

Lignin Structure Dynamics: Advanced Real-Time Molecular Sensing Strategies

Chi Ho Lee^{a,b,†}, Juhyeon Kim^{a,b,†}, Jiae Ryu^{c,†}, Wangyun Won^d, Chang Geun Yoo^c, Joseph Sang-Il Kwon^{a,b,*}

^a Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77845 USA

^b Texas A&M Energy Institute, Texas A&M University, College Station, TX 77845 USA

^c Department of Chemical Engineering, State University of New York College of Environmental Science and Forestry, Syracuse, NY 13210, USA

^d Department of Chemical and Biological Engineering, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea

[†]These authors have contributed equally.

Corresponding authors: kwon075@tamu.edu (Joseph Sang-Il Kwon)

Abstract

Lignin, with its abundant and high energy density, offers potential for sustainable biofuels and chemicals. However, current research faces two primary challenges: limited experimental sampling frequency, impeding the real-time analysis and a lack of theoretical activation energies to elucidate thermodynamic and kinetic properties associated with structural changes under various processing conditions including reaction temperature (343 to 363 K), time (15 to 30 min), and chip size (1.0 to 5.0 mm). To address these challenges, we introduce the "real-time molecular-level sensing approach.", which considers delignification, redeposition, and de/repolymerization over time and various scales, leveraging sparse experimental data and theoretical activation energies. Our approach combines density

functional theory, ab-initio molecular dynamics, and kinetic Monte Carlo simulations to deeply explore dynamic nature of lignin. Specifically, we gain insights into lignin content within biomass, molecular weight distribution, and monolignol ratio, improving lignin's utilization during processing. This work establishes a new standard for comprehensive lignin analysis, fostering efficient lignin utilization by tracking their properties at the molecular level in the real-time manner.

Keywords: Lignin, Real-time, Molecular-level sensing approach, DFT, AIMD, and kMC

1. Introduction

Recently, lignin has emerged as a renewable and sustainable biopolymer with immense potential in the production of value-added chemicals and alternative fuels, considering its abundance, aromatic nature, and high energy content.^[1, 2] This positions lignin as a robust alternative to petrochemical resources, particularly given the current energy crisis and environmental concerns.^[3] Despite its great potential, lignin is still underutilized due to its poor processability.^[4, 5] For example, a substantial amount of lignin currently produced in pulp and paper industry is directly combusted for generating energy rather than being utilized in other commercial applications.^[6, 7] To improve the separation efficiency and quality of lignin, various biomass processing solvents have been investigated. Hydrotropic solvents are solvents containing hydrophilic and hydrophobic molecular groups in a single structure and show an effective lignin fractionation from biomass due to their amphiphilic structures.^[8] In our recent study, phenol-4-sulfonic acid (PSA) was successfully applied to separate lignin from woody biomass under mild conditions.^[9] However, as lignin has a complicated and heterogeneous structure formed by random polymerization of its phenolic units, its structure and behavior have not been thoroughly understood in its extraction processes.^[10]

During fractionation, three key criteria – remaining lignin content in biomass, molecular weight distribution, and monolignol ratio of the extracted lignin chains – can be assessed.^[11, 12] However, existing lignin research encounters fundamental challenges in predicting such properties, all of which are limited by the following reasons:

1. Current experimental research has a clear limitation of sampling frequency, restricting the real-time analysis of the dynamic nature of lignin structures between sparsely sampled data points.
2. Current theoretical models exhibit the absence of key parameters (e.g., activation energy barrier), which constrains our ability to elucidate the thermodynamic and kinetic properties associated with temperature-induced changes in molecular structures.

These challenges significantly limit our understanding of the dynamic behavior of

de/repolymerization processes that inevitably change with variations in the structural form of the lignin system between sparsely sampled data points. As a prime example, it is essential to concurrently evaluate different monomers, such as syringyl (S) and guaiacyl (G), which are predominant in lignin. While lignin generally comprises three types of monomers, including the less abundant hydroxyphenyl (H) monomer, most research primarily centers on the S and G monomers due to their significantly higher prevalence.^[13, 14] However, it is important to note that with just these two components, researchers can have the capability to elucidate the multitude of monomeric combinations within a lignin structure, enabling them to describe the intricate and complex nature of lignin composition.^[15-17] Since the kinetics of de/repolymerization kinetics depends on such monomeric configuration, predicting bond scission and formation for every possible structure becomes an intricate undertaking.^[18-20] To accomplish this, theoretical parameters are necessary, and their absence prevents systematic kinetic analysis, thereby posing significant obstacles to real-time monitoring. Furthermore, the unresolved challenge persists among researchers, stemming from the complexities of effectively integrating diverse theoretical methodologies that allow us to do the systematic kinetic analysis.^[19, 20]

To briefly mention the previous research efforts on the valorization of lignocellulosic materials, the kinetic Monte Carlo (kMC) algorithm has extensively been applied^[21-24] which is further described in the following sections. Due to the stochastic nature of the biopolymer reactions, kMC is considered as a promising tool for simulating lignin over several decades. For instance, kMC has been utilized to provide insights into the lignin structure^[25-27], average molecular weights^[28], and product yields^[28-31]. Nevertheless, there is a noticeable absence of dedicated reports focusing on biomass fractionation and, more specifically, lignin depolymerization that elucidates the complete molecular weight distribution and their monomeric ratios. Since the foundation of kMC simulation lies in the reaction kinetics, such knowledge gaps are attributed to the lack of the kinetic parameters for the lignin depolymerization reactions. Hence, it is worth noticing that the introduction of theoretical parameters and a new model based on them would eliminate such issues, significantly enhancing our understanding of lignin reactions. To achieve this goal, we present a new theoretical model named

“real-time molecular-level sensing approach”. This model leverages sampled data over discrete time instants to systematically investigate the dynamics of the overall lignin reactions, encompassing delignification, redeposition, and de/repolymerization processes at the molecular level (**Fig. 1**).

In this regard, the systematic construction of our model necessitates extensive multiscale thermodynamic and kinetic studies, ranging from the atomic to the macroscopic level. Specifically, we employed a combination of density functional theory (DFT), ab-initio molecular dynamics (AIMD), and kMC simulations, constituting a comprehensive framework for assessing various lignin properties. First, DFT simulations were utilized to gain a deeper understanding of the mechanistic aspects of the de/repolymerization processes by analyzing thermodynamic factors (i.e., structural conformation of lignin and activation energy barrier).^[32-35] While investigating these aspects, it is essential to consider the temperature effect on the structural changes in the lignin macromolecules. Specifically, temperature-induced structural changes can result in the repositioning of active sites where radical interactions occur among lignin structures composed of S and G monomers, resulting in substantial alterations to binding energies and activation energy barriers. Hence, we implemented AIMD simulations to discover the intricate interplay between temperature and lignin structure.^[32] Further, we integrated this with DFT method by calculating activation energy barriers required for de/repolymerization processes based on the changed lignin structural conformations with different temperatures.

Subsequently, we deployed a kMC simulation to study three key aspects of the lignin properties. In the kMC simulation, each lignin reaction is executed^[36] based on the DFT-AIMD results and the system configuration, tracking the system status in real-time. Considering the delignification and de/repolymerization on distinct time and length scales, we utilized the multiscale simulation framework. Here, we systematically categorized reactions into macro and microscopic reactions and executed them simultaneously. Through this simulation, we unveiled insights into the residual lignin content in biomass, lignin molecular weight distribution, and the S/G ratio.

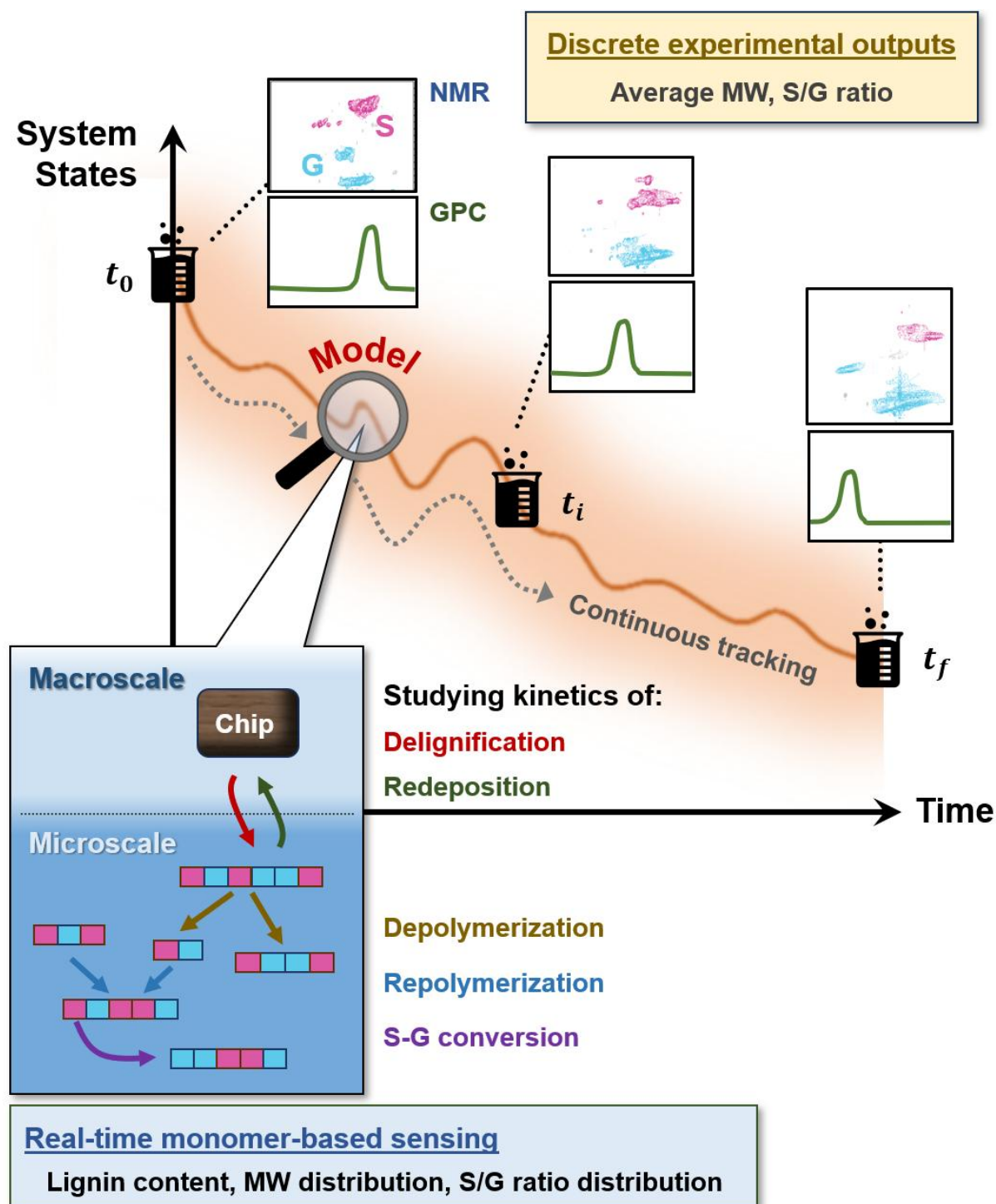


Fig. 1. Real-time molecular-level sensing approach constructed by integrating experimental sampling points in a discrete-time manner and theoretical data that spans from the atomic to the macroscopic level.

