Pan, X., Arato, C., Gilkes, N., Gregg, D., Mabee, W., Pye, K., Xiao, Z., Zhang, X., Saddler, J. 2005. Biorefining of softwoods using ethanol organosolv pulping: Preliminary evaluation of process streams for manufacture of fuel-grade ethanol and co-products. *Biotechnology and bioengineering*, **90**(4), 473–481.
- Patel, R., Rajaraman, T., Rana, P.H., Ambegaonkar, N.J., Patel, S. 2025. A review on techno-economic analysis of lignocellulosic biorefinery producing biofuels and high-value products. *Results in Chemistry*, **13**, 102052.
- Pepper, J.M., Fleming, R.W. 1978. Lignin and related compounds. V. The hydrogenolysis of aspen wood lignin using rhodium-on-charcoal as catalyst. *Canadian Journal of Chemistry*, **56**, 896–898.
- Pepper, J.M., Hibbert, H. 1948. Studies on lignin and related compounds; high pressure hydrogenation of maple wood. *J Am Chem Soc*, **70**(1), 67–71.

- Phongpreecha, T., Hool, N.C., Stoklosa, R.J., Klett, A.S., Foster, C.E., Bhalla, A., Holmes, D., Thies, M.C., Hodge, D.B. 2017. Predicting lignin depolymerization yields from quantifiable properties using fractionated biorefinery lignins. *Green Chemistry*, **19**(21), 5131–5143.
- Ponnusamy, V.K., Nguyen, D.D., Dharmaraja, J., Shobana, S., Banu, J.R., Saratale, R.G., Chang, S.W., Kumar, G. 2019. A review on lignin structure, pretreatments, fermentation reactions and biorefinery potential. *Bioresource technology*, **271**, 462–472.
- Prasad, B., Padhi, R., Ghosh, G. 2023. A review on key pretreatment approaches for lignocellulosic biomass to produce biofuel and value-added products. *International Journal of Environmental Science and Technology*, **20**(6), 6929–6944.
- Qin, H., Hu, X., Wang, J., Cheng, H., Chen, L., Qi, Z. 2020. Overview of acidic deep eutectic solvents on synthesis, properties and applications. *Green Energy & Environment*, **5**(1), 8–21.
- Ragauskas, A.J., Beckham, G.T., Biddy, M.J., Chandra, R., Chen, F., Davis, M.F., Davison, B.H., Dixon, R.A., Gilna, P., Keller, M. 2014. Lignin valorization: improving lignin processing in the biorefinery. *science*, **344**(6185), 1246843.
- Ramos, L.P. 2003. The chemistry involved in the steam treatment of lignocellulosic materials. *Química Nova*, **26**, 863–871.
- Ren, T., You, S., Zhang, Z., Wang, Y., Qi, W., Su, R., He, Z. 2021. Highly selective reductive catalytic fractionation at atmospheric pressure without hydrogen. *Green Chemistry*, **23**(4), 1648–1657.
- Renders, T., Cooreman, E., Van den Bosch, S., Schutyser, W., Koelewijn, S.F., Vangeel, T., Deneyer, A., Van den Bossche, G., Courtin, C.M., Sels, B.F. 2018. Catalytic lignocellulose biorefining in n-butanol/water: a one-pot approach toward phenolics, polyols, and cellulose. *Green Chemistry*, **20**(20), 4607–4619.
- Rietzler, B., Karlsson, M., Kwan, I., Lawoko, M., Ek, M. 2022. Fundamental insights on the physical and chemical properties of organosolv lignin from Norway spruce bark. *Biomacromolecules*, **23**(8), 3349–3358.
- Robak, K., Balcerek, M. 2020. Current state-of-the-art in ethanol production from lignocellulosic feedstocks. *Microbiological Research*, **240**, 126534.
- Rodriguez Rodriguez, N., van den Bruinhorst, A., Kollau, L.J.B.M., Kroon, M.C., Binnemans, K. 2019. Degradation of Deep-Eutectic Solvents Based on Choline Chloride and Carboxylic Acids. *ACS Sustainable Chemistry & Engineering*, **7**(13), 11521–11528.
