

Review

Not peer-reviewed version

---

# Computational Advances in Ionic Liquid Applications for Green Chemistry: A Critical Review of Lignin Processing and Machine Learning Approaches

---

Brian Taylor , [Nikhil Kumar](#) , Dhirendra Kumar Mishra , [Blake A Simmons](#) , [Hemant Choudhary](#) <sup>\*</sup> , [Kenneth L Sale](#) <sup>\*</sup>

Posted Date: 8 October 2024

doi: 10.20944/preprints202410.0430.v1

Keywords: lignin; ionic liquids; quantum chemistry; machine learning; biorefineries



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

# Computational Advances in Ionic Liquid Applications for Green Chemistry: A Critical Review of Lignin Processing and Machine Learning Approaches

Brian R. Taylor <sup>1,2,†</sup>, Nikhil Kumar <sup>1,2,†</sup>, Dhirendra Kumar Mishra <sup>1,2,†</sup>, Blake A. Simmons <sup>1,4</sup>, Hemant Choudhary <sup>1,3,\*</sup> and Kenneth L. Sale <sup>1,2,\*</sup>

<sup>1</sup> Joint BioEnergy Institute, Emeryville, CA, 94608, USA

<sup>2</sup> Department of Biosecurity and Bioassurance, Sandia National Laboratories, Livermore, CA, 94551, USA

<sup>3</sup> Department of Bioresource and Environmental Security, Sandia National Laboratories, Livermore, CA 94550, USA

<sup>4</sup> Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA

\* Correspondence: hchoudhary@lbl.gov or hchoudh@sandia.gov (H.C.); klsale@lbl.gov or klsale@sandia.gov (K.L.S.)

† BRT, NK, and DKM contributed equally.

**Abstract:** The valorization and dissolution of lignin using ionic liquids (ILs) is critical for developing sustainable biorefineries and a circular bioeconomy. This review aims to critically assess the current state of computational and machine learning methods for understanding and optimizing IL-based lignin dissolution and valorization processes reported since 2022. The paper examines various computational approaches, from quantum chemistry to machine learning, highlighting their strengths, limitations, and recent advances in predicting and optimizing lignin-IL interactions. Key themes include the challenges in accurately modeling lignin's complex structure, the development of efficient screening methodologies for ionic liquids to enhance lignin dissolution and valorization processes, and the integration of machine learning with quantum calculations. These computational advances will drive progress in IL-based lignin valorization by providing deeper molecular-level insights and facilitating the rapid screening of novel IL-lignin systems.

**Keywords:** lignin; ionic liquids; quantum chemistry; machine learning; biorefineries

## 1. Introduction

The need for sustainable alternatives to fossil-based resources has driven significant interest in lignocellulosic biomass as a renewable feedstock for biofuels and chemicals. Lignin, which constitutes 15-35% [1] of this feedstock, offers significant potential as a source for high-value products such as biofuels [2], materials [3] and pharmaceuticals [4]. Unlocking this potential is critical to the economic viability of the biobased economy [5].

The valorization of lignin presents several unique challenges. It resists solubilization, depolymerization, and is chemically complex [2]. Furthermore, the structure and composition of lignin vary significantly depending on its biomass source and extraction method, complicating efforts to develop universally applicable conversion technologies. Most current biorefineries often combust the lignin rather than turn it into high-value products [6].

Ionic liquids (ILs) are promising solvents for lignin dissolution and processing, offering a potential solution to the challenges associated with lignin depolymerization. ILs are effective biomass solvents because of their thermostability, ability to dissolve a wide range of substances, tunable nature, and low vapor pressure [7–9]. Their tunability is due to the large number of possible cation-

anion combinations. However, it is costly, labor-intensive, and time-consuming to conduct experimental screenings of ILs due to the wide range of cation and anion combinations. Computational screening, on the other hand, has proven to be a viable option for multi-scale screening of “notable” solvents in biomass processing [10–13].

The vast improvement in computational resources provides an opportunity for computational studies to play a large role in identifying amenable ILs. However, significant challenges remain in applying computational methods to IL-based lignin valorization. These include accurately representing lignin’s structural heterogeneity, balancing computational cost with predictive accuracy, and translating molecular-scale insights to process-level optimization. Additionally, the scarcity of standardized, high-quality experimental data for training and validating computational models poses a barrier to developing robust predictive tools.

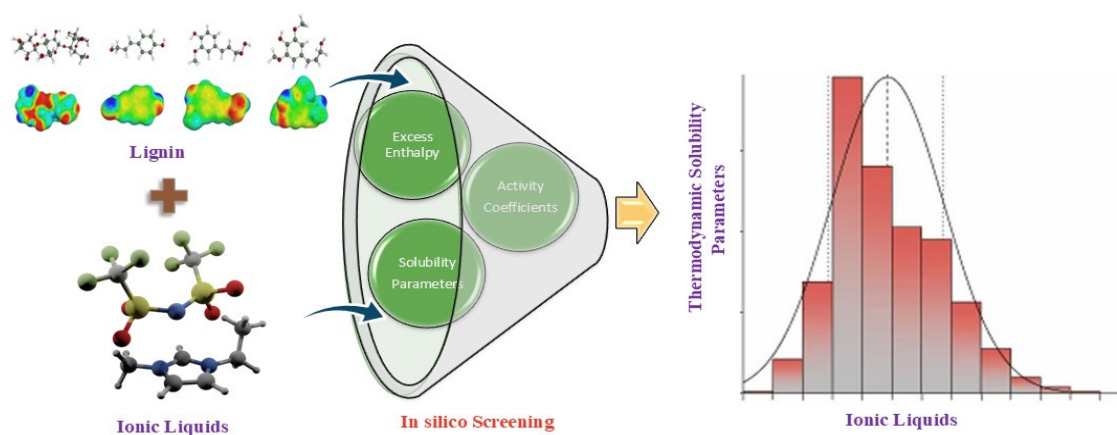
This review aims to critically assess the current state of computational methods for understanding and optimizing IL-based lignin dissolution and valorization processes reported since 2022. We examine the strengths and limitations of various approaches, from quantum chemistry to machine learning. The scope of this review excludes experimental methods, industrial-scale applications, and non-IL-based solvent systems. By highlighting recent advances and identifying key knowledge gaps, we seek to provide a roadmap for future research that will support the design of IL-based systems that allow for lignin dissolution and valorization, ultimately contributing to the development of sustainable biorefinery processes and a circular bioeconomy.