This model has been validated against the experimental outcomes, showcasing that our multifaceted approach significantly aids in elucidating the lignin reaction dynamics. That is, by synergistically integrating DFT-AIMD-kMC simulation framework and experimental datasets, we successfully developed a real-time molecular-level sensing approach that transcends conventional boundaries. Hence, we believe that our multifaceted exploration would set a new benchmark for an in-depth, accurate, and comprehensive simulation platform for effective lignin valorization in a computationally tractable manner. Additionally, the parameters we have detailed would be instrumental for various lignin-related works, especially those needing insights into how temperature and lignin structure impact reaction rates. This analysis emphasizes the utility of our framework, making it a valuable resource for understanding lignin-based processes across different conditions.

***Remark 1.** In this study, Aspen wood was selected for model validation because of its well-known structural and compositional configuration. This selection can also be justified by its commercial availability in the production of biofuels and biochemicals, having a fast-growing and even-textured anatomical feature.*^[37-43]

2. Material and methods

2.1. DFT and AIMD simulations

All DFT calculations were performed with the Vienna Ab initio Simulation Package (VASP 5.4.4.).^[44] We considered augmented wave (PAW) method^[45] with a generalized gradient approximation (GGA)^[46] to accurately describe chemisorption on the surface. The Monkhorst-Pack^[47] k-point grid (ref) was used, and maximum symmetry was applied to reduce the number of k-points in all calculations. Plane-wave cut-off energy of 500 eV was used. Lattice constants and internal atomic positions were optimized until the residual forces became less than 0.02 eV/Å. The vacuum space of a bulk unit cell in the x-, y-, and z-direction was set to 10 Å to avoid interactions between layers. Given the well-established fully-relaxed structures, we systematically calculated activation energy barriers via the climbing image nudged elastic band (CI-NEB) method^[48], and all calculated results are listed in Table 2 and Table S1. The ab initio molecular dynamics (AIMD) simulations were carried out in a NVT

ensemble with a Nose-Hoover heat bath method^[49, 50] for the fully relaxed geometry over three temperatures (300, 350, and 400 K) to understand temperature effect on the structural changes of lignin molecules. A gamma-centered 1x1x1 k-point mesh was utilized, and the total simulation time of each AIMD simulation was determined by monitoring energy fluctuations, ensuring they remained within 0.5 eV (with approximate time steps of 6 to 8 picoseconds (ps)).

2.2. Multiscale kMC simulations

While the DFT-AIMD simulation provided intricate thermodynamic insights into the lignin reactions, we integrated the multiscale kMC simulation framework for kinetic analysis. Foremost, biorefineries aim to extract lignin from bulk biomass to yield high-quality cellulosic fibers.^[51] Additionally, the simultaneous tailor of properties of the liberated lignin macromolecule is of utmost importance for its further utilization.^[52, 53] These dual goals necessitate simultaneous consideration, a task aptly handled by the multiscale simulation framework plays a role. Given that the micro and macroscopic reactions discussed below operate on different length and time scales, we designed our simulation approach in a bilayer form, as detailed further.^[12]

The first layer focuses on simulating delignification, a macroscopic reaction that detaches lignin from the bulk biomass and subsequently causes dissolved lignin to redeposit. Within this layer, we deterministically solved the first-order reversible reaction to compute the remaining lignin content in the biomass. Comprehensive discussions on the mass and energy balance equations are provided in the Supporting information.

In addition, the second layer delves into the microscopic dynamics of dissolved lignin. At any given moment, these lignin chains undergo de/repolymerization and conversion between monolignols within lignin chains. Furthermore, vast numbers of lignin molecules undergo de/repolymerization simultaneously, each at their unique reaction rate. These events are considered stochastic phenomena, and thus, the kMC algorithm simulates the reactions probabilistically, selecting a specific lignin chain and executing a microscopic reaction based on the continually updated reaction rates.

The reaction rates were determined by calculating activation energies for each respective reaction.

Within the kMC simulation, considerations were made for the influence of reaction temperature and the monomeric composition on the lignin structures. This allows us to deduce the precise activation energy values for lignin de/repolymerization. The detailed working principle of the kMC, along with the determination of activation energy, are elucidated in the Supporting Information. Additionally, the schematic diagram of this layer-wise operation is shown in Fig. S4.

2.3. Chemicals and preparation of Aspen wood chips

The air-dried Aspen wood chips were provided by SUNY ESF. The wood chips were processed to pre-determined thicknesses of 1.0, 3.0, and 5.0 mm for all directions using a Wiley mill and manual cutting. The prepared chips were stored in the cold room until the experiment started. The chips were placed in deionized water at room temperature for 48 hours, and the moisture content of the chips was measured before the experiment.

The phenol-4-sulfonic acid (PSA, 85 wt.%) was purchased from TCI America, Inc. (United States). Acetone, 72% sulfuric acid, ethanol, sodium hydroxide, acetic anhydride, pyridine, and tetrahydrofuran (THF) used in this study, were purchased from VWR International LLC. (United States). Dimethyl sulfoxide-*d*₆ (DMSO-*d*₆) was purchased from CIL (Cambridge Isotope Laboratories Inc., United States).

2.4. PSA pretreatment of Aspen wood chips

The pre-saturated Aspen chips were loaded into a 40 mL glass vial. The wood chip loadings were 2.0 ± 0.3 g of dried biomass weight. To adjust the concentration of PSA in the reactor to 72 wt.%, the measured moisture contents of the pre-saturated Aspen chips were considered. The reaction was conducted at 343 – 363 K for 15 – 30 min. Note that the reaction temperature beyond 363 K was not considered to avoid pressure generation from the water in the PSA solvent during the pretreatment. During the reaction, the biomass was manually disintegrated with a glass rod every 5 min. When the reaction ended, 50% acetone solution was loaded to quench the reaction. The solid residues and supernatant were separated and recovered by centrifugation. The collected solid samples were further washed until the pH of the filtrate became neutral. Also, the results at below 343 K were not included

to validate the model because PSA was not fully reacted with Aspen chips in sizes of 3.0 and 5.0 mm at this temperature.

2.5. Lignin recovery

The dissolved lignin in the liquid fraction was recovered using Amicon stirred ultrafiltration cell (UFSC20001, Amicon Corporation) with a 1 kDa regenerated cellulose membrane disc. The liquid sample was diluted with 50% ethanol solution. The 1% sodium hydroxide in 66% ethanol-water mixture (v/v) was added to adjust the pH of overall solution to 3.0-3.5. The pressure of the stirred cell was controlled to 20 psi with nitrogen gas. To avoid creating a filter cake, the liquid sample was constantly agitated at 300 rpm. The residues in the gel state on the membrane disc were recovered and air-dried.

2.6. Klason lignin analysis

The Klason lignin contents of untreated and pretreated Aspen chips were analyzed according to the NREL procedure.^[54] The wood chip samples were screened to 14-mesh prior to the analysis. Two-step acid hydrolysis was conducted to separate acid-insoluble residues from carbohydrate fractions. In brief, biomass samples were treated with 72% sulfuric acid at 303 K for 1 hour. The sulfuric acid concentration in the mixture was adjusted to 4% by adding deionized and then autoclaved at 394 K for 1 hour. The obtained acid-insoluble residue was filtered and neutralized by washing with deionized water. This fraction was oven-dried to measure the weight and then calcinated at 848 K using a muffle furnace for 24 hours. The Klason lignin content was calculated by excluding the remaining ash content after the calcination from the measured acid-insoluble residue weight.

2.7. Molecular weight analysis

To measure the molecular weights for the recovered lignin described in Section 2.5, the lignin was acetylated and analyzed using an Agilent Gel Permeation Chromatography (GPC) SECurity 1200 system equipped with Waters Styragel columns (Waters Corporation, Milford, MA) and a UV detector at 270 nm. In brief, the lignin was acetylated in the mixture of pyridine and acetic anhydride (1:1 v/v) at room temperature for 48 hours. The acetylated lignin was recovered using a rotary evaporator and the residual solvents in the lignin were further washed with ethanol and removed with additional rotary

evaporation. For the GPC analysis, tetrahydrofuran was used as a mobile phase. The weight-average (M_w) and number-average molecular weights (M_n) were obtained by Breeze software v.2.

