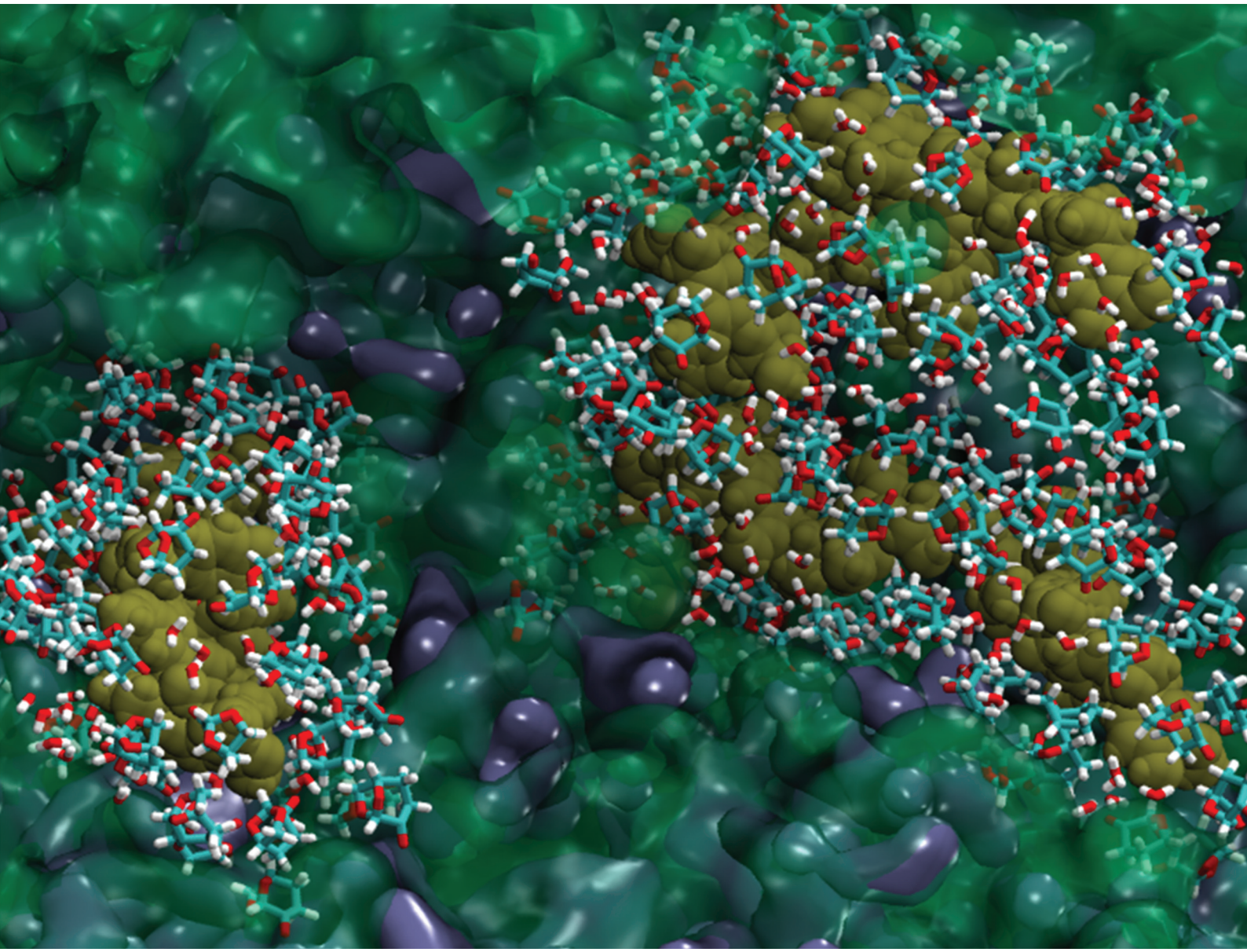
Volume 26
Number 6
21 March 2024
Pages 2933-3528

Green Chemistry

Cutting-edge research for a greener sustainable future

rsc.li/greenchem



ISSN 1463-9262

ROYAL SOCIETY OF CHEMISTRY

PAPER

Yunqiao Pu, Arthur J. Ragauskas *et al.*
Characterization and molecular simulation of lignin in Cyrene pretreatment of switchgrass

25 YEARS ANNIVERSARY

Green Chemistry

PAPER

View Article Online
View Journal | View Issue

Cite this: *Green Chem.*, 2024, **26**, 3170

Characterization and molecular simulation of lignin in Cyrene pretreatment of switchgrass†

Yun-Yan Wang,[b] Yunxuan Wang,[b] Luna Liang,[b] Micholas Dean Smith,[c,d] Xianzhi Meng,[b] Yunqiao Pu,*[e,f] Mitra Mazarei,[g] Rupesh Agarwal,[c,d] Shalini J. Rukmani,[d] Brian H. Davison[e,f] and Arthur J. Ragauskas*[a,b,f]

Biomass-derived solvents have been proposed as a novel pathway in biorefining for the realization of biofuels and bioproducts derived from lignocellulosic biomass. Cyrene derived from cellulose has recently been shown to have a high potential as a green organic solvent for pretreating poplar biomass. However, due to its high dynamic viscosity nature, high Cyrene concentration could cause negative effects on the sugar release of the pretreated biomass as well as driving up the operational cost of the lignin recovery. In this study, we combine experimental and computational approaches to examine the impact of Cyrene pretreatment with reduced Cyrene concentration under mild conditions on switchgrass lignin. Our experimental studies indicated correlation between pretreatment condition and recovery and structure modification of lignin. Switchgrass lignin extracted by Cyrene pretreatment possessed high preservation of β-O-4 ether inter-unit linkage, which could provide versatility in the integration of downstream lignin valorization into the modern biorefinery industries. Molecular modeling examining the solvation of switchgrass lignin polymer and the disaggregation of low-molecular weight lignin aggregates under pretreatment conditions indicated that a preferential interaction exists between Cyrene and lignin, which likely drives lignin release, and that the disruption of inter-lignin contacts can be modulated as a non-monotonic function of Cyrene : water ratio. Further, while Cyrene–lignin interactions permit the solubilization of lignin, simulations with proxy reactive-species reveal that changes to the diffusion of these reactive proxies and their localization near linkage sites under Cyrene conditions may inhibit chemical processes. The results indicated that loss of pretreatment efficacy caused by low Cyrene concentration could be compensated by prolonged pretreatment time and high catalyst dosage.

Received 21st June 2023,
Accepted 3rd October 2023

DOI: 10.1039/d3gc02239k

rsc.li/greenchem

Introduction

Plant-based feedstocks have been widely investigated as one of the most important and viable sources for clean energy and organic carbon to offset climate change caused by human activities and prepare for a carbon neutral economy.[1] Switchgrass (*Panicum virgatum* L.), a local perennial warm-season grass, has been considered as a promising herbaceous energy crop for producing biofuels and bioproducts in the U.S. due to its high productivity, flexible adaptability, and low agronomic inputs.[2,3] The biorefinery of lignocellulosic biomass involves saccharification and fermentation to breakdown polysaccharides and biochemically convert monosaccharides to ethanol or other liquid fuels. Similar to other plant-based biomass, switchgrass evolved a complex plant cell wall to resist enzymatic and microbial degradation in nature, making pretreatment the most expensive and energy-intensive process in lignocellulosic biorefinery. Such recalcitrance can be largely influenced by the content and structure of lignin.[4] Generally, a macromolecular backbone of herbaceous lignin is built up from mostly guaiacyl (G), syringyl (S), and *p*-hydroxyphenyl (H) subunits connected by three predominate inter-unit linkages including β-O-4 ether, β–β resinol and β-5 phenylcoumaran. Lignin is also crosslinked with hemicelluloses and pectin *via* ferulate and *p*-coumarate esters to form lignin–carbohydrate

[a]*Department of Forestry, Wildlife, and Fisheries, Center for Renewable Carbon, University of Tennessee Institute of Agriculture, Knoxville, TN 37996, USA. E-mail: aragausk@utk.edu*
[b]*Department of Chemical & Biomolecular Engineering, University of Tennessee Knoxville, Knoxville, TN 37996, USA*
[c]*Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee Knoxville, Knoxville, TN 37996, USA*
[d]*Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*
[e]*Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. E-mail: puy1@ornl.gov*
[f]*Joint Institute for Biological Sciences, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA*
[g]*Department of Plant Sciences, University of Tennessee, Knoxville, TN 37996, USA*
† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3gc02239k

complexes.[5–8] Unlike cellulose and hemicellulose, biological degradation of lignin remains a challenge to the field although various ligninolytic enzymes have been discovered since 1980s.[9] Thus lignin is generally considered as the most recalcitrant plant cell wall component, and should be extensively removed or altered by biomass pretreatment in order to enhance the cellulose accessible surface area for an efficient enzymatic hydrolysis process.