## 2. Screening of Ionic Liquids (ILs) Based on Modeling

Conducting experimental tests on all the potential cation-anion combinations in ILs is difficult due to their multitude. To effectively handle this level of complexity in ILs, computational screening methods are required. Because lignin is heterogeneous and has different bonding patterns, it is difficult to say how it will solvate. When simplified lignin models are used in simulations, they might not accurately represent how real lignin made from biomass behaves, which lowers the accuracy of screening data. Despite advances in IL screening for lignin solvation, there are no general rules to predict which ILs are most effective in deconstructing biomass. The absence of such guidelines leads to trial-and-error approaches and slows down the discovery of optimal solvents. Accurately predicting critical thermodynamic properties (such as activity coefficients, excess enthalpy, and solubility parameters) for IL-lignin systems is challenging, as these depend on specific interactions that vary widely among ILs and model lignin structures. Robust theoretical techniques are required for efficient screening of the extensive range of ionic liquids (ILs) for lignin dissolution. Recent improvements have facilitated the implementation of several predictive methodologies for simulating the thermodynamic parameters of ionic liquid-containing systems. Earlier prominent techniques encompass the Perturbed Chain Statistical Associating Fluid Theory (PC-SAFT) [14], which was used to calculate the thermodynamic properties of different homologous series of ILs based on the bis(trifluoromethylsulfonyl)imide anion ([NTf<sub>2</sub>]<sup>-</sup>), group contribution approaches [15] were explored for estimating critical properties, normal boiling temperatures and acentric factors for IL, Quantitative Structure-Property Relationships (QSPR) [16] were used to estimate the infinite dilution activity coefficients for organic solutes in IL, Monte Carlo (MC) molecular simulations [17] explored the understanding the gas solubility in IL, Molecular Dynamics (MD) simulations [18] performed to understand the pure and mixed gas absorption in the ionic liquids, and the Conductor-like Screening Model for Real Solvents (COSMO-RS) [19,20] model was applied for screening of solvents for multiple applications such as asphaltene dispersion, gas absorption and extraction of biologically active compounds [21–23]. It follows that the COSMO-RS model is one of the most effective predictive approaches for screening ILs for various applications due to its ability to quickly process vast IL datasets. Experimental investigations might be conducted with much less time, money, and effort if ILs could be screened using predictive techniques that only needed a small number of input parameters.

While emerging methods like COSMO-RS and machine learning (ML) have shown promise, scaling these models to screen thousands of ILs with diverse lignin models requires significant

computational resources and optimization to ensure practical, cost-effective screening [24,25]. Translating predictions into real-world applications is often hindered by the complexity of IL synthesis, cost, and the need for reliable experimental data to find the best solvents. COSMO-RS has previously been successfully used to predict thermodynamic properties such as solubility in IL systems, activity coefficients ( $\gamma$ ), and excess enthalpy (HE). Several research articles emphasize the ability of these parameters to screen ILs for lignin solvation [11,26,27]. The basic schematic diagram for in silico screening of ionic liquids for lignin solvation is shown in Figure 1. Various researchers used basic lignin model compounds, such as p-coumaryl alcohol, sinapyl alcohol, and coniferyl alcohol, to represent the lignin structure. The main links between the structural units of lignin are  $\beta$ -O-4 ( $\beta$ -aryl ether),  $\beta$ - $\beta$  (resinol), and  $\beta$ -5 (phenylcoumaran) [28]. As a result, researchers developed various lignin models, which consist of three basic units (H, G, and S) with some basic linkages, to investigate the impact of lignin structure on dissolving behavior. However, the current understanding of ILs and lignin models remains relatively basic, posing a challenge in formulating comprehensive guidelines for identifying the best solvents for lignin.



**Figure 1.** The schematic diagram for screening of solvents for lignin dissolution.

Mohan et al. conducted an extensive screening of 5670 ionic liquids (ILs), generated from 63 cations and 90 anions, to assess their effectiveness in lignin dissolution [11]. Lignin structure was built by joining all the major lignin linkages ( $\beta$ -O-4,  $\beta$ - $\beta$ , 4-O-5,  $\alpha$ -O-4, and  $\beta$ -5) present in the native lignin and screening results were validated using the experimental data. They identified that anions like acetate, methyl carbonate, glycinate, alaninate, and lysinate were particularly efficient in breaking down lignin, especially when paired with cations such as tetraalkylammonium ([TBA]<sup>+</sup>), tetraalkylphosphonium ([TBP]<sup>+</sup>), and pyridinium. Specifically, ILs containing [TBA]<sup>+</sup> and [TBP]<sup>+</sup> showed great promise as solvents due to their low viscosity, favorable Hansen solubility parameter alignment, and enhanced hydrogen-bonding basicity. Yu et al. further studied 3886 ILs using COSMO-RS and used the 19 model structures of Lignin (monomers, dimers, and trimers) for predictions and validated them through experimental solubility tests [24]. The monomeric, dimeric, and trimer structures of lignin fail to correctly represent the lignin molecule due to the absence of several different linkages found in lignin. Additionally, Kamlet-Taft parameters were used to clarify the mechanisms involved in lignin dissolution, highlighting the importance of IL screening in optimizing lignin processing. In further extension, Mohan and coworkers explored the molecular mechanisms behind the effectiveness of cholinium lysinate ([Ch][Lys]) IL in lignin processing [29]. Using MD simulations alongside COSMO-RS, they determined that [Ch][Lys] serves as a promising solvent for lignin dissolution. This preference is attributed to multiple factors, including increased molecular interactions, extensive hydrogen bonding, a higher dissociation constant, and reduced viscosity. These characteristics collectively enhance the efficiency of [Ch][Lys] in the dissolution of lignin.