2.8. S/G ratio analysis

The composition of lignin was analyzed by semi-quantitative two-dimensional (2D) heteronuclear single quantum coherence (HSQC) nuclear magnetic resonance (NMR). For the NMR analysis, the lignin sample was dissolved in DMSO- d_6 . The NMR analysis was performed using an 800 MHz NMR (Bruker AVANCE III HD) equipped with a TCI cryoprobe at 298 K. The ^1H - ^{13}C HSQC spectra were obtained by hsqcetgpsp pulse program with specified acquisition parameters: 32 scans, 1s of relaxation delay, with F1 (^{13}C) and F2 (^1H) dimensions at 220 and 12 ppm of low field limit of spectrum, with 2048 and 640 for size of fid.

3. Results and discussion

To realize the real-time molecular-level sensing approach, we initially implemented a comprehensive strategy (Fig. 2) aimed at establishing essential theoretical parameter (i.e., activation energy barrier). This parameter plays a pivotal role in facilitating a systematic kinetic analysis of overall lignin reactions, including delignification, redeposition, and de/repolymerizations. This strategy involves the following sequential steps.

1. DFT is used to pre-optimize reasonable lignin aromatic units for 5 C-radical and 3 O-radical structures, sourced from the literature, which includes bond dissociation energy derived from the hardwood lignin structure.
2. AIMD is used to investigate the impact of temperature (300, 350, and 400 K) on the structural changes of 8 lignin aromatic units. This analysis plays a pivotal role in understanding de/repolymerization processes that are notably influenced by temperature.
3. Based on the AIMD results, DFT is again used to determine reaction rates for the de/repolymerization processes by calculating activation energy barriers for all conceivable pairings among different lignin structures based on the DFT method.

4. Integrating the DFT-derived activation energy barriers into a sophisticated kMC simulation for an all-encompassing kinetic assessment (e.g., M_n , M_w and S/G ratio changes with time).
5. Establishing a real-time molecular-level sensing approach through the combination of theoretical results and sparsely experimental sampling points.

3.1. Thermodynamic analysis of lignin system

3.1.1. Selection of initial lignin structures via literature review

Considering inherent intricacies and atomic configurations of the lignin system, a precise structural framework is vital for a cogent theoretical interpretation. Hence, we first considered the hardwood lignin structure ($C_{228}H_{270}O_{94}$), which has been corroborated through experimental NMR analysis.^[55, 56] Table 1 shows the predominant structural units, bond linkages, and their respective percentages for the hardwood lignin. Notably, this model comprises 14 S units and 6 G units, and as a characteristic of hardwood lignin, this predominantly features the β -O-4 bond linkage, accounting for 84.21% of the total.^[55, 56] While there exists an additional aromatic units (SP) and other bond linkages (β -5, β - β , and 4-O-5), their relative proportions are minimal. Therefore, in order to identify the key factors that determine the overall lignin reaction and to newly propose a theoretical framework based on them, we focused solely on the S and G monomers and the β -O-4 linkage.

Expanding upon this aromatic unit and bond information, we tried to identify reliable initial lignin structures by focusing on the de/repolymerization processes that involve various radical types based on C- and O-radicals. So, we anchored our investigations on the hardwood lignin ($C_{228}H_{270}O_{94}$), which had already been examined for bond dissociation trends. Using the lowest dissociation energy reported in this study, we put forward 5 C-radical (CR) and 3 O-radical (OR) structures, indicating the most favorable products, and optimized them. Specifically, their elemental compositions (Figs. 2a and 2b) are as follows: CR-1 ($C_{11}H_{15}O_5$), CR-2 ($C_{22}H_{29}O_{10}$), CR-3 ($C_{33}H_{43}O_{15}$), CR-4 ($C_{44}H_{57}O_{20}$), CR-5 ($C_{54}H_{69}O_{24}$), OR-1 ($C_{10}H_{11}O_3$), OR-2 ($C_{20}H_{23}O_7$), and OR-3 ($C_{10}H_{11}O_3$).

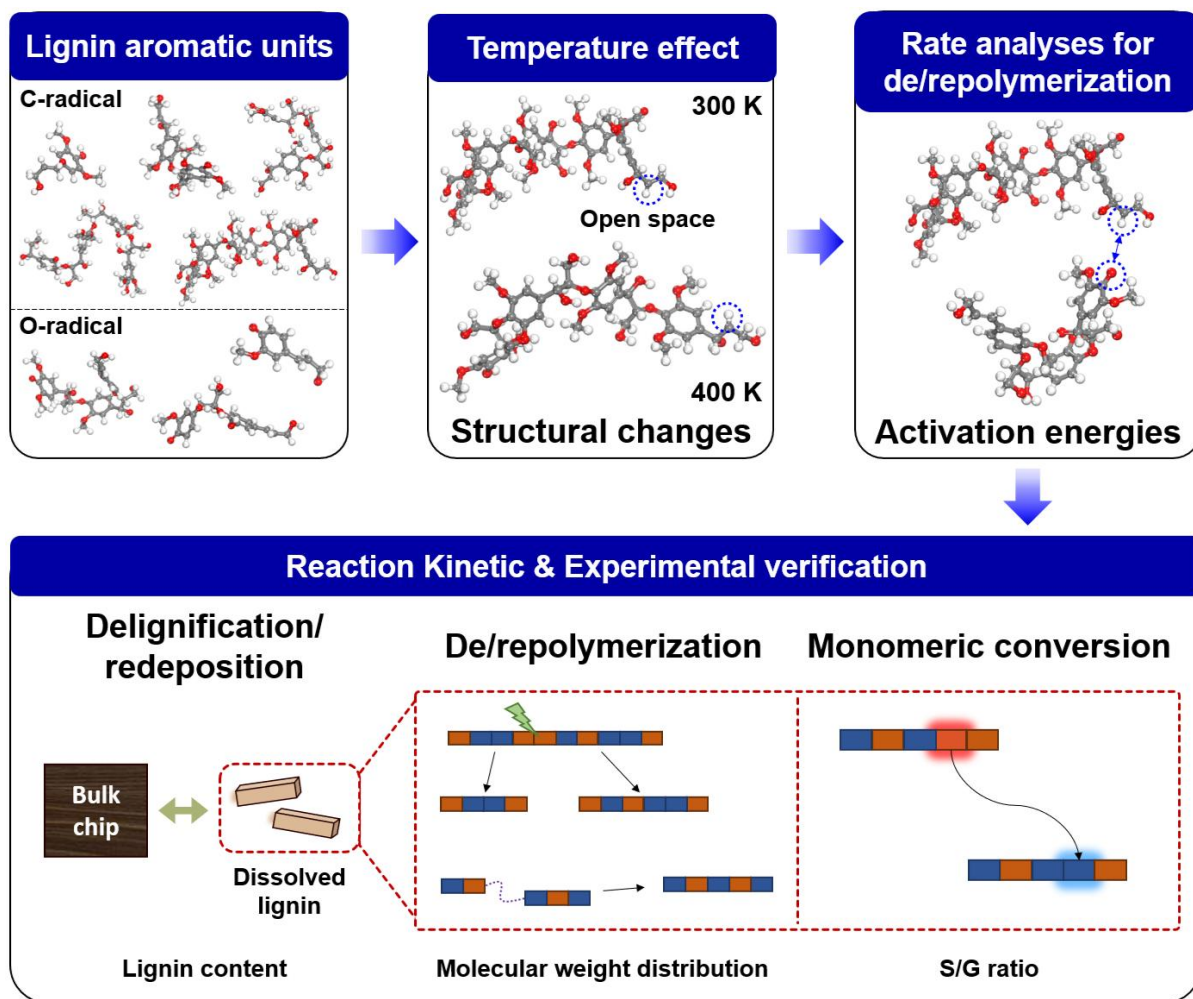


Fig. 2. Multi-step DFT, AIMD, and kMC framework for investigating thermodynamic and kinetic analysis.

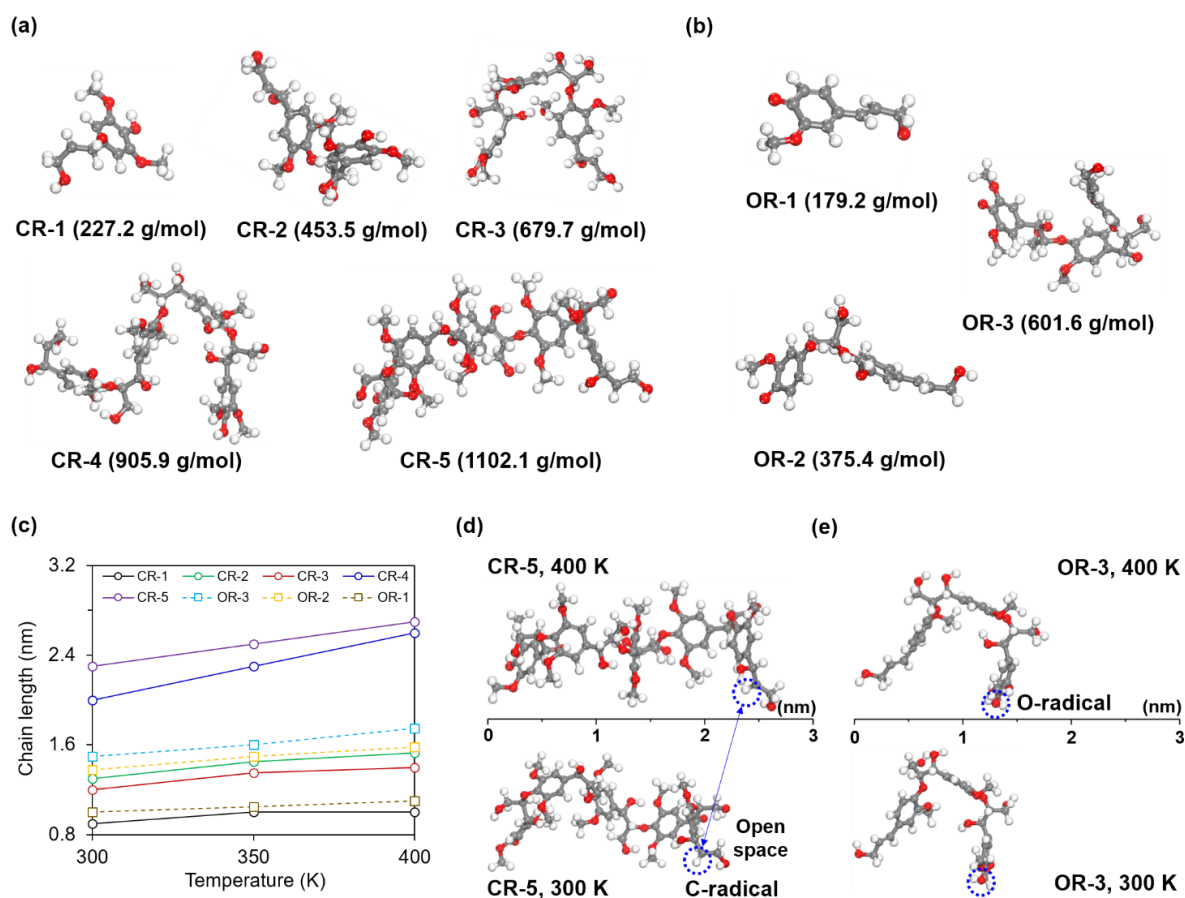


Fig. 3. Optimized (a) 5 C-radical (CR) and (b) 3 O-radical (OR) lignin structures. (c) Changes in chain length of 8 lignin structures depending on temperature (300, 350, and 400 K). Changes in the structure of (d) CR-5 and (e) OR-3 lignins depending on temperature (300 and 400 K).