Organosolv process represents an emerging pretreatment technique where organic solvents or aqueous mixtures of organic solvents are used to disrupt the recalcitrant cell-wall structure by depolymerizing and solubilizing lignin and hemicelluloses, thereby yielding a cellulose-rich residue that could be easily converted to its monosugars.[10] Recently, a novel concept of a closed-loop biorefinery has been introduced as means to increase the profitability (and thus economic feasibility) of biofuels production. Under this approach, lignocellulosic feedstocks are pretreated with biomass-derived solvents such as γ-valerolactone,[11] tetrahydrofuran,[12] 2-methyltetrahydrofuran,[13] or dihydrolevoglucosenone (known commercially as Cyrene).[14] Among these biomass-derived solvents, Cyrene has gained growing attention as a greener substitute for toxic petroleum-based dipolar aprotic solvents such as *N*-methylpyrrolidone (NMP), dimethylformamide (DMF) and sulpholane, which are widely used in chemical and pharmaceutical industries.[15–18] Cyrene can be synthesized from cellulose in two steps, *via* pyrolysis followed by hydrogenation, and high yields have been reported for both processes, making Cyrene a promising candidate as biomass-derived solvent in biorefinery. The cost structure of Cyrene remains dynamic, and currently the production cost is estimated at $3 per kg,[19] with commercial production of a 1000 metric tons per annum (mtpa) plant in France being developed. Circa has a goal of 80 000 mtpa global offtake capacity by 2030. Cleary, future studies will need examine low-cost Cyrene recycling technologies while benefiting from future production efficiencies. Stini *et al.* reported that Cyrene could be easily recovered after the Mizoroki–Heck reaction by diluting the Cyrene-containing filtrate with water.[20] Ethyl acetate was then added to the mixture forming an organic layer. After removing the organic layer, Cyrene could be recovered from the aqueous layer by evaporating the water, resulting in a recovery yield of 70%. Though the effects of recycled Cyrene on biomass pretreatment is not clear, these results demonstrate the feasibility of recovering Cyrene by solvent–solvent extraction. Besides the greener features, Cyrene is a good solvent for a variety of technical lignins, similar to DMF.[14,21] When mixing with water, Cyrene can be found in ketone and geminal diol forms in an equilibrium.[16,22] Abranches *et al.* reported that the solvation capacity of Cyrene/water towards hydrophobic solutes was favored by its ketone form rather than the geminal diol.[22] The amount of the geminal diol reaches maximum (~75% of Cyrene) in the aqueous solution χ_{Cyrene} = 0.3, and 50% of Cyrene was in geminal form when χ_{Cyrene} ~0.38.

We previously showed that Cyrene can successfully be utilized, in both neat Cyrene and 4 : 1 v_{cyrene}/v_{water} aqueous solutions, in the pretreatment of hardwood biomass and achieve nearly 100% cellulose conversion in the following saccharification process.[14] The extent of condensation and loss of β-O-4 ether inter-unit linkage could be significantly reduced in the Cyrene pretreatment under mild conditions due to the exceptional high solubility of lignin macromolecule in Cyrene. The preservation of C–O bonds along with decreased C–C bonds formed by condensation reactions could benefit the downstream depolymerization of lignin into aromatic monomers and oligomers, paving a path towards lignin valorization.[23,24] However, it was also suggested that the application of Cyrene in organosolv pretreatment can be constrained by its high dynamic viscosity, which could inhibit the following enzymatic hydrolysis due to incomplete removal of Cyrene or possible binding to cellulose from the pretreated biomass.[14] For practical consideration, lowering the viscosity of Cyrene/water mixture could reduce energy input on mass mixing in the reactor as well as the energy used on solid–liquid separation. Besides, effects of Cyrene/water ratio at low Cyrene volume fraction on pretreatment of switchgrass, an important herbaceous bioenergy feedstock, have not been studied. In addition, the previous study has only investigated the pretreatment effect of the Cyrene/water based solvent system with high Cyrene volume, while the mechanism of interactions between Cyrene and switchgrass lignin, which could reveal the fundamental of solvent extraction of lignin, was not clear. Therefore, it is necessary to investigate the pretreatment effect of Cyrene on lignin structure and Cyrene–lignin interactions with lower Cyrene volumetric fractions. Hereby, we explored Cyrene pretreatment on switchgrass using low Cyrene volume fraction (Cyrene/water = 2 : 1, 1 : 1 and 1 : 2) under a mild condition (120 °C). The correlation between delignification efficiency and lignin structure was investigated by various analytic techniques including gel permeation chromatography (GPC), ^{31}P nuclear magnetic resonance (NMR), and two-dimensional (2D) heteronuclear single quantum coherence (HSQC) NMR. To further understand the interactions between Cyrene/water and lignin, molecular dynamics simulations were conducted using newly refined Cyrene force-field parameters.

Materials and methods

Materials and chemicals

Switchgrass (*Panicum virgatum* L.) cultivar "Alamo" was from a University of Tennessee-Knoxville experimental field located at the Plant Sciences unit of the East Tennessee Research and Education Center (ETREC). The switchgrass samples used in this analysis were second-year field-grown and were fully senescent upon harvesting. The aboveground plant biomass was harvested in December of 2017 and oven dried at 43 °C for 96 h and chipped into 3.5–12 cm long and 1.2–3.5 mm diameter wide pieces. The switchgrass was Wiley milled and extracted with toluene/ethanol (2 : 1, v/v) for 8 h using a Soxhlet extractor before use for lignin isolation, and then air-dried. Chemicals such as dihydrolevoglucosenone (Cyrene™)

and sulfuric acid were purchased from Sigma-Aldrich and were used as received.

Cyrene pretreatment and lignin isolation

10 g of the extractive-free switchgrass was suspended in 150 mL of Cyrene/water mixture with stirring for 10 min. The pretreatment reactions were carried out at 120 °C for 1–4 h duration time and 33–100 mM sulfuric acid (Table 1) in a 300 mL Parr reactor equipped with a 4848 controller. After quenching to room temperature, the pretreated switchgrass biomass was separated from the slurry by vacuum filtration. The process solid was washed with 75 mL of Cyrene/water (1 : 2, v/v), followed by 800 mL deionized (DI) water. The filtrates were combined and subjected to centrifugation to collect lignin precipitated from the Cyrene/water mixture during water wash.

To isolate cellulolytic enzyme lignin (CEL) from untreated switchgrass, ~0.5 g of biomass was hydrolyzed twice with Accellerase® 1500 (BCA protein content 82 mg ml^{-1}, loadings of 80 mg g^{-1} biomass) in 20 mL of 50 mM sodium acetate buffer for 48 h.[14] The washed and freeze-dried solid residue obtained after enzymatic hydrolysis was extracted twice with 96% (v/v) 1,4-dioxane/water mixture at room temperature for 48 h. The switchgrass CEL was recovered from the combined extracts by rotary evaporation and freeze drying. All switchgrass lignin samples were dried overnight at 40 °C prior to characterization.

Characterization of switchgrass lignin

The molecular weight and distribution of switchgrass lignin were determined by using gel permeation chromatography (GPC) that was performed on an Agilent GPC SECurity 1200 system equipped with Water Styragel columns (Water Corporation, Milford, MA) at 35 °C using tetrahydrofuran as the mobile phase and a flow rate at 0.3 mL min^{-1}. Prior to GPC analysis, the lignin samples were acetylated and purified according to a previous literature.[25]

Table 1 The reaction conditions of Cyrene pretreatment and the yields of lignin and solid residue after pretreatment estimated on the basis of untreated extractive-free switchgrass

Cyrene/water v/v		H_2SO_4, mM	Duration, h	Lignin, wt%	Solid residue, wt%
2 : 1	#1	33	1	8.8	52.3
	#2	67	1	13.1	46.4
1 : 1	#3	0	4	0	87.7
	#4	33	1	5.7	52.4
	#5	67	1	6.0	52.2
	#6	100	1	8.0	47.8
1 : 2	#7	33	1	2.1	63.1
	#8	33	2	4.6	55.5
	#9	67	1	5.4	52.1
	#10	100	1	6.7	50.9
	#11	100	2	7.7	45.0

Two-dimensional ^{1}H–^{13}C heteronuclear single quantum coherence (HSQC) correlation and ^{31}P NMR spectra were acquired on a Bruker Avance III HD 500 MHz spectrometer (Bruker, Billerica, MA, USA) equipped with an N_2 cryoprobe according to the procedures published previously.[25,26] The HSQC and ^{31}P NMR spectra were processed and analyzed using the Bruker TopSpin software (V 4.0.6) and MestReNova (V 11.0.3-18688, Mestrelab Research), respectively. The relative abundance of the lignin subunits and inter-unit linkages detected by 2D HSQC NMR were calculated based on the procedure published previously.[27]

Composition analysis and enzymatic hydrolysis

The raw and pretreated switchgrass samples (~120 mg) were hydrolyzed with 1.5 mL 72 wt% H_2SO_4 for 1 h at 30 °C. Then the hydrolysates were diluted to 4 wt% H_2SO_4 with deionized water and then autoclaved at 121 °C for 1 h for a second hydrolysis. The hydrolysis solution was cooled to room temperature and filtered through a crucible. The residue was dried in an oven overnight and weighted to determine the Klason lignin content. The filtrate was used for sugar content analysis by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) using Dionex ICS-3000 (Dionex Corp., USA).

In the sugar release determination, ~1 g of the pretreated switchgrass sample was post-treated by soaking in 1% NaOH solution overnight, and then washed thoroughly with water. The samples were then mixed with 100 mM sodium acetate buffer at pH 5.0 and Accellerase® 1500 (BCA protein content 82 mg ml^{-1}) was added to the mixture at a loadings of 80 mg g^{-1} biomass. The mixture was incubated at 50 °C at 150 rpm for 72 h. The hydrolysates were filtered and the glucose contents were measured using high-performance liquid chromatography (HPLC) (PerkinElmer, Shelton, CT).