- Romaní, A., Larramendi, A., Yáñez, R., Cancela, Á., Sánchez, Á., Teixeira, J.A., Domingues, L. 2019. Valorization of Eucalyptus nitens bark by organosolv pretreatment for the production of advanced biofuels. *Industrial Crops and Products*, **132**, 327–335.
- Samec, J., Muangmeesri, S., Lebedeva, D., Ramazanov, L., Liu, S., Khalili, H., P Mathew, A., Ralph, J. 2024. *Oxidative Valorization of Spruce Bark to Yield Vanillin*.
- Sannigrahi, P., Miller, S.J., Ragauskas, A.J. 2010. Effects of organosolv pretreatment and enzymatic hydrolysis on cellulose structure and crystallinity in Loblolly pine. *Carbohydrate research*, **345**(7), 965–970.

- Schell, D.J., Farmer, J., Newman, M., McMILLAN, J.D. 2003. Dilute-sulfuric acid pretreatment of corn stover in pilot-scale reactor: investigation of yields, kinetics, and enzymatic digestibilities of solids. *Applied biochemistry and biotechnology*, **105**, 69–85.
- Sella Kapu, N., Trajano, H.L. 2014. Review of hemicellulose hydrolysis in softwoods and bamboo. *Biofuels, Bioproducts and Biorefining*, **8**(6), 857–870.
- Shao, Z., Fu, Y., Wang, P., Zhang, Y., Qin, M., Li, X., Zhang, F. 2020. Modification of the aspen lignin structure during integrated fractionation process of autohydrolysis and formic acid delignification. *International Journal of Biological Macromolecules*, **165**, 1727–1737.
- Shen, G., Yuan, X., Cheng, Y., Chen, S., Xu, Z., Jin, M. 2023. Densification pretreatment with a limited deep eutectic solvent triggers high-efficiency fractionation and valorization of lignocellulose. *Green Chemistry*, **25**(20), 8026–8039.
- Shen, X.-J., Wen, J.-L., Mei, Q.-Q., Chen, X., Sun, D., Yuan, T.-Q., Sun, R.-C. 2019. Facile fractionation of lignocelluloses by biomass-derived deep eutectic solvent (DES) pretreatment for cellulose enzymatic hydrolysis and lignin valorization. *Green chemistry*, **21**(2), 275–283.
- Shuai, L., Amiri, M.T., Questell-Santiago, Y.M., Héroguel, F., Li, Y., Kim, H., Meilan, R., Chapple, C., Ralph, J., Luterbacher, J.S. 2016a. Formaldehyde stabilization facilitates lignin monomer production during biomass depolymerization. *Science*, **354**(6310), 329–333.
- Shuai, L., Amiri, M.T., Questell-Santiago, Y.M., Heroguel, F., Li, Y., Kim, H., Meilan, R., Chapple, C., Ralph, J., Luterbacher, J.S. 2016b. Formaldehyde stabilization facilitates lignin monomer production during biomass depolymerization. *Science*, **354**(6310), 329–333.
- Simmons, T.J., Mortimer, J.C., Bernardinelli, O.D., Pöppler, A.-C., Brown, S.P., Deazevedo, E.R., Dupree, R., Dupree, P. 2016. Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nature communications*, **7**(1), 13902.
- Singhvi, M.S., Gokhale, D.V. 2019. Lignocellulosic biomass: hurdles and challenges in its valorization. *Applied microbiology and biotechnology*, **103**(23), 9305–9320.
- Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., Crocker, D. 2008. Determination of structural carbohydrates and lignin in biomass. *Laboratory analytical procedure*, **1617**(1), 1–16.
- Smit, A.T., Hoek, M., Bonouvrie, P.A., van Zomeren, A., Riddell, L.A., Bruijninx, P.C. 2024. Semicontinuous aqueous acetone organosolv fractionation of lignocellulosic biomass: improved biorefinery processing and output. *ACS Sustainable Chemistry & Engineering*, **12**(11), 4731–4742.