Table 1. Major Features of the hardwood lignin structure ($C_{228}H_{270}O_{94}$) Taken from a previous study by Stewart et al.^[56]

Feature	Notation	Name	Relative amount (%)
Aromatic unit	S	Syringyl	63.63
	G	Guaiacyl	27.27
	SP	Sinapyl <i>p</i> -hydroxy benzoate	9.1
Bond linkage	β -O-4	β -aryl ether	84.21
	β -5	Phenyl coumaran	5.26
	β - β	Resinol	5.26
	4-O-5	Biphenyl ether	5.26

3.1.2. Exploring the effects of temperature on lignin structures via AIMD

Based on the optimized 8 initial structures, our focus is to examine the influence of temperature on structural changes because temperature condition profoundly affects the thermodynamic and kinetic features of the overall lignin reactions. Specifically, structural changes with temperature can relocate the active sites where the interactions between C- and O-radicals occur, leading to significant variations in de/repolymerization processes. With this in mind, we systematically analyzed changes across the 5 CR and 3 OR structures across three temperature ranges (300K, 350K, and 400K) based on the AIMD simulation. The data presented in Fig. 3c underscores that with increasing temperature, all structures clearly undergo structural elongation, leading to an expansion in chain length. Notably, larger structures demonstrate more pronounced changes. A particularly compelling observation is the shift in the position of the radical site, which plays a pivotal role in reactions between lignin aromatic units, as the lignin chain expands. As shown in Fig. 3d, at 300K, the CR site optimally situates itself in an open space, promoting easier reactions. Yet, at 400K, the rotation of C chain appears to obstruct the space essential for the reactivity of CR. On the contrary, Fig. 3e illustrates that OR structures exhibit smaller structural sizes, resulting in a relatively minor temperature-induced impact on structural changes compared to CR structures. It means that OR structures offer a more favorable environment for facilitating polymerization reactions in contrast to CR structures. From these findings, it can be inferred that as the temperature increases from 300 K to 400 K, the repolymerization process is likely to become more challenging due to the closure of reaction spaces induced by structural changes in CR structures, in contrast to depolymerization, which primarily involves bond scission.

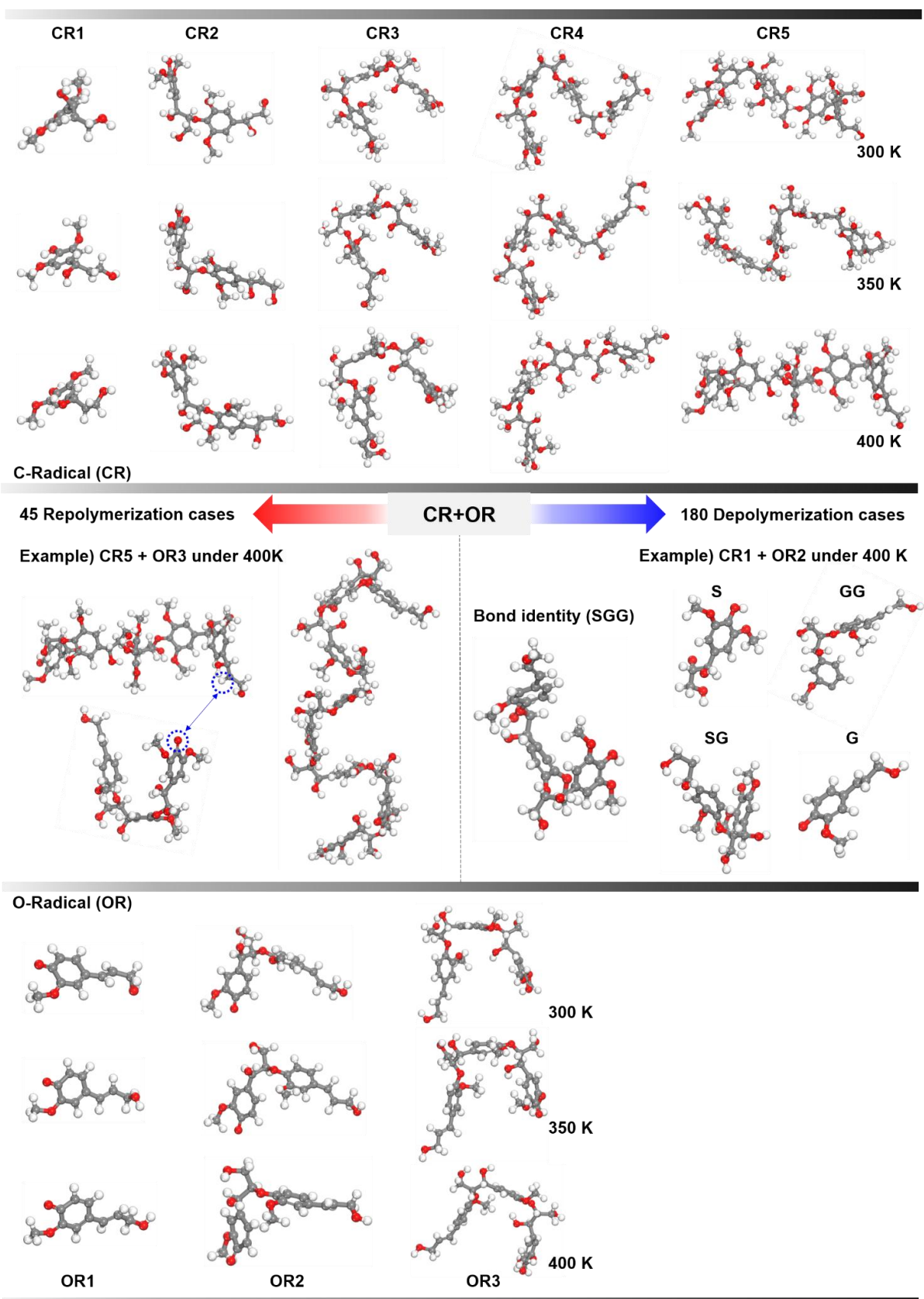


Fig. 4. Optimized 5 C-radical (CR) and 3 O-radical (OR) lignin structures at three distinct temperatures (300, 350, and 400 K) and schematic that illustrates the de/repolymerization processes considering all

possible combinations between CR and OR structures. the total simulation time of each AIMD simulation was determined by monitoring energy fluctuations, ensuring they remained within 0.5 eV (with approximate time steps of 6 to 8 ps).

3.2. Kinetic analysis of lignin system

3.2.1. Calculation of activation energy barriers for lignin de/repolymerization via DFT

kMC simulation can systematically explore the potential structural changes and associated reaction kinetics in response to increasing temperature. It not only bridges the micro and macroscopic levels but also provides a more holistic understanding within a realistic environment.^[57-59] However, a prerequisite for this is the prior determination of the activation energy barrier that accurately describes our specific system.^[60] Hence, as our next step, we calculated activation energy barriers for the de/repolymerization processes by considering possible combinations between lignin structures that could exist at different

Table 2. Calculated activation energy barriers for the de/repolymerization processes at three distinct temperatures (300, 350, and 400 K) considering all possible combinations between CR and OR structures.

Temperature (K)	Combination	Identity	E_{rep} (eV)	Identity	E_{dep} (eV)
300	CR1+OR3	S - SGG	0.78	S - SGG	1.43
	CR1+OR2	S - GG	0.98	S - GG	1.82
	CR1+OR1	S - G	1.17	S - G	0.90
	CR2+OR3	SS - SGG	0.60	SS - SGG	1.36
	CR2+OR2	SS - GG	0.78	SS - GG	1.09
	CR2+OR1	SS - G	0.86	SS - G	0.80
	CR3+OR3	SSS - SGG	0.67	SSS - SGG	1.81
	CR3+OR2	SSS - GG	0.66	SSS - GG	1.98
	CR3+OR1	SSS - G	0.76	SSS - G	1.80
	CR4+OR3	SSSS - SGG	0.72	SSSS - SGG	1.28
	CR4+OR2	SSSS - GG	0.73	SSSS - GG	2.04
	CR4+OR1	SSSS - G	0.57	SSSS - G	1.40
	CR5+OR3	SSSSG - SGG	0.79	SSSSG - SGG	1.96
	CR5+OR2	SSSSG - GG	0.68	SSSSG - GG	1.94
	CR5+OR1	SSSSG - G	0.75	SSSSG - G	2.93