Molecular modeling

To probe, in atomic detail, the interactions between Cyrene–water and Cyrene pretreatments on lignin we developed a new CHARMM-compatible of Cyrene (see ESI† for parameterization and validation details) and performed simulations of three model systems: switchgrass-like lignin aggregates, isolated (single chain) switchgrass lignin, and a G–G β-O-4 linked dimer. Aggregate systems were simulated under bulk Cyrene, bulk water, and 4 : 1 Cyrene : water (v/v), 2 : 1 Cyrene : water (v/v), 1 : 1 Cyrene : water (v/v), and 1 : 2 Cyrene : water conditions at a temperature of 393.15 K and target pressure of 1 bar. The isolated lignin systems were simulated under bulk Cyrene and bulk water conditions, as well as 2 : 1 (v/v), 1 : 1 (v/v), and 1 : 2 (v/v) Cyrene : water environments at a temperature of 393.15 K and pressure of 1 bar. The model dimer system was simulated under 2 : 1 v/v Cyrene : water and bulk water conditions with and without Na^+/HSO_4^- ions present. As the simulations reported here are purely classical in nature, free protons and proton transfer are not permitted; however, in the interest of exploring how acidic anions and water molecules may coordinate near the commonly cleaved β-O-4

linkage, prior to a reaction occurring, we mimic the addition of acid with the use of reactive species proxies: a single Na^+ cation and HSO_4^- anion. Dimer simulations, similar to aggregate and isolated lignin simulations, were performed with a target temperature of 393.15 K and a pressure of 1 bar. Additional aggregate simulations using a softwood (all G subunit) lignin performed as a secondary validation of the compatibility of the new Cyrene parameters with the CHARMM lignin force-field were also performed and are described in the ESI.†

The molecular model used for the switchgrass lignin reported is a linear lignin containing 4 S subunits, 6 G subunits, and 1 H subunit, with a sequence of SSSHGGGGSGG. Additionally, the model contains 7 β-O-4 linkages, 2 β-5 linkages, and 1 β–β linkage. For the initial conditions of the lignin aggregates, three identical lignin chains with the described sequence and linkage composition were randomly placed in a cubic water box with an edge-length of 9 nm and simulated for 100 ns at a temperature of 303 K and a pressure of 1 bar; after 100 ns of simulation time a lignin aggregate structure was formed and this was used as the initial aggregate configuration for the high temperature water, Cyrene, and cosolvent conditions. For the single isolated chain simulations, a single isolated lignin chain was extracted from the aggregate simulation box. Simulation details are noted below.

All simulations were performed using the GROMACS simulation package[28] and CHARMM-family force-fields for lignin[29] and Cyrene. Initial configurations, either dimer, single chain, or aggregate, were placed in cubic boxes and solvated with Cyrene or Cyrene and water at the prescribed volume fractions using the gmx solvate command within GROMACS. Following solvation, each system was energy minimized to (single) machine precision using the steepest decent algorithm. Using the energy minimized structures, five independent NPT relaxation simulations, using the v-rescale[30] and Berendsen thermo/barostats[31] were performed for 2 ns each. Post-relaxation, the coordinates and velocities from the last frame of each of the relaxation simulations were used as initial conditions for NPT production simulations which made use of the v-rescale thermostat and Parrinello–Rahman barostat.[32] For the lignin aggregate systems, five independent production simulations were generated, each with a simulation length of 70 ns. For case of the isolated lignin chain, five independent production simulations were performed, each with a simulation length of 30 ns. For the model lignin dimer systems, 20 independent simulations were performed, each for a length of 15 ns. The shorted simulations times for the isolated chain and dimer simulations were selected based on the observed first exit times for lignin–lignin disassociation in the simulations of lignin aggregates under cosolvent conditions.

All simulations treat non-bonded interactions with a smooth force-switch cutoff of 1.2 nm, and long-range electrostatics with the particle-mesh Ewald (PME) formalism. Additionally, all simulations make use of constraints (using the LINCS algorithm[33]) for hydrogens. An integration timestep of 2 fs was used for all relaxation and production simulations. All simulation analyses were performed making use of the built in GROMACS analysis modules and are described within the Results and discussion section, where appropriate.

Results and discussion

Cyrene pretreatment condition and yields

The chemical stability and fluid viscosity were the two major concerns in the screening of Cyrene pretreatment condition.[14] Mild reaction temperature (120 °C) and low Cyrene/water ratio were employed in this work as listed in Table 1. Although Cyrene possesses a high boiling temperature around 230 °C, it was reported that Cyrene decomposed at 140 °C.[16] In fact, the pressure in the Parr reactor increased rapidly when temperature raised above 140 °C during the Cyrene pretreatment. Exothermic reactions were also observed at low temperatures; therefore, the heating was stopped around 60–65 °C depending on the Cyrene/water ratio and H_2SO_4 concentration to allow the temperature increase and stabilization, and then external heating was provided to keep the reaction running at 120 °C.

The geminal diol formed by reaction between Cyrene carbonyl group and water partially dissociates in water and lowers the pH of Cyrene/water mixture. As shown in Fig. 1, the pH dropped sharply initially and remained around 2.85 when the volume fraction of Cyrene surpassed water. Yet, the inherent acidity of Cyrene/water mixture is insufficient to catalyze delignification of switchgrass at 120 °C even the duration time was extended to 4 h (Table 1). Control tests without biomass were carried out to investigate changes of Cyrene in the solvent system. In brief, pure Cyrene, a mixture of Cyrene/water, and a mixture of Cyrene/water/sulfuric acid were heated at 120 °C for 1 h, and the resulting mixtures were characterized by 2D HSQC NMR. The NMR spectra are shown in Fig. S1.† It can be observed from Fig. S1A and S1B† that pure Cyrene is stable at 120 °C. New cross-signals were detected at δ_C/δ_H 66.0/

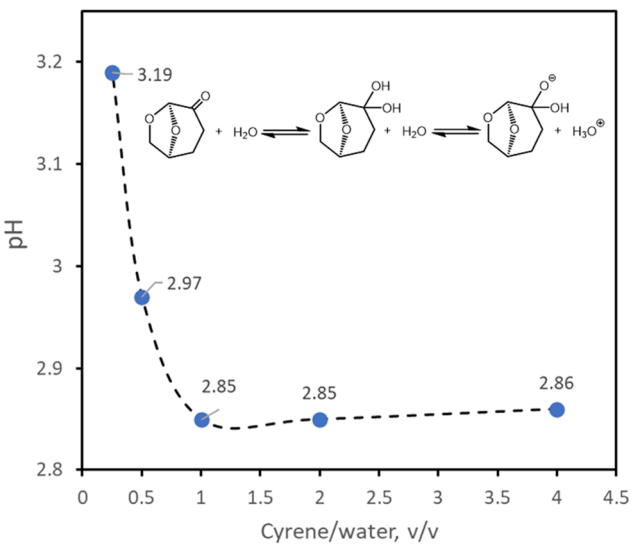


Fig. 1 The correlation between pH and Cyrene/water ratio.

Table 2 Composition of pretreated switchgrass biomass, glucose release and cellulose conversion of the post-treated biomass by enzymatic hydrolysis

Sample	Composition, %				Delignification %	Glucose g g^{-1} biomass[a]	Cellulose Conversion[b]
	Lignin	Arabinose	Glucose	Xylose			
Switchgrass	25.2%	2.1%	36.4%	20.1%		0.201	40.5%
#1	18.4%	1.3%	52.3%	10.5%	61.8%	0.454	67.7%
#2	14.5%	0.0%	53.0%	6.4%	73.3%	0.368	57.4%
#7	25.1%	1.1%	44.0%	14.1%	37.2%	0.401	65.1%
#8	23.3%	1.3%	51.8%	9.9%	48.7%	0.439	64.0%
#9	21.8%	0.8%	54.9%	8.6%	54.9%	0.432	59.6%
#10	21.0%	1.6%	56.5%	6.6%	57.6%	0.391	54.4%

[a] Glucose released from the biomass obtained 1% NaOH post-treatment. [b] Calculated based on the glucose content in pretreated biomass.

3.5–3.7 ppm, 71.6/4.4 ppm, and 102.9/4.7 ppm (Fig. S1C†) after heating Cyrene/water 2 : 1 (v/v) mixture at 120 °C for 1 h, which were assigned to geminal diol. The addition of 33 mM sulfuric acid to the Cyrene/water mixture gave rise to a number of new peaks after heating, as shown in Fig. S1D.† These peaks are associated with the exothermic behavior observed during heating, and future study is needed to understand this reaction and its impact on the pretreatment process.

Though the inherent acidity of Cyrene/water mixture was unable to extract lignin from switchgrass, the system demonstrated high delignification after the addition of sulfuric acid, even at a low acid concentration of 33 mM, as shown in Table 2. Delignification reached 61.8% after 1 h pretreatment using Cyrene/water 2 : 1 mixture with 33 mM H_2SO_4 (pretreatment #1), while increasing the acid concentration to 67 mM further improved delignification to 73.3% (pretreatment #2). Li *et al.* reported 69.2% delignification from switchgrass after 3 h pretreatment at 160 °C using the ionic liquid 1-ethyl-3-methylimidazolium acetate, while dilute acid pretreatment at 160 °C for 20 min only led to 22.4% delignification.[34] Previous studies have demonstrated that Cyrene is an alternative to polar aprotic solvents like dimethylsulfoxide (DMSO) and DMF for lignin dissolution and processing,[35] and results of this study indicate that Cyrene is a promising solvent for extracting lignin from the biomass. After pretreatment #1, 56.5% of the extracted lignin was recovered, accounting for 34.9% of the total lignin in the switchgrass. From Table 1, it can be speculated that with the same volumetric ratio between Cyrene and water, extending the pretreatment time led to improved lignin recovery yield, along with decreases in the recovery yield of the solid residue. Comparing the results of pretreatment #1, 2, 4, 5, 7, and 9 shows that Cyrene/water 2 : 1 mixture is more efficient in delignification than the 1 : 1 and 1 : 2 mixtures, while when the ratio between Cyrene and water was 1 : 1, high sulfuric acid concentration is needed to achieve similar results. These results indicate that the volumetric fraction of Cyrene in the mixture plays an important role in biomass delignification, which was further investigated *via* molecular simulation and discussed in a later section.