#### *Limitations of Solvent Screening for Lignin*

**Complexity of Lignin Structure:** Most computational screening models simplify lignin's structure to make simulations feasible, but this can lead to inaccurate predictions. It is challenging to make a plausible model of how lignin interacts with ILs because it is a very heterogeneous polymer with many different functional groups and bonding patterns. Simplifying lignin into small model compounds (such as guaiacol or syringol) doesn't capture the full complexity of the polymer [24]. Future research should focus on improving computational power or algorithms to allow simulations of larger, more complex lignin structures without compromising accuracy. Its structure varies depending on the source and processing method. This structural variability makes it difficult to identify universal ILs for lignin dissolution and valorization. Future research should concentrate on tailoring ILs to specific lignin sources, and systematic studies are necessary to understand the impact of structural diversity on IL-lignin interactions.

**Solubility vs. Selectivity:** While some ILs are excellent at dissolving lignin, their selectivity toward different lignin fractions and derivatives remains a challenge. Current screening methods often prioritize solubility over selectivity, potentially resulting in inefficient lignin fractionation or subpar product quality. Future research should focus on ILs that are both highly soluble and selective for specific lignin components.

**Computational screening and predictive models:** While computational tools such as COSMO-RS and MD simulations can predict IL-lignin interactions, their accuracy remains limited. There is a need for improved algorithms and more comprehensive datasets that account for the variability in lignin structure and IL chemistry. Future efforts should aim at integrating machine learning models with quantum calculation data from first principle calculation to enhance screening efficiency. Fixing these issues will significantly advance future research in creating ILs for lignin valorization. This will result in biorefinery processes that work better and last longer. In the context of IL screening for lignin, computational screening and predictive models are becoming vital tools. However, to improve their accuracy and efficiency, we must address several challenges. One of the primary challenges in developing accurate predictive models for IL-lignin interactions is the scarcity of high-quality "standardized" experimental data. Most computational models, including machine learning (ML) approaches, require extensive datasets to predict IL performance. Because there isn't any standardized experimental data for how lignin dissolves or depolymerizes in different ILs, it's challenging to make reliable models. Future work should focus on generating large, consistent datasets that capture both the structural variability of lignin and the wide range of IL chemistry.

**Accuracy of force fields and interaction parameters:** MD simulations and other computer methods utilize pre-established force fields to illustrate the interactions between lignin and ILs. The complexity of IL chemistry and lignin variability often limits the accuracy of these force fields, which is critical. Current force fields might not capture subtle interactions such as hydrogen bonding or  $\pi$ - $\pi$  stacking accurately. We need better machine-learned force fields that are made for ILs interacting with biopolymers like lignin in order to make MD simulations more accurate.

**Computational cost:** High-level computational methods, like ab initio MD (AIMD) or quantum mechanical (QM) calculations, can give us a lot of information about how IL and lignin interact, but they are time- and computer-intensive. Large-scale screening, which involves evaluating hundreds or thousands of ILs, makes these methods impractical. To enhance the overall efficiency of the screening process, it could be beneficial to strike a balance between the costs associated with each step's calculations and how accurate they are. We could achieve this by developing hybrid methods that blend faster, less detailed models with occasional high-precision calculations.

**Uncertainty in thermodynamic predictions:** We often use tools like COSMO-RS to guess the thermodynamic features of IL-lignin systems, like their activity coefficients, how well they dissolve, and their Gibbs free energy. However, these methods often have limitations for complex mixtures or systems with strong intermolecular interactions, averaging predictions might not accurately capture the specific behavior of the system [30]. Furthermore, predicting the impact of solvent impurities can be challenging. Temperature variations and the impact of IL recycling on lignin solubility continue to pose challenges. These uncertainties can lead to suboptimal IL selection during screening. Future

efforts should focus on refining these thermodynamic models and incorporating uncertainty quantification to improve their predictive capabilities.

### 3. Ionic Liquids for Lignin Valorization

The process of breaking down lignin through the use of ILs is a developing area of study aimed at transforming this intricate biopolymer, which is plentiful in plant biomass, into useful low-molecular-weight substances [31]. The application of ILs in lignin depolymerization has become increasingly popular because of their distinctive characteristics, such as adjustable acidity and basicity, excellent thermal stability, and re-usability to some extent, which makes them perfect for promoting chemical reactions [32].

Recent studies utilizing Density Functional Theory (DFT) and MD simulations have yielded essential insights into the reaction mechanisms, energy barriers, and interaction dynamics between lignin and ILs [33–37]. This theoretical framework allows for the optimization of conditions to enhance the yield and selectivity of targeted phenolic monomers throughout the depolymerization process. Disputes in the discipline frequently center on the effectiveness and ecological consequences of employing ILs in contrast to conventional techniques like reductive and oxidative depolymerization. Although some researchers advocate for the selectivity and reaction control benefits of ILs, others express apprehensions about their overall sustainability for their use in large-scale applications [38].

Furthermore, as discussed in the previous section, the complex nature of lignin's composition and structure requires continuous research to thoroughly understand how various ILs affect the depolymerization results and to enhance computational models for precise predictions. With the ongoing evolution of computational techniques, the combination of DFT and MD simulations is anticipated to significantly improve the comprehension and optimization of lignin depolymerization, leading to more sustainable methods for biomass valorization [39].