350	CR1+OR3	S - SGG	0.83	S - SGG	0.83
	CR1+OR2	S - GG	1.13	SG - G	1.95
	CR1+OR1	S - G	1.19	S - G	1.48
	CR2+OR3	SS - SGG	0.64	SS - SGG	1.40
	CR2+OR2	SS - GG	0.88	SS - GG	1.78
	CR2+OR1	SS - G	0.98	S - SG	2.16
	CR3+OR3	SSS - SGG	0.78	SSS - SGG	1.45
	CR3+OR2	SSS - GG	0.81	SSS - GG	1.96
	CR3+OR1	SSS - G	0.89	SSS - G	1.79
	CR4+OR3	SSSS - SGG	0.84	SSSS - SGG	1.62
	CR4+OR2	SSSS - GG	0.84	SSSS - GG	1.91
	CR4+OR1	SSSS - G	0.67	SSSS - G	1.44
	CR5+OR3	SSSSG - SGG	0.88	SSSSG - SGG	1.29
	CR5+OR2	SSSSG - GG	0.82	SSSSG - GG	2.27
	CR5+OR1	SSSSG - G	0.82	S - SSSGG	1.74
400	CR1+OR3	S - SGG	0.95	S - SGG	1.38
	CR1+OR2	S - GG	1.18	SG - G	1.50
	CR1+OR1	S - G	1.32	S - G	1.92
	CR2+OR3	SS - SGG	0.78	SS - SGG	0.79
	CR2+OR2	SS - GG	1.00	SS - GG	1.02
	CR2+OR1	SS - G	1.12	S - SG	1.77
	CR3+OR3	SSS - SGG	0.84	SSS - SGG	1.68
	CR3+OR2	SSS - GG	0.92	SSS - GG	0.57
	CR3+OR1	SSS - G	0.95	SSS - G	2.00
	CR4+OR3	SSSS - SGG	0.89	SSSS - SGG	1.93
	CR4+OR2	SSSS - GG	0.87	SSSS - GG	1.20
	CR4+OR1	SSSS - G	0.74	SSSS - G	1.73
	CR5+OR3	SSSSG - SGG	1.02	SSSSG - SGG	1.37
	CR5+OR2	SSSSG - GG	0.98	SSSSG - GG	0.88
	CR5+OR1	SSSSG - G	0.94	S - SSSGG	1.67

320

321 temperatures. Specifically, for the repolymerization process, we considered combinations between CR
322 and OR. However, for depolymerization process, since structures resulting from repolymerization could
323 undergo bond dissociation at sites other than the existing CR and OR points, we considered all potential
324 bond dissociation cases by categorizing bond identity based on S and G units. Thus, we calculated
325 activation energy barriers of 45 combinations for repolymerization and 180 combinations for
326 depolymerization based on CR and OR structures (Fig. 4). All calculated results are listed in Tables 2

and S1. Building on these energies, we developed a high-fidelity kMC model that systematically integrates both macro and microscopic perspectives, surpassing the atomic scale.

3.2.2. Tracking the molecular weight and S/G ratio of lignins via kMC

The activation energy barriers calculated in the previous section are used in microscopic layer, which comprises four reactions, de/repolymerization, demethoxylation, and null events. They can affect the microscopic configuration of the system, such as the molecular weight distribution and the S/G ratio. In the microscopic layer, the kMC algorithm computes the reaction rates of those four (Eqs. S6-S8) to select an event to be executed at a given moment. Each reaction is explained below in the order they appeared above.

For the depolymerization and repolymerization rate calculations, one or two dissolved lignin chains are arbitrarily selected, respectively. For the selected chains, finding activation energies is crucial to perform de/repolymerization. However, for such a lignin chain whose molecular weight is up to tens of kDa, it can have a number of combinations of S and G units. Hence, instead of exploring every possible combination, we employed particular strategies for de/repolymerization reactions as follows. For depolymerization, the long chain is regarded as a connection of the unit blocks as illustrated in Fig. S2,

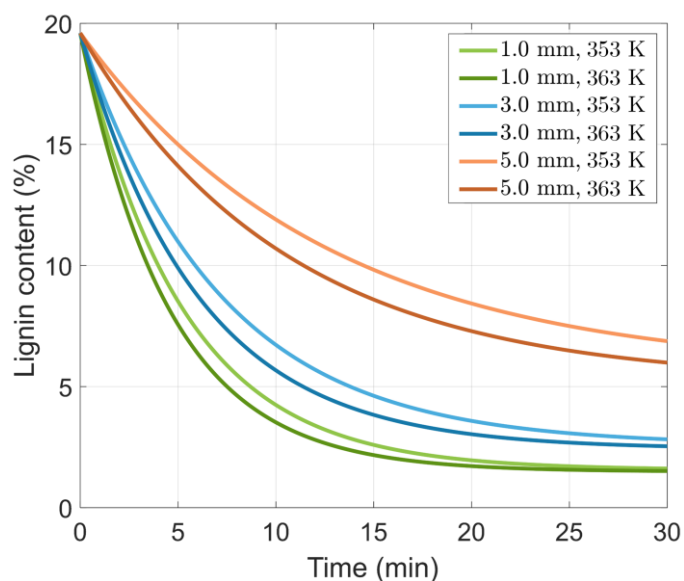


Fig. 5. Remaining lignin content calculated in the macroscopic layer.

and the activation energies were found in Table S1 for each block. Repolymerization becomes simpler since we need to consider only both ends of the connecting chains (Fig. S3). The activation energies required to make new bonds were calculated by DFT calculation, using the eight initial structures exhibited in Fig. 3. The activation energies were calculated over the temperatures and the molecular weights of the combined chains, based on the bond identity (i.e., S-S, S-G, and G-G) (Table 2). Note that repolymerization is considered as lignin condensation. Therefore, a bond created by repolymerization is marked and considered not to be broken again. The detailed procedure for determining the de/repolymerization kinetics is described in Supporting Information.

The demethoxylation converts the S units in the dissolved chains to G unit irreversibly. This primarily influences the S/G ratio and slightly changes the molecular weight of the chain. Note that its reaction rate strongly depends on the S content of the selected chain. The null event stands for any events that do not influence the lignin chains, and it is considered as a zeroth-order reaction. Note that the null event only advances the microscopic simulation time, while it does not cause any change in the system. Additional details for these reactions are provided in Supporting Information.

3.2.3. Investigating the change in the remaining lignin content and temperature via mass/energy balance

In the macroscopic layer, the lignin dissolves out from the bulk biomass (i.e., delignification), and the dissolved lignin attaches back to the biomass (i.e., redeposition). These were governed by the mass and energy balance equations shown in Eqs. S1-S4. As a result of delignification/redeposition, the amounts of remaining lignin on the biomass, and lignin content were readily calculated, as displayed in Fig. 5. Note that this layer based on mass/energy balance can be further utilized for a more in-depth exploration of lignin reactions from a microscopic perspective. Specifically, the mass balance determines how many lignin chains will be dissolved out or redeposited. At the same time, the energy balance updates the temperature of the chips and liquor phases. The lignin chains dissolved in this layer underwent the microscopic reactions discussed in the previous section. To perform our simulation properly with DFT and AIMD calculations, we pre-determined the lignin configuration as follows. The experimental results provide the following initial conditions: the lignin content of the biomass is 0.196,

and the S/G ratio is 1.76. We assume that the lignin is a randomly composed linear chain containing S and G monolignols. Considering this, lignin chains of diverse composition were randomly generated using 5 CR and 3 OR structures, ensuring that their average ratio is around 1.76. Note that the molecular weight of the chains is set to ~13 kDa, according to the experimental observation. The lignin chains can vary in the degree of polymerization due to different monomeric compositions, but their molecular weights were kept the same.

To perform delignification, the macroscopic layer accumulates the dissolved amount of lignin for every time step. Then, the chip releases one lignin chain to the liquor phase once the accumulated mass reaches the threshold. When it comes to selecting chains to be dissolved, it is known that the chain with a higher S/G ratio can easily be delignified.^[61-63] In this regard, the dissolving chain was selected randomly but weighted based on their respective S/G ratios. The redeposition occurs oppositely. One of the dissolved chains is selected, and the redeposition event is executed when the accumulated mass stacks up to the mass of the selected chain.

3.2.4. Multiscale model validation

The multifaceted simulation framework was developed based on the discussion above, and its results were validated with the experimental results. We have examined the molecular weight distribution (Fig. 6) and the S/G ratio (Fig. 7) in six conditions for 1.0/3.0/5.0 mm biomass under the temperatures of 80/90 °C, with the experimental sampling points presented. Note that every reaction involved in the biorefinery processes, regardless of its length and time scale, is interrelated to affect the lignin content, molecular weight distribution and the S/G ratio.

As depicted in Fig. 6, delignification and redeposition of lignin can change its molecular weight distribution since those reactions can take and release the lignin chains of diverse lengths and compositions between both phases. Also, de/repolymerization of lignin chains are the primary factors for the rapid fluctuation in the average molecular weight as they directly affect the chain lengths in a short time. Such fluctuations are pronounced during the first few minutes, since we evaluate the molecular weight distribution using a limited number of the dissolved lignin chains. It can be further

supported by seeing the chip size effect. The larger chips release the lignin chains slowly, so the rapid changes of average molecular weight are observed for a longer time.

Notably, the average molecular weights decrease faster with the higher temperature. It suggests that the depolymerization reactions dominantly occur over the other microscopic reactions. This also agrees with the discussion in Section 3.1.2., that is, the repolymerization becomes challenging under higher temperatures as the temperature influences the lignin geometries to make the formulation of a chain-to-chain bonding unfavorable. The chip size effect can also be verified with the trend of the average molecular weight. In Fig. 6, it is seen that the average molecular weight reached such a lower bound under all conditions. Also, note that it was achieved earlier with the larger chip size. As previously discussed, (i) the microscopic reactions happen in a short time, (ii) depolymerization rate dominates the other reaction rates, and (iii) the larger chips slowly release the long lignin chains (13 kDa) into the liquor phase. Therefore, in the case of the larger chips, the average molecular weight initially drops faster due to rapid depolymerization (Fig. 6) and slow delignification (Fig. 5). When it comes to the S/G ratio (Fig. 7), a substantial initial upshift was observed in all conditions.