Meng *et al.* reported that the cellulose conversion to glucose by enzymatic hydrolysis of Cyrene pretreated poplar was inhibited by free lignin that was previously dissolved in Cyrene and deposited back onto the biomass surface during water wash.[14] Post-treatment by 1% NaOH solution was able to remove the residual lignin and boost the cellulose conversion as significant weight losses were observed after post-treatment (Table 2). The untreated control switchgrass with 1% NaOH post-treatment released ~0.2 g glucose per g biomass with ~40% cellulose conversion during enzymatic hydrolysis. Overall, the Cyrene pretreated switchgrass in this study doubled glucose release and improved cellulose conversion by up to 27% at a mild pretreatment temperature of 120 °C as compared to the control switchgrass without Cyrene pretreatment. It should be noted that the highest cellulose conversion in this study was less than 70% (Table 2) under the conditions investigated. To achieve higher glucose release yield, further study on optimization of the Cyrene pretreatment process is needed. In addition, the recovery and utilization of hemicellulose dissolved into the pretreatment liquor after Cyrene pretreatment needs further investigation in future studies.

The impacts of Cyrene pretreatment on switchgrass lignin chemical structure

In an acid-catalyzed organosolv pretreatment, the lignin–carbohydrate matrix is broken down to release lignin fragments into the solution. Our previous study on co-product lignin obtained from THF organosolv pretreatment showed that the degree of lignin depolymerization depended on the reaction temperature, acid concentration and duration, among which reaction temperature was the key factor that can be employed to tune the molecular weight and structure of lignin.[26] In this work, the impact of reaction temperature was excluded considering the chemical instability of Cyrene at elevated temperatures, and preliminary data suggest that lower temperature is not effective in terms of delignification. The correlation between switchgrass lignin molecular weight and Cyrene pretreatment condition was demonstrated in Fig. 2. The acid-catalyzed Cyrene pretreatment was able to reduce switchgrass lignin molecular weight by 20–60% comparing to CEL, but interestingly, the average molecular weight dispersity ($Đ$) of these switchgrass lignin was 2.67 ± 0.3.2, close to CEL ($Đ$ = 2.64). When Cyrene/water volume ratio was 1 : 1 and 1 : 2, the lignin molecular weight decreased as more H_2SO_4 was added to the reaction system, showing an opposite trend to pretreat-

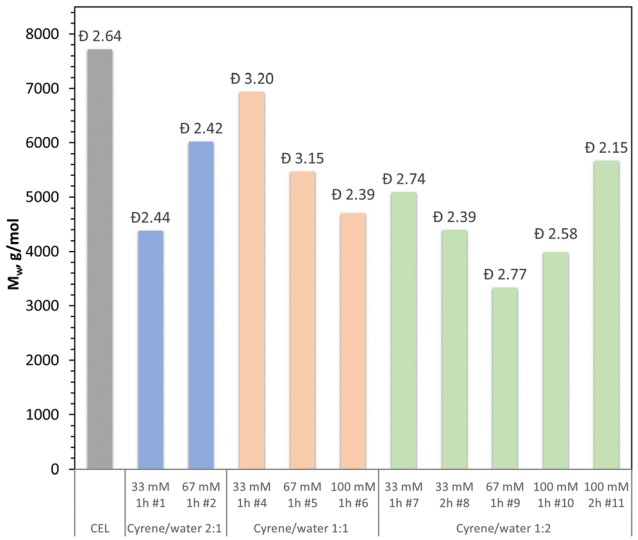


Fig. 2 The weight-average molecular weight (M_w) and molecular dispersity index ($Đ$) of switchgrass CEL and switchgrass Cyrene lignin extracted under different pretreatment conditions.

ment #1 and #2 (Cyrene/water = 2 : 1). Switchgrass lignin samples obtained from pretreatment #4 and #7 possessed unexpected high molecular weights and low yields, indicating that their macromolecular backbone underwent mild cleavage reactions. However, as the duration time or acid concentration increased, both lignin yield and molecular weight of pretreatment #9–11 started to show an upward trend.

Hydroxyl groups are the most important functionality for lignin, providing the latter with not only sites for various chemical modification, but also strong hydrogen bonding capacity and anti-oxidative/UV-blocking properties. ^{31}P NMR analysis involving reacting hydroxyl groups with phosphorylation reagent (Fig. 3A) can distinguish and quantify various hydroxyl groups found in lignin macromolecules.[36–38] The quantitative information of switchgrass lignin hydroxyl groups including aliphatic, phenolic and carboxylic ones is summarized in Fig. 3B. The content of aliphatic hydroxyl groups tended to decrease as the Cyrene pretreatment severity increases by adding more H_2SO_4 and/or extending duration time. Compared to CEL, the total hydroxyl content of Cyrene pretreated switchgrass lignin decreased dramatically due to the heavy loss (up to 50%) of aliphatic hydroxyl group. The fraction of free phenolic hydroxyl groups in total hydroxyl content was boosted from 13% (CEL) to 17–25% with similar quantitative results, indicating that the change was attributed mainly to the decrease of aliphatic hydroxyl groups. Fig. 3C shows the composition of phenolic hydroxyl groups including C_5-substituted, guaiacyl and *p*-hydroxylphenyl. The content of

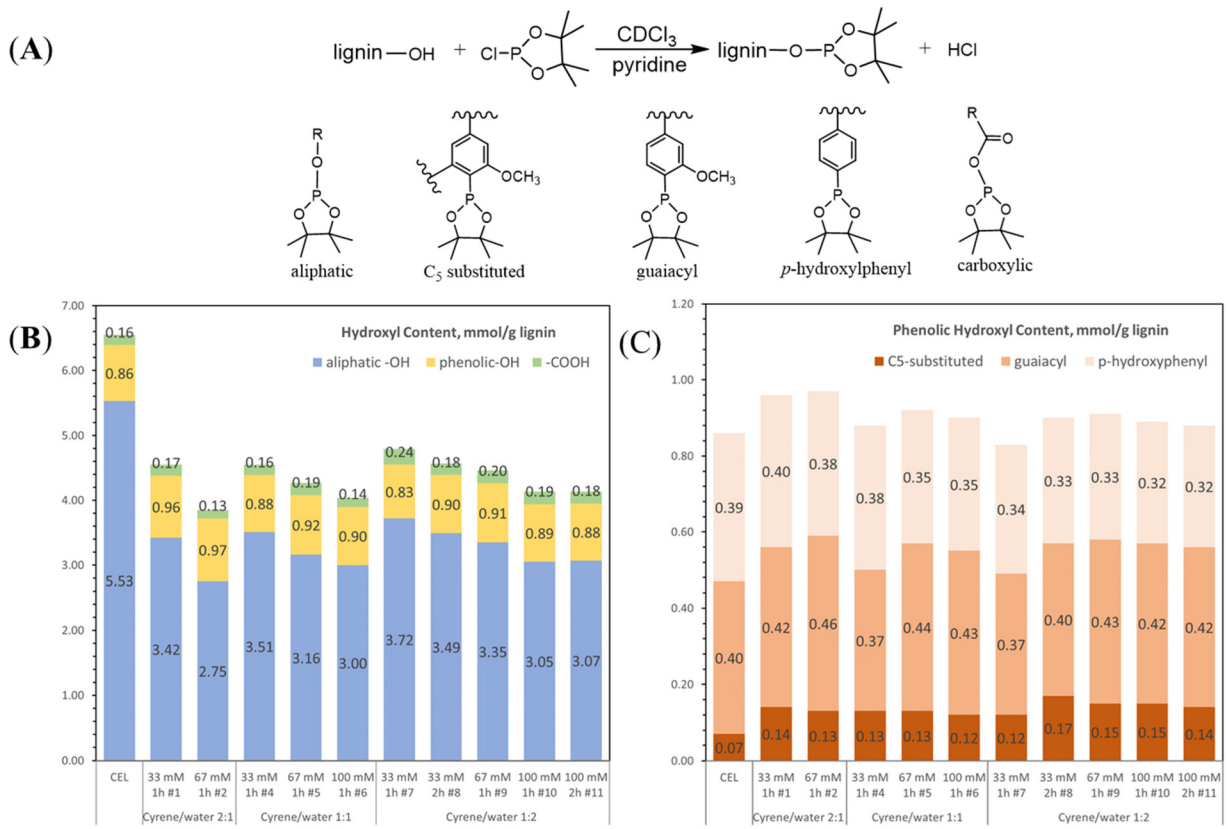


Fig. 3 Quantitative ^{31}P NMR analysis of switchgrass CEL and Cyrene lignin. (A) Phosphitylation of lignin hydroxyl groups with 2-chloro-4,4,5,5 tetramethyl-1,3,2-dixoaphospholane (TMDP), (B) contents of aliphatic, phenolic, and carboxylic hydroxyl groups, (C) composition of phenolic hydroxyl groups.

C_5-substituted hydroxyl group was doubled compared to CEL as the syringyl subunits (C_5-substituted by methoxyl group) were exclusively connected with β-O-4 ether.