The comprehension of lignin depolymerization in the presence of ILs has consistently been a captivating area of research; however, insights into the process via QM methods have been limited when compared with experimental investigations. Nevertheless, there have been limited theoretical investigations aimed at enhancing comprehension of the process. This builds on several earlier investigations aimed at elucidating the cleavage mechanism of traditional C-O bonds in lignin when exposed to different combinations of ILs [34,35,40]. The earlier research focused on elucidating the catalytic mechanism for the depolymerization of lignin model compounds in the presence of ILs. This was achieved by examining the thermodynamics of the process, investigating the activation energy barriers, and analyzing the bonding parameters within the lignin-IL complex system. Building on a similar methodology, a recent investigation conducted by Liu et al. examined the electrocatalytic depolymerization of lignin into valuable chemicals, presenting a promising avenue for sustainable biorefineries [41]. The study utilizes dispersion-corrected DFT calculations to elucidate the electrocatalytic reduction mechanism associated with the depolymerization of lignin model dimers and oxidized lignin. The computational study was completely focused on understanding the depolymerization of  $\beta$ -O-4 based model compounds. The calculations clarify the role of the trimetallic ILs-PdNiBi catalytic system in promoting the generation of a crucial radical intermediate and the release of products. The anticipated energies of the reaction pathways validated the function of ionic liquids with metals in achieving total substrate conversion and elevated yields of phenols and acetophenones by reducing the activation energy barriers, thereby aiding in C-O bond cleavage. The development of C-O bond cleavage can enhance lignin valorization and utilization by facilitating the selective breakdown of lignin into valuable chemicals while preserving functional groups necessary for material applications. For example, demethylation, a specific form of C-O bond cleavage, involves breaking the bond between the methyl group and oxygen in methoxy (-OCH<sub>3</sub>) substituents on lignin's aromatic rings. The process results in increased phenolic content and reactivity of lignin fragments serving as valuable precursors for materials, phenolic resins, polyurethanes, etc [42].

With recent advancements for conversion of lignin into materials, DFT calculations have been critical in elucidating the electron transport process and material's electronic structure. Chen et al. recently published a study in which they used lignosulfonate as a carbon source from biomass to

make ultrahigh-energy density supercapacitor electrodes. Calculations were performed to gain a deeper insight into the process of electron transport. Simulations grounded in DFT also contributed to a deeper understanding of the electronic and ionic characteristics, as well as the adsorption energies of the fabricated supercapacitor electrodes. This kind of research concerning material synthesis involving lignin and ILs is experimentally accessible, yet it remains limited from a theoretical standpoint, owing to the challenges posed by lignin's complicated structure that requires resolution [43]. DFT based optimization of lignin modifications, and studying structure-property relationships can be achieved and would be of great help in the future. By revealing lignin's complex molecular behavior, these multidisciplinary approaches help speed the development of lignin-based materials and sustainable technologies.

The integration of QM calculations with ML techniques has proven to be a significant advantage in this domain, accelerating the prediction and validation of reaction mechanisms while minimizing computational effort and associated expenses. A recent investigation by Ding et al. effectively utilized machine learning alongside QM calculations to predict the bond cleavage efficiency of approximately a set of  $10^3$  possible ILs on the  $\beta$ -O-4 bond of guaiacylglycerol- $\beta$ -guaiacyl ether (GGE), achieving a strong correlation between the predicted bond dissociation energy and the experimental yield of guaiacol, emphasizing the reliability of the ML model [44]. The study pinpointed imidazolium-based cations paired with tyrosine anions as ideal candidates for future applications in the precise cleavage of lignin. This approach showcases an innovative method that bypasses the necessity for intricate activation barrier energy calculations while greatly minimizing the time required for evaluating potential ILs. DFT calculations have been helpful to generate a lot of useful molecular descriptors which can be used by the ML models. In another study performed, DFT generated dataset were used as input parameters for training ML models to understand the photocatalytic cleavage of C-C bonds in lignin-derived structures [45]. The extent of datasets for training ML models is restricted by the high computational cost of DFT. The generalization of ML models to a variety of systems may be impeded by the scarcity of data. Furthermore, while ML can expedite predictions, it frequently compromises accuracy in comparison to DFT. Nevertheless, these studies were limited to model monomeric/dimeric molecules and there is an urgent need to expand these calculations to various model polymeric lignin structures to gain further insights. The primary challenge in combining these methods is the delicate balance between the reliability of results and computational efficiency, especially when dealing with a combination of ILs and lignin systems.

On the contrary, methods such as MD simulations have emerged as a key tool in deciphering the intricate process of lignin depolymerization. These simulations offer a deeper understanding of the structural dynamics, interactions with solvents. In a study by Hackenstrass et al. the authors conducted classical MD simulations to uncover intricate details about the structural dynamics of lignin dimers, illustrating how different linkages influence lignin conformation and its interaction with water [33]. The structures used for predictions were based on dimers composed of common linkages found in lignin, including 4-O-5, 5-5,  $\alpha$ -O-4,  $\beta$ -1,  $\beta$ -5,  $\beta$ - $\beta$ , and  $\beta$ -O-4. This information is essential for predicting the solubility and reactivity of lignin in polar solvents. The study reveals that the  $\beta$ -O-4 linkage exhibits unique characteristics regarding its conformational flexibility and water interaction, which are essential for its solubility and depolymerization processes. Similar predictions can be noted in the presence of ionic solvents. In a recent study, the behavior of lignin with multiple chains was investigated within the 1-ethyl-3-methylimidazolium acetate and water system. The analysis of simulation trajectories revealed that lignin formed an aggregated complex in pure water, preserving its hydrophobic regions. Lignin chains were observed to disentangle and extend their conformation at varying concentrations of ILs in the system. The authors investigated the interaction patterns of multiple lignin chains with guaiacyl decamer in the presence of ILs, which can offer insights into the dissolution phenomenon of lignin at the atomic level [12]. Additionally, reactive force field (ReaxFF) simulations have emerged as an essential asset to the community. The ability to realistically replicate the dynamics, reactivity, and effects of solvents on the solute is a further advantage of using ReaxFF-based simulations [46]. As a result of the ease of simulating small chemical structures, ReaxFF-based simulations have been conducted on lignin-derived structures.