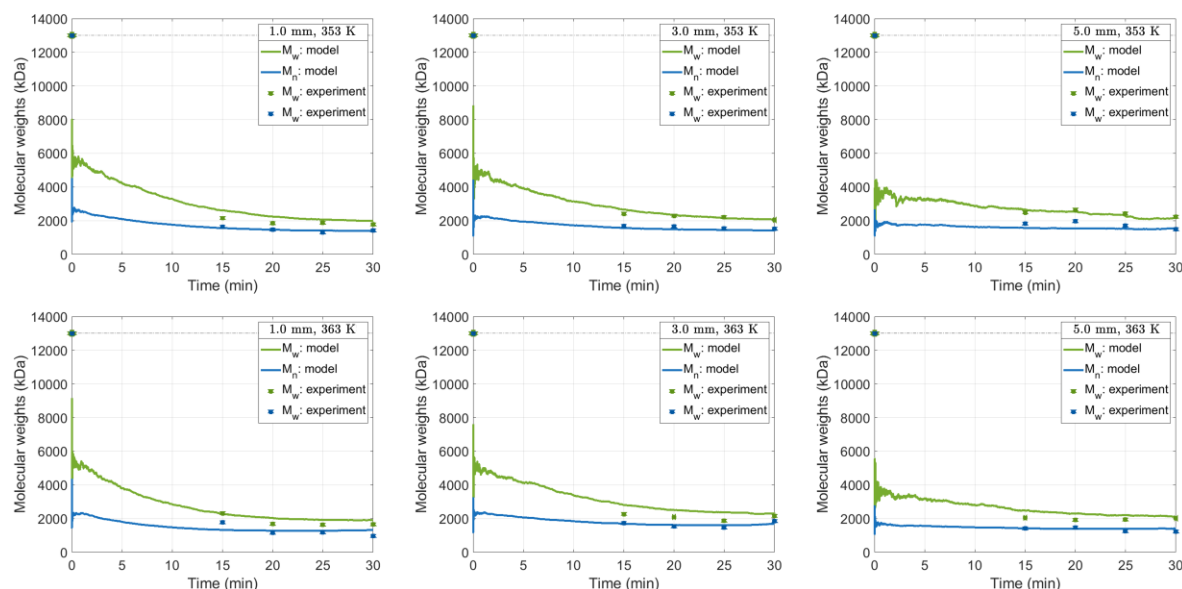


Fig. 6. Remaining lignin content calculated in the macroscopic layer.

This is attributed to the fast and favorable dissolution of the chains with a relatively high S/G ratio (Section 3.2.3). After the first few minutes, the S/G ratio gradually decreases due to the dissolution of the lignin chains with a low S/G ratio and the S-G conversion reactions. While the chip size has no visible impact on the S/G ratio, we can observe that the S/G ratio drops faster under higher temperatures, implying the high sensitivity of the S/G ratio to the system temperature. To demonstrate the prediction performance of our proposed model, the root-mean-square errors (RMSEs) are listed in Table 3. Note that the molecular weights of S/G monolignols are 227.2 and 179.2 Da, respectively. This suggests that the maximum prediction error in average molecular weight is less than the equivalent of 2 monomers per chain, demonstrating our model's capability of tracking the molecular weight accurately. The model also showed that the RMSE for the S/G ratio remains within 0.1712, which is considered negligible. These results imply that the model predictions are closely aligned with the experimental outcomes.

Table 3. The RMSEs evaluated for each reaction condition.

Chip Size (mm)	Temperature (K)	RMSE (Da)		RMSE (N/A)
		M_w	M_n	S/G ratio
1	353	333.0	74.41	0.1428
	363	236.3	288.3	0.1699
3	353	119.4	118.1	0.1438
	363	250.5	250.1	0.1103
5	353	132.2	270.6	0.1551
	363	320.9	126.5	0.1712

Along with the real-time sensing feature, our model can also track the lignin in detail. The multiscale reactions alter the length and composition of the dissolved lignin chains during the simulation. For example, Fig. S5 gives an insight into how the average molecular weight is evaluated. We tracked the molecular weight distribution of the dissolved lignin chains and updated the average molecular weights based on this result for every time step. In addition to the length information, our simulation also kept an eye on their composition. Fig. S6 highlights how the model represents the lignin chains in the solution. All of this information is implemented collectively to calculate the molecular weight of the chains, evaluate the activation energy barriers, and find the S/G ratio. Overall, the model output aligns well

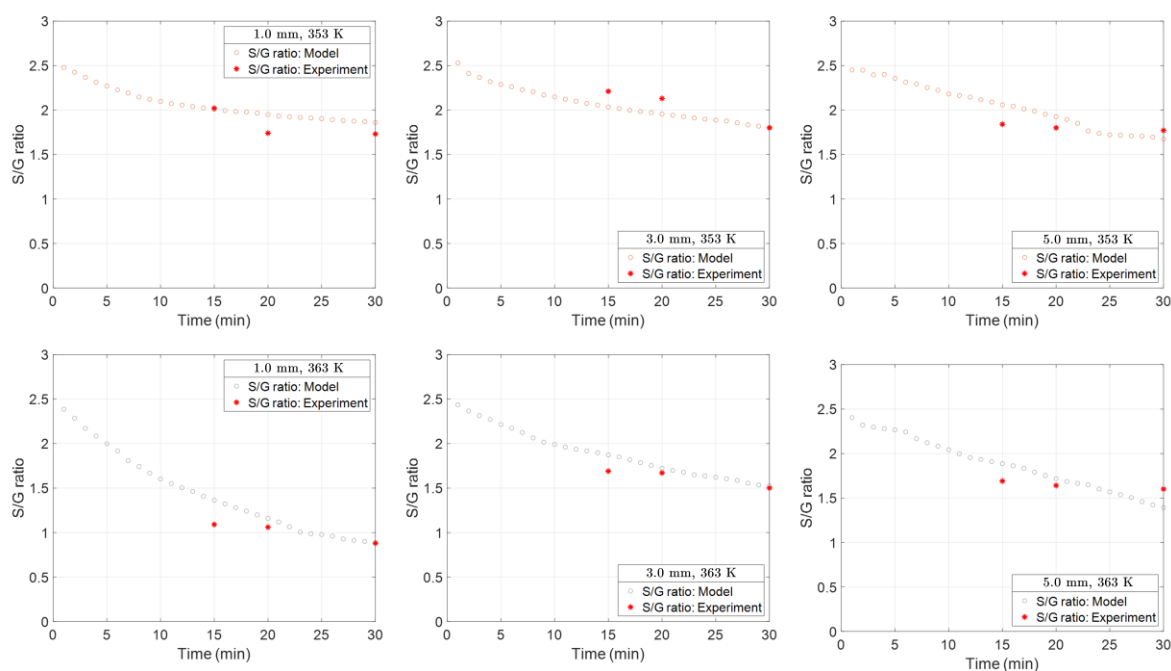


Fig. 7. Experimental validation of the multiscale kMC model for the S/G ratio under each condition.

with the experiments. It implies that the developed multiscale model explains the complicated lignin reactions well. While investigating the key lignin properties, as highlighted above, every component of this model – DFT, AIMD, and kMC – had its own contribution to interpreting the lignin behavior extensively during fractionation. Specifically, the parameters derived from DFT and AIMD calculations would be valuable, and they provide a solid basis for future research on lignin that requires data on how rates change with temperature and S/G ratio. Further, based on these parameters, the developed model can provide real-time information on the fractionation system, which will play a crucial role in lignin property tracking for further applications.

Remark 2. Based on the suggested model, further development of a controller to regulate lignin properties is possible. While the design of model-based controllers^[64-67] and process design and scale-up^[68-70] are common strategies, existing models specifically for lignin in biorefineries have been limited. With the solid foundation of lignin reaction dynamics presented here, we foresee various advanced applications including the novel process design and the development of a dedicated controller to tailor lignin properties, such as molecular weight distribution and S/G ratio.

4. Conclusions

Our study introduces a “real-time molecular-level sensing approach”, a theoretical model designed to address key challenges in studying lignin reactions, such as the limited frequency of experimental sampling and the lack of theoretical parameters to understand lignin behavior under various temperatures. Specifically, the model successfully captures the dynamic nature of lignin reactions, including delignification, redeposition, and de/repolymerization, by integrating sparse experimental sampling with theoretical activation energy barriers. This synergistic integration of DFT, AIMD, and kMC methodologies provided comprehensive information on lignin from atomic to macroscopic scales. Starting from the initial 8 lignin structures (5 CR and 3 OR units), our results clearly demonstrate how temperature variations induce structural transformation in lignin. These changes lead to shifts in reactive radical sites, significantly impacting the reactivity in de/repolymerization. Notably, repolymerization of lignin becomes challenging with higher temperatures due to the reduced reaction space caused by alterations in CR structures, while we observed that depolymerization dominates over the others. Furthermore, we found that the S-G conversion reaction is temperature-sensitive. In addition to such reactions, delignification and redeposition reactions provided global mass/heat flow between phases, and then the remaining lignin content in the chip and the system temperature were calculated. As a result, we obtained real-time molecular-level information about the dissolved lignin chains, including the molecular weight distributions and S/G ratios under diverse reaction conditions. Overall, this comprehensive study advances our understanding of the lignin system and expands our perspective on sustainable energy applications. Moreover, by selecting an appropriate set of simulation parameters, we believe that our approach can serve as a flexible model for simulating lignin and cellulose behaviors. This versatility can extend beyond the hardwood species, encompassing softwoods and gramineous biomasses.

CRedit authorship contribution statement

Chi Ho Lee: Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Juhyeon Kim:** Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Jiae Ryu:** Validation, Writing – original draft. **Wangyun Won:** Investigation, Methodology, writing – review & editing. **Chang Geun Yoo:** Conceptualization, Investigation, Methodology, Validation, Writing – review & editing. **Joseph Sang-II Kwon:** Conceptualization, Funding acquisition, Investigation, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgements

The authors gratefully acknowledge financial support from Artie McFerrin Department of Chemical Engineering and the Texas A&M Energy Institute. J. Ryu and C.G. Yoo were supported by the National Science Foundation CBET-2027125. W. Won was supported by the Industrial Technology Innovation Program (RS-2022-00144051, Development of operation control technology for carbon reduction) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea).