HSQC NMR has been the most common technique used to study the relative abundance of inter-unit linkages (β-O-4 ether, β–β resinol and β-5 phenylcoumaran), subunits (guaiacyl, syringyl, *p*-hydroxyphenyl) and end-groups (*p*-coumarate and ferulate) of lignin macromolecules. The semiquantitative analysis of lignin subunits and inter-unit linkages are shown in Fig. 4A and B, while example 2D HSQC NMR spectra can be found in Fig. S2.† β-O-4 ether is the most abundant inter-unit linkage found in switchgrass CEL (Fig. 4A). The relative abundance of β-O-4 linkage in CEL was 48/100 aromatic ring (AR). This chemically liable linkage was well-preserved after pretreatment #1 (45/100 AR), #7 (46/100 AR), and #8 (43/100 AR). All three pretreatments were conducted with 33 mM sulfuric acid. The β-O-4 content decreased with increasing acid concentration, as can be seen from pretreatment #2 (29/100 AR), #5 (36/100AR), and #9 (21/100 AR). The ratio between Cyrene and water also affects the lignin structure, as shown in Fig. 4A. Under the same acid loading, pretreatment with Cyrene/water 1 : 1 resulted in lower β-O-4 linkage retention than pretreatment using Cyrene/water 2 : 1 and 1 : 2, while the β-O-4 content was less affected by increasing acid loading in Cyrene/water 1 : 2 than in the 2 : 1 mixture. The cleavage of ether bonds was the main reaction during acid-catalyzed delignification; C–C bonded inter-unit linkages such as β–β resinol and β-5 phenylcoumaran, are more difficult to be degraded.[39] As shown in Fig. 4A, the relative abundance of β–β resinol and β-5 phenylcoumaran linkage in Cyrene pretreated switchgrass lignin saw a slight increase compared to CEL. The frequency of β-O-4 ether inter-unit decreased as the pretreatment severity was leveled up by increasing duration time and/or H_2SO_4 concentration using a fixed Cyrene/water ratio. Yet, increasing Cyrene volume fraction seems to be able to effectively improve the rate of β-O-4 ether removal under the same reaction conditions (*i.e.* duration time and H_2SO_4 concentration). Nevertheless, the β-O-4 retention is inversely related to delignification, especially in the case of Cyrene/water 2 : 1, where delignification increased from 61.8% to 73.3%, and the relative abundance of β-O-4 linkage decreased from 45/100 AR to 29/100 AR. This phenomenon was less pronounced when the ratio between Cyrene and water was 1 : 2, as the β-O-4 content was 40/100 AR when the highest delignification under this condition was achieved (54.9%, pretreatment #9). Overall, pretreatment with Cyrene/water 2 : 1 led to 60–73% delignification in 1 h at 120 °C, while 29-45/100 AR of β-O-4 linkage was retained in the lignin structure, which is much higher than pretreatment with ethanol or dilute acid,[40,41] and comparable with pretreatments using deep eutectic solvents,[42] demonstrating the potential of Cyrene as a solvent for lignin extraction. To overcome the barrier between delignification and lignin structural integrity, future studies with Cyrene might consider the integration of stabilization strategies.

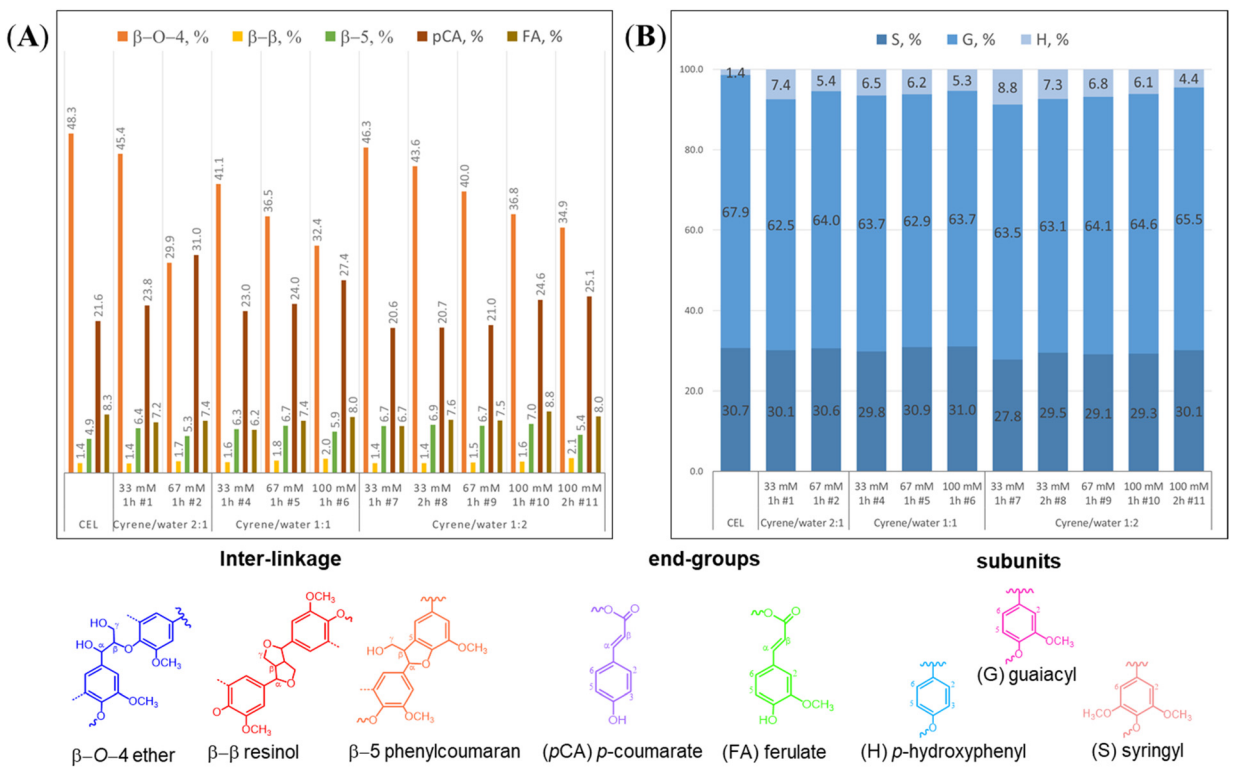


Fig. 4 Semiquantitative HSQC analysis of switchgrass CEL and Cyrene lignin. Frequencies of lignin inter-unit linkages (β-O-4 ether, β–β resinol and β-5 phenylcoumaran) and end-groups (*p*-coumarate and ferulate) (A), and subunits (guaiacyl, syringyl, *p*-hydroxyphenyl) (B) per 100 (S + G + H) units.

A similar observation was documented in the change of *p*-courmarate content in switchgrass lignin in response to the Cyrene pretreatment condition: *p*-courmarate ester tended to be recalcitrant under elevated pretreatment severity, but the impact of the latter effect was weakened by reducing Cyrene volume fraction. It was reported that *p*-courmarate end-group were incorporated into grass lignin mainly *via* acetylation with the aliphatic side chain γ-OH, and *p*-courmaration preferred to occur on S subunits.[43] The subunits composition shown in Fig. 4B indicated that G subunits seemed to be the major contributor to the loss of aromatic rings during mild Cyrene pretreatment. The portion of H subunits underwent an initial boost, and then decreased as lignin depolymerization proceeded.

Since lignin is the largest renewable source for platform aromatic chemicals and current strategies to derive such value-added products from lignin involves depolymerization, the amount of liable ether linkages like the β-O-4 linkages in lignin becomes a crucial factor.[44] Shuai *et al.* have demonstrated that by preserving the lignin structure and preventing condensation, near theoretical monomer yields could be obtained after depolymerization.[45] Though the lignin structure could be well-preserved after the current Cyrene pretreatment, the ratio between Cyrene/water, pretreatment temperature and time still need tuning to achieve both high delignification and lignin recovery yield along with structurally well-preserved lignin.

Molecular dynamics examination of lignin aggregates, isolated chains, and model dimers in Cyrene and Cyrene/water mixtures

Molecular dynamics (MD) simulations are a powerful tool for gaining insights into solvent–polymer interactions and biomass pretreatments.[46–49] Recent work by Mohan *et al.* provided an account of lignin–Cyrene and lignin–Cyrene : water interactions using a combination of COSMO-RS and all-atom molecular dynamics simulations and suggested that Cyrene has a preferential interaction with lignin when in mixtures with water and caused isolated lignin chains to adopt random-coil like conformation.[50] However, the prior work made use of non-optimized force-field for Cyrene obtained from CGenFF, modelled a generic lignin, did not consider inter-lignin chain interactions, *i.e.* the behaviour of lignin aggregates, and performed MD simulations of only two of the experimental ratios considered here. Here, simulations were performed with an updated parameterization of Cyrene (see ESI† for details) and the influence of Cyrene, water, and Cyrene : water mixtures on the behaviour of both (low-molecular weight) switchgrass-like lignin aggregates and isolated switchgrass-like lignin chains. Additional, simulations with softwood-like lignin were also performed and are reported in the ESI.† Fig. 5 summarizes the results of simulations of pre-formed switchgrass-lignin aggregates under Cyrene, mixtures of Cyrene with water (4 : 1 v/v, 2 : 1 v/v, 1 : 1 v/v, 1 : 2 v/v), and bulk water at elevated tempera-