- Stone, M.L., Anderson, E.M., Meek, K.M., Reed, M., Katahira, R., Chen, F., Dixon, R.A., Beckham, G.T., Román-Leshkov, Y. 2018. Reductive Catalytic Fractionation of C-Lignin. *ACS Sustainable Chemistry & Engineering*, **6**(9), 11211–11218.
- Su, S., Shen, Q., Wang, S., Song, G. 2023a. Discovery, disassembly, depolymerization and derivatization of catechyl lignin in Chinese tallow seed coats. *International Journal of Biological Macromolecules*, **239**, 124256.

- Su, S., Shen, Q., Wang, S., Song, G. 2023b. Discovery, disassembly, depolymerization and derivatization of catechyl lignin in Chinese tallow seed coats. *Int J Biol Macromol*, **239**, 124256.
- Su, S., Wang, S., Song, G. 2021. Disassembling catechyl and guaiacyl/syringyl lignins coexisting in Euphorbiaceae seed coats. *Green Chemistry*, **23**(18), 7235–7242.
- Sun, J., Ding, R., Yin, J. 2022. Pretreatment of corn cobs and corn stalks with tetrabutyl phosphate hydroxide ionic liquid to enhance enzymatic hydrolysis process. *Biochemical Engineering Journal*, **177**, 108270.
- Sun, Z., Fridrich, B., de Santi, A., Elangovan, S., Barta, K. 2018. Bright Side of Lignin Depolymerization: Toward New Platform Chemicals. *Chemical Reviews*, **118**(2), 614–678.
- Syvitski, J., Waters, C.N., Day, J., Milliman, J.D., Summerhayes, C., Steffen, W., Zalasiewicz, J., Cearreta, A., Gałuszka, A., Hajdas, I. 2020. Extraordinary human energy consumption and resultant geological impacts beginning around 1950 CE initiated the proposed Anthropocene Epoch. *Communications Earth & Environment*, **1**(1), 32.
- Terrett, O.M., Lyczakowski, J.J., Yu, L., Iuga, D., Franks, W.T., Brown, S.P., Dupree, R., Dupree, P. 2019. Molecular architecture of softwood revealed by solid-state NMR. *Nature communications*, **10**(1), 4978.
- Usia, M., Kara, S. 1997. The chemical composition of wood and bark of *Cedrus libani* A. Rich. *Holz als Roh- und Werkstoff*, **55**(2), 268–268.
- Van Aelst, K., Van Sinay, E., Vangeel, T., Cooreman, E., Van den Bossche, G., Renders, T., Van Aelst, J., Van den Bosch, S., Sels, B.F. 2020. Reductive catalytic fractionation of pine wood: elucidating and quantifying the molecular structures in the lignin oil. *Chemical Science*, **11**(42), 11498–11508.
- Van den Bosch, S., Schutyser, W., Vanholme, R., Driessen, T., Koelewijn, S.F., Renders, T., De Meester, B., Huijgen, W.J.J., Dehaen, W., Courtin, C.M., Lagrain, B., Boerjan, W., Sels, B.F. 2015. Reductive lignocellulose fractionation into soluble lignin-derived phenolic monomers and dimers and processable carbohydrate pulps. *Energy & Environmental Science*, **8**(6), 1748–1763.
- Vanderghem, C., Brostaux, Y., Jacquet, N., Blecker, C., Paquot, M. 2012. Optimization of formic/acetic acid delignification of *Miscanthus × giganteus* for enzymatic hydrolysis using response surface methodology. *Industrial Crops and Products*, **35**(1), 280–286.
- Vangeel, T., Renders, T., Van Aelst, K., Cooreman, E., Van den Bosch, S., Van den Bossche, G., Koelewijn, S.-F., Courtin, C., Sels, B. 2019a. Reductive catalytic fractionation of black locust bark. *Green Chemistry*, **21**.