References

1. C. C. Azubuike, M. N. Allemann and J. K. Michener, Microbial assimilation of lignin-derived aromatic compounds and conversion to value-added products, *Current Opinion in Microbiology*, 65 (2022), 64-72, 10.1016/j.mib.2021.10.014
2. Z. W. Cao, M. Dierks, M. T. Clough, I. B. D. de Castro and R. Rinaldi, A Convergent Approach for a Deep Converting Lignin-First Biorefinery Rendering High-Energy-Density Drop-in Fuels, *Joule*, 2 (2018), 6, 1118-1133, 10.1016/j.joule.2018.03.012
3. L. Dessbesell, M. Paleologou, M. Leitch, R. Pulkki and C. Xu, Global lignin supply overview and kraft lignin potential as an alternative for petroleum-based polymers, *Renewable & Sustainable Energy Reviews*, 123 (2020), 10.1016/j.rser.2020.109768
4. M. S. Ganewatta, H. N. Lokupitiya and C. B. Tang, Lignin Biopolymers in the Age of Controlled Polymerization, *Polymers*, 11 (2019), 7, 10.3390/polym11071176
5. X. Yang, X. Z. Hu, Z. J. Yang, Q. Wang, A. Zaman, F. Huang and M. Jiang, Preparation of single O-methoxyphenol from lignin and related liquor products as reinforcement for epoxy resin, *Renewable Energy*, 162 (2020), 1285-1291, 10.1016/j.renene.2020.07.143
6. A. Grossman and W. Vermerris, Lignin-based polymers and nanomaterials, *Current Opinion in Biotechnology*, 56 (2019), 112-120, 10.1016/j.copbio.2018.10.009
7. J. Zakzeski, P. C. A. Bruijninx, A. L. Jongerius and B. M. Weckhuysen, The Catalytic Valorization of Lignin for the Production of Renewable Chemicals, *Chemical Reviews*, 110 (2010), 6, 3552-3599, 10.1021/cr900354u
8. S. Jeong, J. Ryu, Q. Yang, J.-Y. Zhu, and C. G. Yoo, Recent advances in hydrotropic solvent systems for lignocellulosic biomass utilization, *Green Chemistry*, Advance Article (2024), 10.1039/D3GC03309K

9. D. He, Y. Wang, C. G. Yoo, Q.-J. Chen, and Q. Yang, The fractionation of woody biomass under mild conditions using bifunctional phenol-4-sulfonic acid as a catalyst and lignin solvent, *Green Chemistry*, 22 (2020), 5414-5422, 10.1039/D0GC01722A
10. B. M. Upton and A. M. Kasko, Strategies for the Conversion of Lignin to High-Value Polymeric Materials: Review and Perspective, *Chemical Reviews*, 116 (2016), 4, 2275-2306, 10.1021/acs.chemrev.5b00345
11. M. M. Abu-Omar, K. Barta, G. T. Beckham, J. S. Luterbacher, J. Ralph, R. Rinaldi, et al, Guidelines for performing lignin-first biorefining, *Energy & Environmental Science*, 14 (2021), 1, 262-292, 10.1039/d0ee02870c
12. P. Prinsen, J. Rencoret, A. Gutiérrez, T. Liitiä, T. Tamminen, J. L. Colodette, et al, Modification of the Lignin Structure during Alkaline Delignification of Eucalyptus Wood by Kraft, Soda-AQ, and Soda-O Cooking, *Wood Science and Technology*, 52 (2013), 45, 15702-15712, 10.1021/ie401364d
13. Y. Musha and D. A. I. Goring, Distribution of Syringyl and Guaiacyl Moieties in Hardwoods as Indicated by Ultraviolet Microscopy, *Wood Science and Technology*, 9 (1975), 1, 45-58.
14. T. R. Pang, G. H. Wang, H. Sun, W. J. Sui and C. L. Si, Lignin fractionation: Effective strategy to reduce molecule weight dependent heterogeneity for upgraded lignin valorization, *Industrial Crops and Products*, 165 (2021), 10.1016/j.indcrop.2021.113442.
15. S. C. D. Eswaran, S. Subramaniam, U. Sanyal, R. Rallo and X. Zhang, Molecular structural dataset of lignin macromolecule elucidating experimental structural compositions, *Scientific Data*, 9 (2022), 1, 10.1038/s41597-022-01709-4.
16. R. B. Santos, E. A. Capanema, M. Y. Balakshin, H. M. Chang and H. Jameel, Lignin Structural Variation in Hardwood Species, *Journal of Agricultural and Food Chemistry*, 60 (2012), 19, 4923-4930, 10.1021/jf301276a

17. T. Phongpreecha, N. C. Hool, R. J. Stoklosa, A. S. Klett, C. E. Foster, A. Bhalla, et al, Predicting lignin depolymerization yields from quantifiable properties using fractionated biorefinery lignins, *Green Chemistry*, 19 (2017), 21, 5131-5143, 10.1039/c7gc02023f
18. J. Kim, S. Pahari, J. Ryu, M. Zhang, Q. Yang, C. G. Yoo, and J. S. I. Kwon, Advancing biomass fractionation with real-time prediction of lignin content and MWd: A kMC-based multiscale model for optimized lignin extraction, *Chemical Engineering Journal*, 479 (2024), 147226, 10.1016/j.cej.2023.147226.
19. K. H. Kim, K. Jeong, S. S. Kim and R. C. Brown, Kinetic understanding of the effect of Na and Mg on pyrolytic behavior of lignin using a distributed activation energy model and density functional theory modelling, *Green Chemistry*, 21 (2019), 5, 1099-1107, 10.1039/C8GC02948B
20. S. Pahari, J. Kim, H.-K. Choi, M. Zhang, A. Ji, C. G. Yoo, and J. S.-I. Kwon, Multiscale kinetic modelling of biomass fractionation in an experiment: Understanding individual reaction mechanisms and cellulose degradation, *Chemical Engineering Journal*, 467 (2023), 143021, 10.16/j.cej.2023.143021
21. H.-K. Choi, and J. S.-I. Kwon, Multiscale modelling and control of Kappa number and porosity in a batch-type pulp digester, *AIChE Journal*, 65 (2019), e16589, 10.1002/aic.16589
22. H.-K. Choi, and J. S.-I. Kwon, Multiscale modelling and multiobjective control of wood fiber morphology in batch pulp digester, *AIChE Journal*, 66 (2020), e16972, 10.1002/aic.16972
23. S. H. Son, H.-K. Choi, and J. S.-I. Kwon, Multiscale modelling and control of pulp digester under fiber-to-fiber heterogeneity, *Computers & Chemical Engineering*, 143 (2020), 107117, 10.1016/j.compchemeng.2020.1071

24. J. Jung, H.-K. Choi, S. H. Son, J. S.-I. Kwon, and J. H. Lee, Multiscale modelling of fiber deformation: Application to a batch pulp digester for model predictive control of fiber strength, *Computers & Chemical Engineering*, 158 (2022), 107640, 10.1016/j.compchemeng.2021.107640
25. A. J. Yanez, W. Li, R. Mabon, and L. J. Broadbelt, A stochastic method to generate libraries of structural representations of lignin, *Energy & Fuels*, 30 (2016), 5835-5845, 10.1021/acs.energyfuels.6b00966
26. M. J. Orella, T. Z. H. Gani, J. V. Vermaas, M. L. Stone, E. M. Anderson, G. T. Beckham, F. R. Brushett, and Y. Román-Leshkov, Lignin-kMC: A toolkit for simulating lignin biosynthesis, *ACS Sustainable Chemistry & Engineering*, 7 (2019), 18313-18322, 10.1021/acssuschemeng.9b03534
27. S. c. d. Eswaran, S. Subramaniam, U. Sanyal, R. Rallo, and X. Zhang, Molecular structural dataset of lignin macromolecule elucidating experimental structural compositions, *Scientific data*, 9 (2022), 647, 10.1038/s41597-022-01709-4
28. A. J. Yanez, P. Natarajan, W. Li, R. Mabon, and L. J. Broadbelt, Coupled structural and kinetic model of lignin fast pyrolysis, *Energy & Fuels*, 32 (2018), 1822-1830, 10.1021/acs.energyfuels.7b03311
29. R. M. Prunescu, M. Blanke, J. G. Jakobsen, and G. Sin, Dynamic modelling and validation of a biomass hydrothermal pretreatment process—a demonstration scale study, *AIChE Journal*, 61 (2015), 4235-4250, 10.1002/aic.14954
30. M. T. Reza, W. Yan, M. H. Uddin, J. G. Lynam, S. K. Hoekman, C. J. Coronella, and V. R. Vásquez, Reaction kinetics of hydrothermal carbonization of loblolly pine, *Bioresource Technology*, 139 (2013), 161-169, 10.1016/j.biortech.2013.04.028