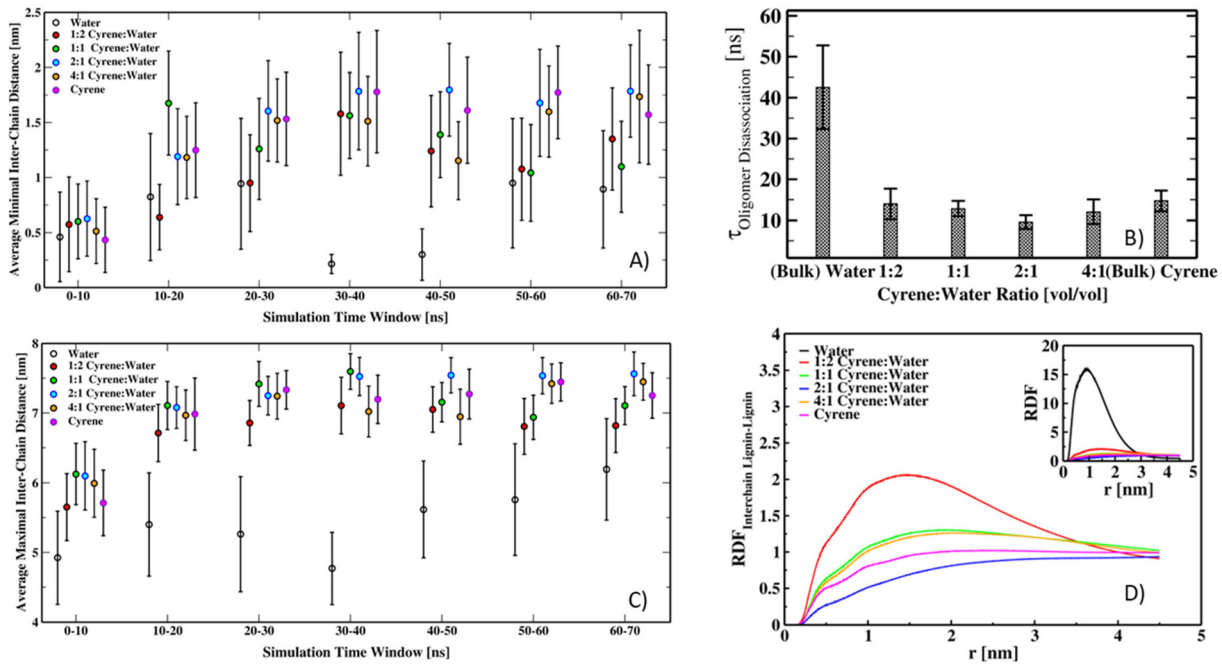


Fig. 5 Lignin aggregate characterization. (A) Average minimal distances between inter-lignin chains obtained over 10 ns windows from 5 independent simulations. (B) Bar-graph representation of mean first exit times for the disassociation of a lignin chain from other lignins. A disassociation event is taken as the time when the minimal distance to any other lignin chains is more than 1.5 nm. (C) Average maximal distances between inter-lignin chains obtained over 10 ns windows from 5 independent simulations. (D) Inter-chain, monomer-centered, lignin–lignin radial distribution functions; radial distribution functions were calculated using the last 25 ns of each of the five-independent 70 ns trajectories per solvent condition. Error-bars presented in subfigures are standard deviations.

tures (~393 K). Fig. 5A and C report the average minimal and maximal distances in 10 ns time-windows between lignin chains, respectively. These sub-figures clearly show that in the presence of Cyrene and Cyrene : water mixtures, lignin aggregates are disrupted, with minimal inter-chain distances exceeding 1.5 nm within 30 ns of simulation time. By setting a dissociation distance threshold of having two independent chains be more than 1 nm apart, and tracking the "first exit times" for such an event to occur, Fig. 5B provides an alternative quantification of aggregate disruption under the varied solvent conditions. Interestingly, the average first exit times show that for the tested solvents, 2 : 1 Cyrene : water mixtures have the fastest disruption of lignin aggregates. Furthermore, by calculating the atomic inter-chain radial distribution functions of the lignin chains over the last 30 ns of the simulations (Fig. 5D), it is clear that the interchain correlations are similar, though not identical, to bulk solutions, with the exceptions of lignins in bulk water and 1 : 2 Cyrene : water environments, where aggregates may spontaneously regenerate.

While the aggregate simulations provide clear evidence that when lignin is placed in environments consisting of bulk Cyrene or Cyrene : water mixtures with moderate Cyrene present, these simulations are limited in providing a fair comparison of the impact of these conditions on isolated lignin chains and an idealized view of lignin–cosolvent interactions. To circumvent the aforementioned limitation, simulations of isolated lignin were performed, as outlined in Methods, under bulk Cyrene, bulk water, and 2 : 1 v/v, 1 : 1 v/v, and 1 : 2 v/v Cyrene : water stresses. Fig. 6 summarizes the impact of these solvent conditions on lignin morphology and short-range (Lennard-Jones) intra-chain interaction strengths. As expected, given the previously observed trends to that Cyrene and Cyrene-enriched environments result in the disruption of lignin aggregates, isolated lignin chains adopt expanded, open-chain conformations, as evidenced by the shifts, compared to bulk water, towards equilibrium geometries with increased radius of gyration and solvent accessible surface area measures (as computed using the GROMACS gmx gyrate and sasa tools). Further, the energetics of intra-chain short-range (Lennard-Jones) interactions, obtained using the gmx energy tool and reported in Fig. 6C, indicate that in the tested Cyrene containing environments, lignin–lignin self-interactions are screened. While all Cyrene containing conditions tested are found to be disruptive, similar to the aggregate simulations, the largest changes in (single chain) lignin morphology and self-interactions are found for systems containing 2 : 1 v/v Cyrene : water.

The relative positioning of the cosolvents themselves (Fig. 7A & B), as shown by the per monomer center of mass-cosolvent molecule center of mass radial distribution functions (obtained from the last 10 ns of each of the 30 ns isolated lignin chain production simulations) show that, in the presence of water, Cyrene is readily enriched near lignin. Additionally, it is interesting to note that for all the tested mixtures, except for the 2 : 1 v/v Cyrene : water condition, water is generally displaced from the local lignin environment. Exploring the short-range contact (Lennard-Jones) energetics of the cosolvent components with lignin (Fig. 7C & D) indicate

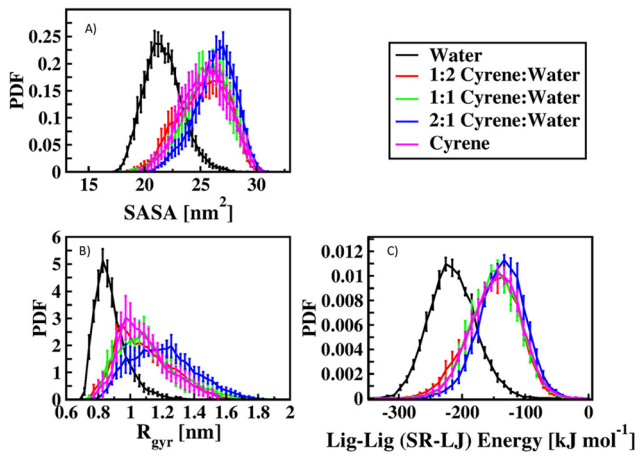


Fig. 6 Isolated lignin chain characteristics under water, Cyrene, and cosolvent conditions. (A) Mean lignin solvent accessible surface area distributions. (B) Mean lignin radius of gyration distributions. (C) Mean short-range intra-chain (lignin–lignin) Lennard-Jones interaction energy distributions. All distributions were obtained by histogramming calculated features over the final 10 ns of each (five) independent trajectory for each of the solvent conditions. Error-estimates are standard-errors for each bin assignment. Lennard-Jones energies are used to approximate the strength and number of close intra-chain contacts. This is necessarily an incomplete measure of contact energies as it excludes electrostatic contributions though it does provide a window into potential contact screening.

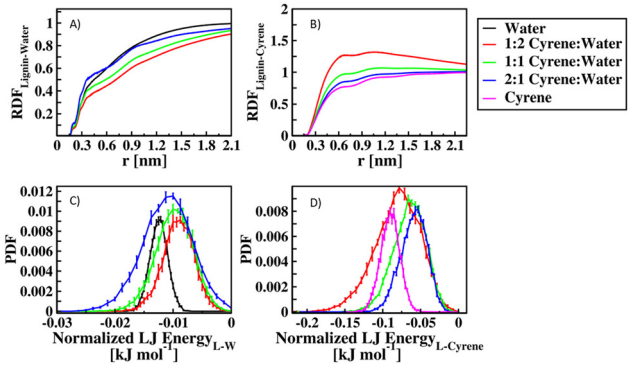


Fig. 7 Isolated lignin chain-(co)solvent interactions. (A) Mean lignin monomer center-of-mass-water molecule center-of-mass radial distribution functions. (B) Mean lignin monomer center-of-mass-water molecule center-of-mass radial distribution functions. All presented data was obtained from the last 10 ns of each of the five independent simulations per solvent condition. (C) Normalized (by total number of water molecules in simulation box) Lignin–water short-range (within 1.2 nm cutoff) Lennard-Jones interaction energy distributions. (D) Normalized (by total number of water molecules in simulation box) lignin–Cyrene short-range (within 1.2 nm cutoff) Lennard-Jones interaction energy distributions. As before in Fig. 6, Lennard-Jones energies are used to approximate the strength and number of close intra-chain contacts and is necessarily an incomplete measure of contact energies as it excludes electrostatic contributions; however, the choice still provides a window into short-range contacts.

that when normalized by the number of each solvent type in the system (*i.e.* number of water or Cyrene molecules) water–lignin interactions are indeed screened in the presence of Cyrene; however, this screening is incomplete for 2 : 1 v/v mixtures. Additionally, the strength of lignin–Cyrene interactions are themselves reduced in the tested cosolvent systems containing water. From the radial distribution functions presented in Fig. 7A & B, along with the changes in the morphology and screening of lignin intra-chain self-interactions/contracts reported in Fig. 6 and disruption of lignin aggregates (Fig. 5), it may be inferred that Cyrene : water cosolvents act in a similar manner to THF : Water cosolvent mixtures, where one of the cosolvent components has a preferential solute–solvent interaction with lignin, resulting in a potentially theta-like, localized solvent environment.[49]