Since it is difficult to replicate the complex and heterogeneous nature of lignin, model compounds are frequently used by researchers. A model compound of lignin serves as a simplified representation of a particular structural unit, allowing for a focused investigation of specific bonds or functional groups present in natural lignin. Nonetheless, a few studies have investigated the use of this method for the thermo-oxidation of lignin or lignin-based products. ReaxFF simulations were employed in a study to investigate thermal decomposition in an oxygen environment for models that represented the most prevalent linkages found in softwood. This revealed the reaction pathways of dominant reaction products [47]. ReaxFF reactive force field to investigate the oxidative stabilization of softwood lignin for carbon fiber production. It focuses on lignin fragments derived from coniferyl alcohol units, revealing that the 5-5 linkage exhibits the highest reactivity towards cyclization and dehydrogenation, crucial for forming rigid connections necessary for stabilization [48]. The structures used in this study consist of 5 to 6 coniferyl alcohol units connected through different linkages (namely 4-O-5, 5-5,  $\alpha$ -O-4,  $\beta$ -1,  $\beta$ -5,  $\beta$ - $\beta$ , and  $\beta$ -O-4). Two structures were preferred, including 9 lignin fragments containing 5 monolignols, 9 lignin fragments containing 6 monolignols, and 396 oxygen molecules. The cost-effectiveness associated with ReaxFF in comparison to conventional MD simulations facilitated the utilization of large structures. In a recent study, Ahmad et al. presented a technique for identifying products and their pathways in thermo-oxidation reactions through ReaxFF simulations, aiming to elucidate the dynamics and reactivity of the thermo-oxidation of two model structures of modified lignin [49]. Nonetheless, employing ReaxFF simulations to comprehend the utilization and valorization of lignin with ILs requires significant advancement, particularly in the development of accurate force fields for extensive lignin structures and ILs.

Recently a new theoretical strategy named “Advanced real-time molecular sensing strategy” was introduced, which is claimed to be capable of capturing the dynamic nature of lignin reactions, including delignification, and de/repolymerization. The work integrated the combination of DFT, AIMD, and kinetic Monte Carlo (kMC) simulations, constituting a comprehensive framework for assessing various lignin properties. This approach was capable of providing real-time molecular level information about dissolved lignin chains during delignification, including molecular weight distributions and S/G ratio under diverse reaction conditions. However, this study was not conducted with ILs, but surely can be a usable synergistic integration of theoretical approaches with experimental data for understanding of the lignin systems in presence of ILs and their potential for sustainable biorefineries in future [50]. A comprehensive grasp of this subject can aid in the formulation of a cost-effective process for lignin dissolution using ILs and even depolymerization, ultimately facilitating an improved biorefinery operation.

#### 4. Challenges

Despite the advances made using DFT, and MD simulations, several challenges remain. One significant challenge is the computational cost associated with simulating large lignin structures. While DFT is excellent for small systems, its computational expense becomes prohibitive for larger systems, necessitating the use of simplified lignin models, which may not fully capture the complexity of real lignin.

One significant challenge lies in the accuracy of the models used to represent lignin's complex structure. Different approaches, such as using repeating building blocks or a stochastic approach to vary linkages and monomers, aim to create representative lignin models. However, these models often come with inherent uncertainties, especially regarding the linkage distributions, which can have a margin of error of up to 5% [51]. Such discrepancies can lead to variations in the predicted behavior of lignin during depolymerization processes. Furthermore, the MD simulations are influenced by the time step chosen, with common practices employing smaller time steps. However, lower Metropolis temperatures or increased Monte Carlo (MC) events might improve simulated statistics but also increase computational complexity.

Additionally, accurately modeling ILs is challenging due to the wide range of interactions such as hydrogen bonding, van der Waals forces, and electrostatic interactions present in these systems. High-level quantum chemical methods are often required to model these interactions accurately,

making simulations computationally intensive. Moreover, the dynamic behavior of ILs, especially when mimicking the real experimental conditions, adds another layer of complexity to the simulations. Reactive simulation techniques can be useful to understand the dynamics structure of lignin and ILs, as well as being capable of tracking the reaction in presence of solvents and catalysts. Polarizable force fields were introduced in conjunction with reactive force field simulations to address simulation concerns related with ILs, and they may be more accurate in capturing local electronic effects than non-polarizable force fields. Their limited DFT use to date is certainly due to increasing computing cost or difficulty in implementation, despite their capacity to mimic more delicate electronic characteristics [52,53]. The major advantage of the method of ReaxFF is its ability to show how chemical reactions occur with no preassigned connectivity, which is useful for a number of dynamic studies in chemistry and materials science field. However, the method has limitations, especially when applied on larger structures. These are advanced processes of parameterization, lower precision in relation to QM calculations, and difficulties in interpreting simulation results of complex systems. Also, it should be noted that while ReaxFF allows carrying out simulations that are many times larger than QM calculations allow, there are also limitations in practical terms for the size of the model and for the duration that can be simulated [46].

## 5. Future Perspectives

Numerous prospective avenues may aid in addressing these challenges and enhancing the theoretical comprehension of lignin depolymerization through the use of ionic liquids:

1. Integration of ML: By training models on existing data from DFT and MD simulations, researchers could forecast the properties of new ILs and structural properties of lignin without relying on resource-intensive quantum chemical calculations. This method has the potential to greatly accelerate the identification of novel ILs for the depolymerization of lignin.
2. Advanced Multi-scale Modeling: The combination of DFT and MD within a unified multi-scale framework would enable researchers to effectively capture both the electronic and dynamic behaviors involved in the depolymerization of lignin in ILs. This method may offer a comprehensive perspective on the depolymerization process and contribute to the formation of more effective ILs. Development of accurate reactive force fields for lignin and ILs can also be a solution to this field as they can capture the reactive dynamics of the system at a lower computational cost in comparison to classical MD simulations.
3. Selective Catalysis: Formulating tailored ILs that focus on particular lignin linkages (e.g., C-O, and C-C) while reducing unwanted side reactions will be essential for enhancing lignin depolymerization. The computational data analyzed through DFT and MD simulations are quite limited and have primarily been applied to a specific group of ILs, which restricts the overall understanding and should be expanded.

## 6. ML Methods for Lignin Processes

ML techniques are increasingly being applied to optimize and predict outcomes in lignin valorization processes using ILs as also highlighted in the previous sections. This emerging approach offers significant potential to accelerate research and development by reducing the need for time-consuming and costly experimental work across various scales - from molecular interactions to process development.