31. J. B. McDermott and M. T. Klein, Chemical and probabilistic modelling of complex reactions: a lignin depolymerization example, *Chemical Engineering Science*, 41 (1986), 1053-1060, 10.1016/0009-2509(86)87192-5
32. L. Petridis, R. Schulz and J. C. Smith, Simulation Analysis of the Temperature Dependence of Lignin Structure and Dynamics, *Journal of the American Chemical Society*, 133 (2011), 50, 10.1021/ja206839u
33. C. H. Lee, S. Pahari, N. Sitapure, M. A. Barteau, and J. S.-I. Kwon, Investigating high-performance non-precious transition metal oxide catalysts for nitrogen reduction reaction: A multifaceted DFT-kMC-LSTM approach, *ACS Catalysis*, 13 (2023), 8336-8346, 10.1021/acscatal.3c01360
34. C. H. Lee, S. Pahari, N. Sitapure, M. A. Barteau, and J. S.-I. Kwon, DFT-kMC analysis for identifying novel bimetallic electrocatalysts for enhanced NRR performance by suppressing HER at ambient conditions via active-site separation, *ACS Catalysis*, 12 (2022), 15609-15617, 10.1021/acscatal.2c04797
35. S. Pahari, Y.-T. Lin, S. Liu, C. H. Lee, M. Akbulut, J. S.-I. Kwon, Stochastic optimal control of mesostructure of supramolecular assemblies using dissipative particle dynamics and dynamic programming with experimental validation, *Chemical Engineering Journal*, 475 (2023), 145087, 10.1016/j.cej.2023.145087
36. A. J. Yanez, P. Natarajan, W. J. Li, R. Mabon and L. J. Broadbelt, Coupled Structural and Kinetic Model of Lignin Fast Pyrolysis, *Energy & Fuels*, 32 (2018), 2, 1822-1830, 10.1021/acs.energyfuels.7b03311
37. J. D. DeMartini, M. Foston, X. Meng, S. Jung, R. Kuman, A. J. Ragauskas, and C.E. Wyman, How chip size impacts steam pretreatment effectiveness for biological conversion of poplar wood into fermentable sugars, *Biotechnology for Biofuels*, 8 (2015), 209, 10.1186/s13068-015-0373-1

38. I. D. Hartley, F. A. Kamke, and H. Peemoeller, Absolute moisture content determination of aspen wood below the fiber saturation point using pulsed NMR, *Holzforschung*, 48 (1994), 474-479, 10.1515/hfsg.1994.48.6.474
39. A. Jun, U. W. Tschirner, and Z. Tauer, Hemicellulose extraction from aspen chips prior to kraft pulping utilizing kraft white liquor, *Biomass and Bioenergy*, 37 (2012), 229-236, 10.1016/j.biombioe.2011.12.008
40. Y. Wu, S. Ge, C. Xia, L. Cai, C. Mei, C. Sonne, Y.-K. Park, Y.-M. Kim, W.-H. Chen, J.-S. Chang, and S.S. Lam, Using low carbon footprint high-pressure carbon dioxide in bioconversion of aspen branch waste for sustainable bioethanol production, *Bioresource Technology*, 313 (2020), 123675, 10.1016/j.biortech.2020.123675
41. S. S. Jagtap, S. S. Dhiman, T.-S. Kim, J. Li, J.-K. Lee, and Y. C. Kang, Enzymatic hydrolysis of aspen biomass into fermentable sugars by using lignocellulases from *Armillaria gemina*, *Bioresource Technology*, 133 (2013), 307-314, 10.1016/j.biortech.2013.01.118
42. J. R. Jensen, J. E. Morinelly, K. R. Gossen, M. J. Brodeur-Campbell, and D. R. Shonnard, Effects of dilute acid pretreatment conditions on enzymatic hydrolysis monomer and oligomer sugar yields for aspen, balsam, and switchgrass, *Bioresource Technology*, 101 (2010), 2317-2325, 10.1016/j.biortech.2009.11.038
43. M. J. Brodeur-Campbell, J. Klinger, and D. Shonnard, Feedstock mixture effects on sugar monomer recovery following dilute acid pretreatment and enzymatic hydrolysis, *Bioresource Technology*, 116 (2012), 320-326, 10.1016/j.biortech.2012.03.090
44. G. Kresse and J. Hafner, Lignin utilization: Ab-Initio Molecular-Dynamics for Open-Shell Transition-Metals, *Physical Review B* 48 (1993), 13115, 10.1103/PhysRevB.48.13115
45. P. E. Blochl, Projector Augmented-Wave Method, *Physical Review B*, 50 (1994), 24, 17953-17979, 10.1103/PhysRevB.50.17953

46. J. P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple, Physical Review Letters, 77 (1996), 18, 3865-3868, 10.1103/PhysRevLett.77.3865.
47. H. C. Chow and S. H. Vosko, Special Points for Two-Dimensional Brillouin-Zone or Wigner-Seitz Cell Integrations, Canadian Journal of Physics, 58 (1980), 4, 497-503, 10.1139/p80-070
48. G. Henkelman, B. P. Uberuaga and H. Jónsson, A climbing image nudged elastic band method for finding saddle points and minimum energy paths, Journal of Chemical Physics, 113 (2000), 22, 9901-9904, 10.1063/1.1329672
49. W. G. Hoover, Canonical Dynamics - Equilibrium Phase-Space Distributions, Physical Review A, Physical Review A, 31 (1985), 3, 1695-1697. 10.1103/PhysRevA.31.1695.
50. S. Nose, Constant Temperature Molecular-Dynamics Methods, Progress of Theoretical Physics Supplement, (1991), 103 1-46, 10.1143/PTPS.103.1
51. K. M. Davis, M. Rover, R. C. Brown, X. L. Bai, Z. Y. Wen and L. R. Jarboe, Recovery and Utilization of Lignin Monomers as Part of the Biorefinery Approach, Energies, 9 (2016), 10, 10.3390/en9100808.
52. C. L. Chio, M. Sain and W. S. Qin, Lignin utilization: A review of lignin depolymerization from various aspects, Renewable & Sustainable Energy Reviews, 107 (2019), 232-249, 10.1016/j.rser.2019.03.008.
53. C. F. Zhang and F. Wang, Catalytic Lignin Depolymerization to Aromatic Chemicals, Accounts of Chemical Research, 53 (2020), 2, 470-484, 10.1021/acs.accounts.9b00573
54. A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter, D. Templeton, and D. Crocker. Determination of structural carbohydrates and lignin in biomass. Technical Report NREL 2008; NREL/TP- 510-42618.
55. T. Azad, M. L. Auad, T. Elder and A. J. Adamczyk, Toward Native Hardwood Lignin Pyrolysis: Insights into Reaction Energetics from Density Functional Theory, Energy & Fuels, 37 (2023),

1, 401-433, 10.1021/acs.energyfuels.2c03025.

56. J. J. Stewart, T. Akiyama, C. Chapple, J. Ralph, S. D. Mansfield, The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar, *Plant physiology*, 150 (2009), 2, 621-635, 10.1104/pp.109.137059.

57. Z. Nasrun, L. S. Osman, N. H. Abd Latif, N. H. H. Elias, M. Saidin, S. Shahidan, et al, Conversion of archeological iron rust employing coconut husk lignin, *International Journal of Biological Macromolecules*, 253 (2023), 10.1016/j.ijbiomac.2023.126786.

58. J. P. DeVita, L. M. Sander and P. Smereka, Multiscale kinetic Monte Carlo algorithm for simulating epitaxial growth, 72 (2005), 20, 10.1103/PhysRevB.72.205421.

59. S. Pahari, B. Bhadriraju, M. Akbulut and J. S. I. Kwon, A slip-spring framework to study relaxation dynamics of entangled wormlike micelles with kinetic Monte Carlo algorithm, *Journal of Colloid and Interface Science*, 600 (2021), 550-560, 10.1016/j.jcis.2021.05.032.

60. H. K. Choi and J. S. I. Kwon, Modeling and control of cell wall thickness in batch delignification, *Computers & Chemical Engineering*, 128 (2019), 512-523, 10.1016/j.compchemeng.2019.06.025.

61. S. K. Bose, R. C. Francis, M. Govender, T. Bush and A. Spark, Lignin content versus syringyl to guaiacyl ratio amongst poplars, *Bioresource Technology*, 100 (2009), 4, 1628-1633, 10.1016/j.biortech.2008.08.046

62. G. Papa, P. Varanasi, L. Sun, G. Cheng, V. Stavila, B. Holmes, et al., Exploring the effect of different plant lignin content and composition on ionic liquid pretreatment efficiency and enzymatic saccharification of L mutants, *Bioresource Technology*, 117 (2012), 352-359, 10.1016/j.biortech.2012.04.065

63. A. Sequeiros and J. Labidi, Characterization and determination of the S/G ratio via Py-GC/MS of agricultural and industrial residues, *Industrial Crops and Products*, 97 (2017), 469-476,

10.1016/j.indcrop.2016.12.056.

64. H.-K. Choi, S. H. Son, and J. S.-I. Kwon, Inferential model predictive control of continuous pulping under grade transition, *Industrial & Engineering Chemistry Research*, 60 (2021), 3699-3710, 10.1021/acs.iecr.0c06216

65. S. H. Son, H.-K. Choi, and J. S.-I. Kwon, Application of offset-free Koopman-based model predictive control to a batch pulp digester, 67 (2021), e17301, 10.1002/aic.17301

66. S. H. Son, H.-K. Choi, J. Moon, and J. S.-I. Kwon, Hybrid Koopman model predictive control of nonlinear systems using multiple EDMD models: An application to a batch pulp digester with feed fluctuation, *Control Engineering Practice*, 118 (2022), 104956, 10.1016/j.conengprac.2021.104956

67. P. Shah, H.-K. Choi, and J. S.-I. Kwon, Achieving optimal paper properties: a layered multiscale kMC and LSTM-ANN-based control approach for Kraft pulping, *Processes*, 11 (2023), 809, 10.3390/pr11030809

68. B. Ahn, C. Park, J. J. Liu, Y. S. Ok, and W. Won, Maximizing the utilization of lignocellulosic biomass: Process development and analysis, *Renewable Energy*, 215 (2023), 119004, 10.1016/j.renene.2023.119004

69. C. Gong, S. Lee, and W. Won, Process development and analyses for the co-production of 2-methyltetrahydrofuran and 1,4-pentanediol from lignocellulosic biomass, *GCB Bioenergy*, 15 (2023), 900-915, 10.1111/gcbb.13058

70. S. Jeong, J. Kim, J. J. Liu, and W. Won, A lignocellulosic 2,3-butanediol strategy: Process development and integrative analyses, *Journal of Environmental Chemical Engineering*, 11 (2023), 110085, 10.1016/j.jece.2023.110085