- Vangeel, T., Renders, T., Van Aelst, K., Cooreman, E., Van den Bosch, S., Van den Bossche, G., Koelewijn, S.F., Courtin, C.M., Sels, B.F. 2019b. Reductive catalytic fractionation of black locust bark. *Green Chemistry*, **21**(21), 5841–5851.
- Vassilev, S.V., Baxter, D., Andersen, L.K., Vassileva, C.G. 2010. An overview of the chemical composition of biomass. *Fuel*, **89**(5), 913–933.
- Wang, J.-X., Asano, S., Kudo, S., Hayashi, J.-i. 2020a. Deep Delignification of Woody Biomass by Repeated Mild Alkaline Treatments with Pressurized O₂. *ACS Omega*, **5**(45), 29168–29176.

- Wang, L., Min, M., Li, Y., Chen, P., Chen, Y., Liu, Y., Wang, Y., Ruan, R. 2010. Cultivation of green algae *Chlorella* sp. in different wastewaters from municipal wastewater treatment plant. *Applied biochemistry and biotechnology*, **162**, 1174–1186.
- Wang, P., Wang, S., Su, S., Zhang, D., Liao, Y., Song, G., Wang, L. 2025. A techno-economically feasible and sustainable C-lignin biorefinery. *Green Chemistry*, **27**(12), 3207–3216.
- Wang, S., Dai, G., Yang, H., Luo, Z. 2017. Lignocellulosic biomass pyrolysis mechanism: A state-of-the-art review. *Progress in Energy and Combustion Science*, **62**, 33–86.
- Wang, S., Su, S., Xiao, L.-P., Wang, B., Sun, R.-C., Song, G. 2020b. Catechyl Lignin Extracted from Castor Seed Coats Using Deep Eutectic Solvents: Characterization and Depolymerization. *ACS Sustainable Chemistry & Engineering*, **8**(18), 7031–7038.
- Wang, S., Zhang, K., Li, H., Xiao, L.-P., Song, G. 2021. Selective hydrogenolysis of catechyl lignin into propenylcatechol over an atomically dispersed ruthenium catalyst. *Nature Communications*, **12**(1), 416.
- Wang, T., Hong, M. 2016. Solid-state NMR investigations of cellulose structure and interactions with matrix polysaccharides in plant primary cell walls. *Journal of experimental botany*, **67**(2), 503–514.
- Wang, T., Salazar, A., Zabolina, O.A., Hong, M. 2014. Structure and dynamics of Brachypodium primary cell wall polysaccharides from two-dimensional ¹³C solid-state nuclear magnetic resonance spectroscopy. *Biochemistry*, **53**(17), 2840–2854.
- Wang, Z., Qiu, S., Hirth, K., Cheng, J., Wen, J., Li, N., Fang, Y., Pan, X., Zhu, J. 2019a. Preserving both lignin and cellulose chemical structures: flow-through acid hydrotropic fractionation at atmospheric pressure for complete wood valorization. *ACS Sustainable Chemistry & Engineering*, **7**(12), 10808–10820.
- Wang, Z., Qiu, S., Hirth, K., Cheng, J., Wen, J., Li, N., Fang, Y., Pan, X., Zhu, J.Y. 2019b. Preserving Both Lignin and Cellulose Chemical Structures: Flow-Through Acid Hydrotropic Fractionation at Atmospheric Pressure for Complete Wood Valorization. *ACS Sustainable Chemistry & Engineering*, **7**(12), 10808–10820.
- Wang, Z., Winstrand, S., Gillgren, T., Jönsson, L.J. 2018. Chemical and structural factors influencing enzymatic saccharification of wood from aspen, birch and spruce. *Biomass & Bioenergy*, **109**, 125–134.
- Warren-Walker, D., Ravella, S.R., Gallagher, J., Winters, A., Charlton, A., Bryant, D.N. 2024. Optimising parameters for pilot scale steam explosion and continuous pressurised disc refining of *Miscanthus* and sugarcane bagasse for xylose and xylo-oligosaccharide release. *Bioresource Technology*, **405**, 130932.