Taking the above results together, it is possible to devise a ranking of cosolvent conditions by simulation derived disaggregation first exit times (2 : 1 < 4 : 1–1 : 1 < 1 : 2 ∼ bulk Cyrene < bulk water) or degree of morphological differences from lignin morphologies found under bulk water (2 : 1 > 4 : 1–1 : 1 > 1 : 2 ∼ bulk Cyrene > bulk water) and from this it is clear that the simulations do find the same optimal cosolvent mixture as that reported found experimentally. However, given the similarity of the Cyrene : water cosolvent's apparent mechanism of action to that of THF : water cosolvent's mechanism, it is, at first glance, quite surprising that the experimental measures of lignin release and deconstruction show such sensitivity to acid content and pretreatment length. One potential explanation for this is that the presence of Cyrene near the lignin solvation shell merely encourages physical disruption, but inhibits chemical reactions, as seen by the reduction of water in the first solvation shell for all tested Cyrene cosolvent mixtures, except for the 2 : 1 v/v mixture which has a minor enrichment of 'trapped' water near the lignin surface as seen from Fig. 7A. While short-range (Lennard-Jones) lignin–water interactions energies on average were screened at the 2 : 1 v/v concentration, compared to other cosolvent mixtures (1 : 1 v/v and 1 : 2 v/v) this screening was much weaker, with a substantial portion of strong water–lignin contacts still possible, as evident by the large left-ward shoulder in lignin–water Lennard-Jones energies. Considering that the first exit times for lignin disaggregation are fastest in the 2 : 1 v/v Cyrene : water mixture, it may be posited that the observed weakening of Cyrene–lignin interactions at 2 : 1 v/v ratio may 'lubricate' the solvation shell which prevents Cyrene's inherent viscous nature from slowing polymer reorganization, as is found in the bulk Cyrene exit times. It is important to caution that this analysis is excluding electrostatics and is using the short-range Lennard-Jones interaction energies merely as a proxy for the ability of cosolvents to make energetically favorable short-range contacts with lignin.

To better quantify whether the solvation of the experimentally most cleaved linkage (*i.e.*, β-O-4) is substantially altered under the identified optimal condition (2 : 1 v/v Cyrene : water) in comparison to bulk water, we performed simulations of a single β-O-4 dimer with and without "classical" reactive species models. We reiterate that reactions are not explicitly modelled here; the aim is to primarily examine how acid moieties and water may coordinate around the α-O-4 atom in a pre-attack state in 2 : 1 v/v Cyrene : water and bulk water. Fig. 8 summarizes the results from these dimer simulations. The cumulative radial distribution functions (*i.e.* running coordination number function) indicates that in the event that the α-O-4 site is exposed, as would be promoted under Cyrene : water environments, water and (the classical) reactive acidic species (HSO_4^-) can localize near α-O-4, which may facilitate β-O-4 cleavage; however, this localization is reduced in the 2 : 1 v/v Cyrene : water condition compared to an idealized water environment; with waters being largely displaced and coordination of the reactive anion being slightly reduced. When this screening is taken together with the observation that water and reactive species diffusion dynamics are substantially slower in the 2 : 1 v/v Cyrene : water system compared to bulk water (Fig. 8B & D), the data suggests that while Cyrene–water cosolvents may readily alter the physical state of lignin chains and aggregates, the diffusion of reactive species and localization of these species near linkages sites is still retarded compared to an idealized water environment, and that this in turn may inhibit chemical processes. Taken together, this may explain the experimental observation that it is necessary to increase the severity of the reaction conditions, either by increasing reactive species content or reaction time,

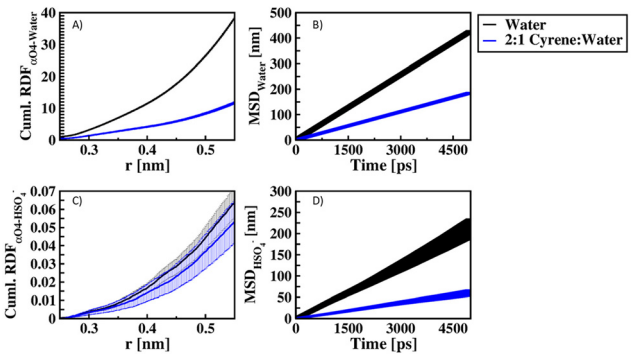


Fig. 8 Model lignin-dimer water and HSO_4^- coordination and dynamics characterization. (A) Mean cumulative radial distribution function between water and the model dimeric lignin's α-O-4 oxygen in proximity to the model β-O-4 linkage. (B) Average mean squared displacements of water molecules within the simulations over a time window of 5 ns. (C) Mean cumulative radial distribution function between HSO_4^- and the model dimeric lignin's α-O-4 oxygen in proximity to the model β-O-4 linkage. (D) Average mean squared displacements of water molecules within the simulations over a time window of 5 ns. The mean cumulative radial distributions (cRDFs) were computed over the final 5 ns of each of 20 independent 15 ns production simulations per solvent conditions. To facilitate equivalent binning during the calculations of the cRDFs, the cRDFs were only computed up-to a maximum distance of 2.1 nm. Error-bars in subfigures (A) & (C) represent the standard-error for each distance bin in the cRDFs. Subfigure (A)'s errors are roughly equivalent to the thicknesses of the shown lines. Errors-bars in subfigures (B) & (D) are standard errors for each lag-time over the 20 independent trajectories. The entire 15 ns trajectories were used to estimate the MSD profiles in (B) & (D).

to obtain substantial changes in lignin molecular cleavage under these pretreatments.

Conclusions

Mild Cyrene pretreatment has advantages in preserving lignin integrity with relatively high delignification. In this work, to overcome the high viscosity nature of Cyrene, the pretreatment on switchgrass was conducted using a low volume fraction of Cyrene (Cyrene/water = 2 : 1, 1 : 1 and 1 : 2) under selected conditions. It was demonstrated that Cyrene pretreatment could effectively break down the recalcitrant switchgrass cell-wall matrix, reducing lignin molecular weight by 20–60% and increasing the cellulose conversion to up to ~70% depending on the pretreatment conditions. At 120 °C, higher Cyrene volume fraction (Cryene/water = 2 : 1) yielded better lignin removal and cellulose conversion rate, yet the efficacy of Cyrene pretreatment could be compensated by prolonging duration time and adopting high catalyst dosage. The NMR analyses indicated that the switchgrass lignin macromolecules underwent mild modification with high preservation of β-O-4 ether inter-unit linkages and functional groups during Cyrene pretreatment, which could benefit integrating lignin valorization into modern biorefinery. Molecular dynamics simulations probed the possible molecular mechanisms behind the efficiency of the Cyrene–water cosolvent mixtures and found that Cyrene readily disrupts lignin aggregate structures and modulates intra-chain lignin–lignin contacts and results in lignin adopting expanded "open" states. Additionally, simulation analysis of lignin solvation suggests that Cyrene has a preferential interaction with lignin. Interestingly, if we rank the cosolvent conditions based on the simulation-based prediction of lignin aggregate disruption, our simulations indicate that 2 : 1 v/v mixtures of Cyrene : water, similar to experiment, will be optimal. Simulations comparing the localization of classical reactive-species stand-ins in water and 2 : 1 v/v Cyrene : water showed that water, and to a very minor extent anionic reactive species, would be screened near lignin, compared to an idealized water environment. When taken together with reduced solvent dynamics, this may explain the necessity to make use of either longer reaction times or higher reaction severity to generate yields of low molecular-weight lignin under Cyrene : water pretreatments. It should be noted that alternative explanations behind the need for longer reaction times or increased reaction severity (acid loadings) may exist; while the current modeling appears consistent with the observed experimental trends, it is important to note that the presented computational results do not explicitly treat reaction events, and as a result the role of acid in modulating the interconversion of Cyrene with its transient germinal diol form within the lignin solvation shell and what role this may play in retarding the cleavage of inter-unit linkages remains unknown. Future studies will likely be necessary to obtain a complete understanding of this promising and complex cosolvent pretreatment.

Author contributions

Yun-Yan Wang: investigation, writing – original draft, writing – review & editing. Yunxuan Wang: investigation, writing – original draft, writing – review & editing. Luna Liang: investigation, writing – review & editing. Micholas Dean Smith: investigation, writing – review & editing. Xianzhi Meng: writing – review & editing. Yunqiao Pu: conceptualization, methodology, supervision, writing – review & editing. Mitra Mazarei: resources, writing – review & editing. Rupesh Agarwal: writing – review & editing. Shalini Jayaraman Rukmani: writing – review & editing. Brian H. Davison: writing – review & editing, funding acquisition. Arthur J. Ragauskas: conceptualization, supervision, writing – review & editing, funding acquisition.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We are grateful for funding from the Genomic Science Program, Office of Biological and Environmental Research, Office of Science, US Department of Energy (DOE), under Contract FWP ERKP752. This research used the Oak Ridge Leadership Computing Facility resources at the Oak Ridge National Laboratory under an INCITE award, which is supported by the Office of Science of the U.S. Department of Energy along with the computational resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory supported by the Office of Science of the U.S. Department of Energy. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the US DOE under Contract DE-AC05-00OR22725. Additional portions of the computations for this work were performed on the computational resources at the Infrastructure for Scientific Applications and Advanced Computing (ISAAC) supported by the University of Tennessee and the National Energy Research Scientific Computing Center (NERSC) under the ALCC award ERCAP0025417, project ID m4196.