The synthetic flexibility of ILs lends itself to large datasets. As a result, they provide ample opportunity for ML methods. As a result, a growing amount of literature has been published on the use of ML for ILs [54]. The use of ML for ILs in general has been comprehensively reviewed before [55]. Furthermore, the use of IL with ML for applications such as CO<sub>2</sub> capture [56] and thermo-chemical properties [57] has also been reviewed. But in this section of the review, we focus specifically on ML applications for lignin extraction and valorization.

At the molecular level, ML is proving valuable for predicting fundamental properties crucial for IL-based lignin processing. More sophisticated approaches using graph-based genetic algorithms and graph neural networks have been developed for solvent design in lignin-first biorefineries. This

framework designed numerous solvents with high potential for lignin solubilization, demonstrating solubilities between 20 and 60 wt% across different lignin types, including Kraft lignin from softwood, organosolv lignin from hardwood, and herbaceous lignin from corn cob by mild acidolysis [13]. Complementing this, other researchers have developed predictive toolsets that combine Hansen solubility parameters, COSMO-RS predictions, and ML models to identify effective lignocellulosic pretreatment solvents. These approaches not only predict solvent effectiveness but also provide insights into the dissolution mechanism through quantum chemical calculations [25]. Furthermore, recent work has utilized explainable ML models based on Kamlet-Taft and polarity parameters to predict and optimize lignocellulose pretreatment efficiency, offering both high predictive accuracy and interpretability of solvent properties crucial for lignin processing [58].

Moving to catalyst design, ML is also being applied to optimize both chemical and biological catalysts for lignin processing. While not specifically focused on ILs, work on using ML to engineer cytochrome P450 enzymes for optimal bioconversion of lignin fragments demonstrates the potential of this approach [59]. Similar techniques could be adapted to design catalytic ILs or to optimize IL-enzyme systems for lignin valorization. In the realm of chemical catalysis, ML models have been developed to predict the catalytic activity of ILs for cleaving the  $\beta$ -O-4 bond in lignin. The best model achieved an  $R^2$  value over 0.93, allowing rapid screening of potential IL catalysts [44]. This work highlights how ML can accelerate the discovery and optimization of IL-based catalytic systems for lignin depolymerization.

Considering larger scales, ML is being applied to optimize reaction conditions and predict process outcomes. Random forest regression has been used to predict bio-oil yield and char yield from IL-catalyzed lignin depolymerization, achieving  $R^2$  scores of 0.91 and 0.94, respectively [60]. These types of predictive models can help researchers identify promising reaction conditions without extensive experimental trials. The power of ML in uncovering non-obvious relationships in lignin processing has been further demonstrated. Using algorithms like XGBoost to model lignin hydrogenolysis revealed that parameters such as lignin-to-solvent ratio and catalyst pore size had large impacts on outcomes [61]. This type of insight can guide more targeted experimental design and process optimization. At the process development level, Bayesian optimization has been applied to simultaneously optimize multiple lignin properties in an IL-based biorefinery process. This allowed efficient navigation of the multi-dimensional process parameter space and identification of optimal conditions, significantly accelerating the development and scale-up of IL-based lignin valorization processes [62].

An important consideration in applying ML to lignin valorization is the selection and engineering of appropriate descriptors or input features. While many studies use readily available parameters like IL cation/anion identities and reaction conditions, more sophisticated approaches are emerging. Various molecular representation techniques for ILs have been reviewed, highlighting the potential of natural language processing methods that can learn directly from SMILES strings or other text-based molecular notations [55]. Building on this, Transformer Convolutional Neural Networks operating on SMILES inputs have been found to outperform traditional descriptor-based models for predicting IL melting points [63]. This suggests that advanced text-based ML models could offer improved predictive power for IL properties relevant to lignin processing. The use of explainable ML models is particularly noteworthy. Models have been developed to predict lignocellulose pretreatment efficiency based on Kamlet-Taft and polarity parameters [58]. Such approaches not only achieve high predictive accuracy but also provide insights into the relative importance of different solvent properties. This type of interpretable ML model is crucial for gaining scientific insights and building trust in ML predictions for lignin valorization processes.

Beyond reaction optimization, ML is also being applied to analytical techniques relevant to lignin research. ML models have been used to predict lignin content in poplar wood using Raman spectroscopy data [64]. This showcases how ML can enhance rapid characterization methods, potentially enabling real-time monitoring and control of lignin processing for downstream processing.

Despite these advances, challenges remain in applying ML to IL-based lignin valorization. Available datasets are often small and fragmented, and the complex, heterogeneous nature of lignin poses difficulties for developing universally applicable models. Proper validation is crucial, as random splitting of limited IL datasets can lead to overly optimistic performance estimates [63].

Looking forward, integrating ML with high-throughput experimentation could enable rapid closed-loop optimization of IL-lignin systems. Incorporating mechanistic knowledge into ML architectures may improve accuracy and interpretability. There's also significant potential in applying ML to analyze spectroscopic and analytical data from IL-lignin reactions, potentially yielding new chemical insights. ML methods have already demonstrated clear value in accelerating research on IL-based lignin valorization across multiple scales, from molecular design to process optimization. As datasets grow and algorithms improve, ML is likely to play an increasingly central role in this field.

## 7. Conclusion

Given the need to valorize lignin for a sustainable biorefinery, this review highlights the expanding influence of computational methods and machine learning approaches in lignin processing using ILs. The substantial progress in computational resources, including the integration of ML techniques, has significantly reduced the timeline of optimizing deconstruction through data-driven methods to gaining atomic- and molecular-level insights of complex lignin-IL interactions. This review additionally identifies the major research gaps and potential future directions to tackle unknown challenges in lignin valorization using ILs. For instance, the need for a) extensive experimental validation of simulated/predicted data, b) introduction of complexity in lignin structure, c) standardization of simulated data and open access data-sharing, and d) integrating ML with high-throughput experimentation. Ongoing advancement and expansion of computational methods to enhance lignin dissolution and valorization, particularly through next-generation predictive paradigms, are crucial for establishing the fundamental science needed to develop and optimize effective lignin conversion processes. In summary, realizing the full potential of lignin as a renewable feedstock will require continued collaboration between data scientists, chemists, and chemical engineers to develop ML tools tailored for the unique challenges of IL-lignin systems.