- Welsby, D., Price, J., Pye, S., Ekins, P. 2021. Unextractable fossil fuels in a 1.5 C world. *Nature*, **597**(7875), 230–234.
- Wildschut, J., Smit, A.T., Reith, J.H., Huijgen, W.J. 2013. Ethanol-based organosolv fractionation of wheat straw for the production of lignin and enzymatically digestible cellulose. *Bioresource technology*, **135**, 58–66.
- Woiciechowski, A.L., Neto, C.J.D., de Souza Vandenberghe, L.P., de Carvalho Neto, D.P., Sydney, A.C.N., Letti, L.A.J., Karp, S.G., Torres, L.A.Z., Soccol, C.R. 2020. Lignocellulosic biomass: Acid and alkaline pretreatments and their effects on biomass

- recalcitrance—Conventional processing and recent advances. *Bioresource technology*, **304**, 122848.
- Xia, W., Cui, C., Shao, L., Liu, Y., Li, X., Wang, C., Zhao, D., Xu, F. 2025. Efficient separation of catechyl lignin from castor seed coats via molten salt hydrate. *Separation and Purification Technology*, **353**, 128487.
- Xu, L., Zhang, J., Zong, Q.-J., Wang, L., Xu, T., Gong, J., Liu, Z.-H., Li, B.-Z., Yuan, Y.-J. 2022. High-solid ethylenediamine pretreatment to fractionate new lignin streams from lignocellulosic biomass. *Chemical Engineering Journal*, **427**, 130962.
- Xu, X., Wang, K., Zhou, Y., Lai, C., Zhang, D., Xia, C., Pugazhendhi, A. 2023. Comparison of organosolv pretreatment of masson pine with different solvents in promoting delignification and enzymatic hydrolysis efficiency. *Fuel*, **338**, 127361.
- Yamakawa, C.K., Qin, F., Mussatto, S.I. 2018. Advances and opportunities in biomass conversion technologies and biorefineries for the development of a bio-based economy. *Biomass and Bioenergy*, **119**, 54–60.
- Yamashita, Y., Sasaki, C., Nakamura, Y. 2010. Effective enzyme saccharification and ethanol production from Japanese cedar using various pretreatment methods. *Journal of Bioscience and Bioengineering*, **110**(1), 79–86.
- Yan, J., Oyediji, O., Leal, J.H., Donohoe, B.S., Semelsberger, T.A., Li, C., Hoover, A.N., Webb, E., Bose, E.A., Zeng, Y. 2020. Characterizing variability in lignocellulosic biomass: a review. *ACS Sustainable Chemistry & Engineering*, **8**(22), 8059–8085.
- Yang, H., Yoo, C.G., Meng, X., Pu, Y., Muchero, W., Tuskan, G.A., Tschaplinski, T.J., Ragauskas, A.J., Yao, L. 2020. Structural changes of lignins in natural *Populus* variants during different pretreatments. *Bioresource technology*, **295**, 122240.
- Yin, W.-Z., Zou, S.-L., Xiao, L.-P., Sun, R.-C. 2024. Catechyl lignin extracted from candlenut by biphasic 2-methyltetrahydrofuran/water: Characterization and depolymerization. *Chemical Engineering Science*, **288**, 119828.
- Yoo, C.G., Li, M., Meng, X., Pu, Y., Ragauskas, A.J. 2017. Effects of organosolv and ammonia pretreatments on lignin properties and its inhibition for enzymatic hydrolysis. *Green Chemistry*, **19**(8), 2006–2016.
- Yu, G., Li, B., Liu, C., Zhang, Y., Wang, H., Mu, X. 2013. Fractionation of the main components of corn stover by formic acid and enzymatic saccharification of solid residue. *Industrial crops and products*, **50**, 750–757.
- Yu, Q., Zhang, A., Wang, W., Chen, L., Bai, R., Zhuang, X., Wang, Q., Wang, Z., Yuan, Z. 2018. Deep eutectic solvents from hemicellulose-derived acids for the cellulosic ethanol refining of *Akebia*’herbal residues. *Bioresource technology*, **247**, 705–710.