References

1 A. J. Ragauskas, C. K. Williams, B. H. Davison, G. Britovsek, J. Cairney, C. A. Eckert, W. J. Frederick, J. P. Hallett, D. J. Leak, C. L. Liotta, J. R. Mielenz, R. Murphy, R. Templer and T. Tschaplinski, *Science*, 2006, **311**, 484–489.
2 K. David and A. J. Ragauskas, *Energy Environ. Sci.*, 2010, **3**, 1182–1190.
3 M. Li, Y. Pu, C. G. Yoo, E. Gjersing, S. R. Decker, C. Doeppke, T. Shollenberger, T. J. Tschaplinski, N. L. Engle, R. W. Sykes, M. F. Davis, H. L. Baxter, M. Mazarei, C. Fu, R. A. Dixon, Z.-Y. Wang, C. Neal Stewart Jr and A. J. Ragauskas, *Biotechnol. Biofuels*, 2017, **10**, 12.
4 L. Alexander, C. Hatcher, M. Mazarei, E. Haynes, H. Baxter, K. Kim, C. Hamilton, R. Sykes, G. Turner, M. Davis, Z.-Y. Wang, N. Labbé and C. Neal Stewart Jr, *Euphytica*, 2020, **216**, 25.
5 J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P. F. Schatz, J. M. Marita, R. D. Hatfield, S. A. Ralph, J. H. Christensen and W. Boerjan, *Phytochem. Rev.*, 2004, **3**, 29–60.
6 N. Giummarella, Y. Pu, A. J. Ragauskas and M. Lawoko, *Green Chem.*, 2019, **21**, 1573–1595.
7 K. H. Caffall and D. Mohnen, *Carbohydr. Res.*, 2009, **344**, 1879–1900.
8 H. G. Jung and R. L. Phillips, *Crop Sci.*, 2010, **50**, 403–418.
9 G.-S. Nguyen, A. S. Lewin, F. Di Bartolomeo and A. Wentzel, *Enzymes for Solving Humankind's Problems*, 2021, pp. 439–471.
10 J. Baruah, B. K. Nath, R. Sharma, S. Kumar, R. C. Deka, D. C. Baruah and E. Kalita, *Front. Energy Res.*, 2018, **6**, 141.
11 L. Shuai, Y. M. Questell-Santiago and J. S. Luterbacher, *Green Chem.*, 2016, **18**, 937–943.
12 C. M. Cai, T. Zhang, R. Kumar and C. E. Wyman, *Green Chem.*, 2013, **15**, 3140–3145.
13 S. X. Li, M. F. Li, J. Bian, S. N. Sun, F. Peng and Z. M. Xue, *Bioresour. Technol.*, 2017, **243**, 1105–1111.
14 X. Meng, Y. Pu, M. Li and A. J. Ragauskas, *Green Chem.*, 2020, **22**, 2862–2872.
15 J. Sherwood, M. De bruyn, A. Constantinou, L. Moity, C. R. McElroy, T. J. Farmer, T. Duncan, W. Raverty, A. J. Hunt and J. H. Clark, *Chem. Commun.*, 2014, **50**, 9650–9652.
16 L. Hughes, C. R. McElroy, A. C. Whitwood and A. J. Hunt, *Green Chem.*, 2018, **20**, 4423–4427.
17 R. A. Milescu, A. Zhenova, M. Vastano, R. Gammons, S. Lin, C. H. Lau, J. H. Clark, C. R. McElroy and A. Pellis, *ChemSusChem*, 2021, **14**, 3367–3381.
18 A. Marathianos, E. Liarou, E. Hancox, J. L. Grace, D. W. Lester and D. M. Haddleton, *Green Chem.*, 2020, **22**, 5833–5837.
19 J. E. Camp, S. B. Nyamini and F. J. Scott, *RSC Med. Chem.*, 2020, **11**, 111–117.
20 N. A. Stini, P. L. Gkizis and C. G. Kokotos, *Org. Biomol. Chem.*, 2023, **21**, 351–358.
21 A. Duval and L. Avérous, *Green Chem.*, 2022, **24**, 338–349.
22 D. O. Abranches, J. Benfica, S. Shimizu and J. A. P. Coutinho, *Ind. Eng. Chem. Res.*, 2020, **59**, 18649–18658.
23 T. Renders, S. Van den Bosch, S. F. Koelewijn, W. Schutyser and B. F. Sels, *Energy Environ. Sci.*, 2017, **10**, 1551–1557.
24 M. M. Abu-Omar, K. Barta, G. T. Beckham, J. S. Luterbacher, J. Ralph, R. Rinaldi, Y. Román-Leshkov, J. S. M. Samec, B. F. Sels and F. Wang, *Energy Environ. Sci.*, 2021, **14**, 262–292.
25 Y.-Y. Wang, M. Li, C. E. Wyman, C. M. Cai and A. J. Ragauskas, *ACS Sustainable Chem. Eng.*, 2018, **6**, 6064–6072.
26 Y.-Y. Wang, P. Sengupta, B. Scheidemantle, Y. Pu, C. E. Wyman, C. M. Cai and A. J. Ragauskas, *Front. Energy Res.*, 2020, **8**, 149.
27 D. S. Zijlstra, A. de Santi, B. Oldenburger, J. de Vries, K. Barta and P. J. Deuss, *J. Visualized Exp.*, 2019, **143**, e58575.
28 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1–2**, 19–25.
29 L. Petridis and J. C. Smith, *J. Comput. Chem.*, 2009, **30**, 457–467.
30 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
31 H. J. Berendsen, J. v. Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.
32 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.
33 B. Hess, *J. Chem. Theory Comput.*, 2008, **4**, 116–122.
34 C. Li, B. Knierim, C. Manisseri, R. Arora, H. V. Scheller, M. Auer, K. P. Vogel, B. A. Simmons and S. Singh, *Bioresour. Technol.*, 2010, **101**, 4900–4906.
35 A. Duval and L. Avérous, *Green Chem.*, 2022, **24**, 338–349.
36 Y. Pu, B. Hallac and A. J. Ragauskas, in *Aqueous Pretreatment of Plant Biomass for Biological and Chemical Conversion to Fuels and Chemicals*, 2013, DOI: **10.1002/9780470975831.ch18**.
37 X. Meng, C. Crestini, H. Ben, N. Hao, Y. Pu, A. J. Ragauskas and D. S. Argyropoulos, *Nat. Protoc.*, 2019, **14**, 2627–2647.
38 X. Meng, B. R. Evans, C. G. Yoo, Y. Pu, B. H. Davison and A. J. Ragauskas, *ACS Sustainable Chem. Eng.*, 2017, **5**, 8004–8010.
39 M. R. Sturgeon, S. Kim, K. Lawrence, R. S. Paton, S. C. Chmely, M. Nimlos, T. D. Foust and G. T. Beckham, *ACS Sustainable Chem. Eng.*, 2014, **2**, 472–485.
40 L. Liang, Y.-Y. Wang, S. Bhagia, V. Sethuraman, Z. Yang, X. Meng, N. Bryant, L. Petridis, J. C. Smith, S. V. Pingali, N. C. Gallego, Y. Pu, B. R. Evans, H. M. O'Neill, B. H. Davison and A. J. Ragauskas, *ACS Sustainable Chem. Eng.*, 2022, **10**(28), 9041–9052.

41 Y. Pu, F. Hu, F. Huang, B. H. Davison and A. J. Ragauskas, *Biotechnol. Biofuels*, 2013, **6**, 15.
42 X.-J. Shen, J.-L. Wen, Q.-Q. Mei, X. Chen, D. Sun, T.-Q. Yuan and R.-C. Sun, *Green Chem.*, 2019, **21**, 275–283.
43 M. J. Rosado, F. Bausch, J. Rencoret, G. Marques, A. Gutiérrez, T. Rosenau, A. Potthast and J. C. del Río, *Ind. Crops Prod.*, 2021, **174**, 114226.
44 S. Van den Bosch, W. Schutyser, R. Vanholme, T. Driessen, S. F. Koelewijn, T. Renders, B. De Meester, W. J. J. Huijgen, W. Dehaen, C. M. Courtin, B. Lagrain, W. Boerjan and B. F. Sels, *Energy Environ. Sci.*, 2015, **8**, 1748–1763.
45 L. Shuai, M. T. Amiri, Y. M. Questell-Santiago, F. Héroguel, Y. Li, H. Kim, R. Meilan, C. Chapple, J. Ralph and J. S. Luterbacher, *Science*, 2016, **354**, 329–333.
46 P. Moyer, M. D. Smith, N. Abdoulmoumine, S. C. Chmely, J. C. Smith, L. Petridis and N. Labbe, *Phys. Chem. Chem. Phys.*, 2018, **20**, 2508–2516.
47 M. D. Smith, L. Petridis, X. Cheng, B. Mostofian and J. C. Smith, *Phys. Chem. Chem. Phys.*, 2016, **18**, 6394–6398.
48 J. V. Vermaas, M. F. Crowley and G. T. Beckham, *ACS Sustainable Chem. Eng.*, 2020, **8**, 17839–17850.
49 M. D. Smith, B. Mostofian, X. Cheng, L. Petridis, C. M. Cai, C. E. Wyman and J. C. Smith, *Green Chem.*, 2016, **18**, 1268–1277.
50 M. Mohan, K. L. Sale, R. S. Kalb, B. A. Simmons, J. M. Gladden and S. Singh, *ACS Sustainable Chem. Eng.*, 2022, **10**, 11016–11029.