**Acknowledgments:** This material is based upon work supported by the Joint BioEnergy Institute, U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number DE-AC02-05CH11231 with Lawrence Berkeley National Laboratory. The U.S. Government retains, and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript or allow others to do so, for the U.S. Government purpose. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. DOE's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. DOE or the U.S. Government.

**Conflicts of Interest:** BAS has a financial interest in Illium Technologies, Caribou Biofuels, and Erg Bio. All other authors declare the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Ralph, J.; Lapierre, C.; Boerjan, W. Lignin structure and its engineering. *Curr. Opin. Biotechnol.* **2019**, *56*, 240–249, doi:10.1016/j.copbio.2019.02.019.
2. Schutyser, W.; Renders, T.; Van den Bosch, S.; Koelewijn, S.F.; Beckham, G.T.; Sels, B.F. Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem. Soc. Rev.* **2018**, *47*, 852–908, doi:10.1039/c7cs00566k.
3. Yu, O.; Kim, K.H. Lignin to materials: A focused review on recent novel lignin applications. *Appl. Sci.* **2020**, *10*, 4626, doi:10.3390/app10134626.

4. Domínguez-Robles, J.; Cárcamo-Martínez, Á.; Stewart, S.A.; Donnelly, R.F.; Larrañeta, E.; Borrega, M. Lignin for pharmaceutical and biomedical applications – Could this become a reality? *Sustainable Chemistry and Pharmacy* **2020**, *18*, 100320, doi:10.1016/j.scp.2020.100320.
5. Shrestha, S.; Goswami, S.; Banerjee, D.; Garcia, V.; Zhou, E.; Olmsted, C.N.; Majumder, E.L.-W.; Kumar, D.; Awasthi, D.; Mukhopadhyay, A.; Singer, S.W.; Gladden, J.M.; Simmons, B.A.; Choudhary, H. Perspective on lignin conversion strategies that enable next generation biorefineries. *ChemSusChem* **2024**, *17*, e202301460, doi:10.1002/cssc.202301460.
6. Sethupathy, S.; Murillo Morales, G.; Gao, L.; Wang, H.; Yang, B.; Jiang, J.; Sun, J.; Zhu, D. Lignin valorization: Status, challenges and opportunities. *Bioresour. Technol.* **2022**, *347*, 126696, doi:10.1016/j.biortech.2022.126696.
7. Lei, Z.; Dai, C.; Hallett, J.; Shiflett, M. Introduction: ionic liquids for diverse applications. *Chem. Rev.* **2024**, *124*, 7533–7535, doi:10.1021/acs.chemrev.4c00291.
8. Zhang, Z.; Song, J.; Han, B. Catalytic Transformation of Lignocellulose into Chemicals and Fuel Products in Ionic Liquids. *Chem. Rev.* **2017**, *117*, 6834–6880, doi:10.1021/acs.chemrev.6b00457.
9. Wang, B.; Qin, L.; Mu, T.; Xue, Z.; Gao, G. Are ionic liquids chemically stable? *Chem. Rev.* **2017**, *117*, 7113–7131, doi:10.1021/acs.chemrev.6b00594.
10. König-Mattern, L.; Komarova, A.O.; Ghosh, A.; Linke, S.; Rihko-Struckmann, L.K.; Luterbacher, J.; Sundmacher, K. High-Throughput Computational Solvent Screening for Lignocellulosic Biomass Processing. *SSRN Journal* **2022**, doi:10.2139/ssrn.4182541.
11. Mohan, M.; Simmons, B.A.; Sale, K.L.; Singh, S. Multiscale molecular simulations for the solvation of lignin in ionic liquids. *Sci. Rep.* **2023**, *13*, 271, doi:10.1038/s41598-022-25372-2.
12. Radhakrishnan, R.; Manna, B.; Ghosh, A. Molecular insights into dissolution of lignin bunch in ionic liquid-water mixture for enhanced biomass conversion. *Renew. Energy* **2023**, *206*, 47–59, doi:10.1016/j.renene.2023.02.036.
13. König-Mattern, L.; Komarova, A.O.; Ghosh, A.; Linke, S.; Rihko-Struckmann, L.K.; Luterbacher, J.; Sundmacher, K. High-throughput computational solvent screening for lignocellulosic biomass processing. *Chemical Engineering Journal* **2023**, *452*, 139476, doi:10.1016/j.cej.2022.139476.
14. Paduszyński, K.; Domańska, U. Thermodynamic modeling of ionic liquid systems: development and detailed overview of novel methodology based on the PC-SAFT. *J. Phys. Chem. B* **2012**, *116*, 5002–5018, doi:10.1021/jp3009207.
15. Valderrama, J.O.; Robles, P.A. Critical properties, normal boiling temperatures, and acentric factors of fifty ionic liquids. *Ind. Eng. Chem. Res.* **2007**, *46*, 1338–1344, doi:10.1021/ie0603058.
16. Eike, D.M.; Brennecke, J.F.; Maginn, E.J. Predicting Infinite-Dilution Activity Coefficients of Organic Solutes in Ionic Liquids. *Ind. Eng. Chem. Res.* **2004**, *43*, 1039–1048, doi:10.1021/ie034152p.
17. Shah, J.K.; Maginn, E.J. Monte Carlo simulations of gas solubility in the ionic liquid 1-n-butyl-3-methylimidazolium hexafluorophosphate. *J. Phys. Chem. B* **2005**, *109*, 10395–10405, doi:10.1021/jp0442089.
18. Shi, W.; Maginn, E.J. Molecular simulation and regular solution theory modeling of pure and mixed gas absorption in the ionic liquid 1-n-hexyl-3-methylimidazolium bis(trifluoromethylsulfonyl)amide ([hmim][Tf2N]). *J. Phys. Chem. B* **2008**, *112*, 16710–16720, doi:10.1021/jp8075782.
19. Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99*, 2224–2235, doi:10.1021/j100007a062.
20. Klamt, A.; Eckert, F. COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equilib.* **2000**, *172*, 43–72, doi:10.1016/S0378-3812(00)00357-5.
21. Kumar, N.; Mohan, M.; Smith, J.C.; Simmons, B.A.; Singh, S.; Banerjee, T. Inhibition of asphaltene aggregation using deep eutectic solvents: COSMO-RS calculations and experimental validation. *J. Mol. Liq.* **2024**, *400*, 124471, doi:10.1016/j.molliq.2024.124471.
22. Hadj-Kali, M.K.; Althuluth, M.; Mokraoui, S.; Wazeer, I.; Ali, E.; Richon, D. Screening of ionic liquids for gas separation using COSMO-RS and comparison between performances of ionic liquids and aqueous alkanolamine solutions. *Chem. Eng. Commun.* **2020**, *207*, 1264–1277, doi:10.1080/00986445.2019.1680363.
23. Khan, H.W.; Reddy, A.V.B.; Nasef, M.M.E.; Bustam, M.A.; Goto, M.; Moniruzzaman, M. Screening of ionic liquids for the extraction of biologically active compounds using emulsion liquid membrane: COSMO-RS prediction and experiments. *J. Mol. Liq.* **2020**, *309*, 113122, doi:10.1016/j.molliq.2020.113122.
24. Yu, K.; Ding, W.-L.; Lu, Y.; Wang, Y.; Liu, Y.; Liu, G.; Huo, F.; He, H. Ionic liquids screening for lignin dissolution: COSMO-RS simulations and experimental characterization. *J. Mol. Liq.* **2022**, *348*, 118007, doi:10.1016/j.molliq.2021.118007.
25. Achinivu, E.C.; Mohan, M.; Choudhary, H.; Das, L.; Huang, K.; Magurudeniya, H.D.; Pidatala, V.R.; George, A.; Simmons, B.A.; Gladden, J.M. A predictive toolset for the identification of effective lignocellulosic pretreatment solvents: a case study of solvents tailored for lignin extraction. *Green Chem.* **2021**, *23*, 7269–7289, doi:10.1039/D1GC01186C.
26. Zhao, J.; Zhou, G.; Fang, T.; Ying, S.; Liu, X. Screening ionic liquids for dissolving hemicellulose by COSMO-RS based on the selective model. *RSC Adv.* **2022**, *12*, 16517–16529, doi:10.1039/d2ra02001g.