- Zeng, J., Helms, G.L., Gao, X., Chen, S. 2013. Quantification of Wheat Straw Lignin Structure by Comprehensive NMR Analysis. *Journal of Agricultural and Food Chemistry*, **61**(46), 10848–10857.
- Zhai, Q., Long, F., Jiang, X., Hse, C.-y., Jiang, J., Xu, J. 2020. Facile and rapid fractionation of bamboo wood with a p-toluenesulfonic acid-based three-constituent deep eutectic solvent. *Industrial Crops and Products*, **158**, 113018.

- Zhang, C., Shen, X., Jin, Y., Cheng, J., Cai, C., Wang, F. 2023. Catalytic strategies and mechanism analysis orbiting the center of critical intermediates in lignin depolymerization. *Chemical Reviews*, **123**(8), 4510–4601.
- Zhang, J., Zhang, W., Cai, Z., Zhang, J., Guan, D., Ji, D., Gao, W. 2022a. Effect of Ammonia Fiber Expansion Combined with NaOH Pretreatment on the Resource Efficiency of Herbaceous and Woody Lignocellulosic Biomass. *ACS Omega*, **7**(22), 18761–18769.
- Zhang, K., Li, H., Xiao, L.-P., Wang, B., Sun, R.-C., Song, G. 2019. Sequential utilization of bamboo biomass through reductive catalytic fractionation of lignin. *Bioresource technology*, **285**, 121335.
- Zhang, M., Tian, R., Han, H., Wu, K., Wang, B., Liu, Y., Zhu, Y., Lu, H., Liang, B. 2022b. Preparation strategy and stability of deep eutectic solvents: A case study based on choline chloride-carboxylic acid. *Journal of Cleaner Production*, **345**, 131028.
- Zhang, T., Kumar, R., Wyman, C.E. 2013. Sugar yields from dilute oxalic acid pretreatment of maple wood compared to those with other dilute acids and hot water. *Carbohydrate Polymers*, **92**(1), 334–344.
- Zhao, W., Debnath, D., Gautam, I., Fernando, L.D., Wang, T. 2024. Charting the solid-state NMR signals of polysaccharides: A database-driven roadmap. *Magnetic Resonance in Chemistry*, **62**(4), 298–309.
- Zhao, W., Zhu, B., Davis, S.J., Ciais, P., Hong, C., Liu, Z., Gentine, P. 2023. Reliance on fossil fuels increases during extreme temperature events in the continental United States. *Communications Earth & Environment*, **4**(1), 473.
- Zhao, X., Li, S., Wu, R., Liu, D. 2017. Organosolv fractionating pre-treatment of lignocellulosic biomass for efficient enzymatic saccharification: chemistry, kinetics, and substrate structures. *Biofuels, Bioproducts and Biorefining*, **11**(3), 567–590.
- Zhao, Z.-M., Meng, X., Scheidemantle, B., Pu, Y., Liu, Z.-H., Li, B.-Z., Wyman, C.E., Cai, C.M., Ragauskas, A.J. 2022. Cosolvent enhanced lignocellulosic fractionation tailoring lignin chemistry and enhancing lignin bioconversion. *Bioresource Technology*, **347**, 126367.
- Zheng, Q., Zhou, T., Wang, Y., Cao, X., Wu, S., Zhao, M., Wang, H., Xu, M., Zheng, B., Zheng, J., Guan, X. 2018. Pretreatment of wheat straw leads to structural changes and improved enzymatic hydrolysis. *Scientific Reports*, **8**(1), 1321.
- Zhong, R., Ye, Z.-H. 2015. Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. *Plant and cell physiology*, **56**(2), 195–214.
- Zhu, W., Houtman, C.J., Zhu, J., Gleisner, R., Chen, K. 2012. Quantitative predictions of bioconversion of aspen by dilute acid and SPORL pretreatments using a unified combined hydrolysis factor (CHF). *Process Biochemistry*, **47**(5), 785–791.