27. Zhou, L.; Liu, Y.; Zhang, J.; Li, Q.; Yuan, M.; Kang, Z. Ionic liquid screening for lignocellulosic biomass fractionation: COSMO-RS prediction and experimental verification. *J. Mol. Liq.* **2024**, *407*, 125214, doi:10.1016/j.molliq.2024.125214.
28. Ralph, J.; Lundquist, K.; Brunow, G.; Lu, F.; Kim, H.; Schatz, P.F.; Marita, J.M.; Hatfield, R.D.; Ralph, S.A.; Christensen, J.H.; Boerjan, W. Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochemistry Reviews* **2004**, *3*, 29–60, doi:10.1023/B:PHYT.0000047809.65444.a4.
29. Mohan, M.; Huang, K.; Pidatala, V.; Simmons, B.; Singh, S.; Sale, K.L.; Gladden, J. Prediction of Solubility Parameters of Lignin and Ionic Liquids Using Multi-resolution Simulation Approaches. *Green Chem.* **2022**, *24*, 1165–1176, doi:10.1039/D1GC03798F.
30. Yao, A.; Choudhary, H.; Mohan, M.; Rodriguez, A.; Magurudeniya, H.; Pelton, J.G.; George, A.; Simmons, B.A.; Gladden, J.M. Can multiple ions in an ionic liquid improve the biomass pretreatment efficacy? *ACS Sustain. Chem. Eng.* **2021**, *12*, 4371–4376, doi:10.1021/acssuschemeng.0c09330.
31. Bourmaud, C.L.; Bertella, S.; Bosch Rico, A.; Karlen, S.D.; Ralph, J.; Luterbacher, J.S. Quantification of Native Lignin Structural Features with Gel-Phase 2D-HSQC0 Reveals Lignin Structural Changes During Extraction. *Angew. Chem. Int. Ed* **2024**, *63*, e202404442, doi:10.1002/anie.202404442.
32. Nanayakkara, S.; Patti, A.F.; Saito, K. Lignin Depolymerization with Phenol via Redistribution Mechanism in Ionic Liquids. *ACS Sustain. Chem. Eng.* **2014**, *2*, 2159–2164, doi:10.1021/sc5003424.
33. Hackenstrass, K.; Hasani, M.; Wohler, M. Structure, flexibility and hydration properties of lignin dimers studied with Molecular Dynamics simulations. *Holzforschung* **2024**, *78*, 98–108, doi:10.1515/hf-2023-0054.
34. Zhang, T.; Zhang, Y.; Wang, Y.; Huo, F.; Li, Z.; Zeng, Q.; He, H.; Li, X. Theoretical Insights Into the Depolymerization Mechanism of Lignin to Methyl p-hydroxycinnamate by [Bmim][FeCl<sub>4</sub>] Ionic Liquid. *Front. Chem.* **2019**, *7*, 446, doi:10.3389/fchem.2019.00446.
35. Tian, X.-Y.; Zheng, Y.-Z.; Zhang, Y.-C. Molecular design of efficient SO<sub>3</sub>H-functionalized ionic liquid to catalyse chitin into levulinic acid: NMR and DFT study. *J. Mol. Liq.* **2022**, *368*, 120735, doi:10.1016/j.molliq.2022.120735.
36. Mishra, D.K.; Banerjee, B.; Pugazhenth, G.; Banerjee, T. Metal-Free, Ionic Liquid-Mediated Hydrogen Release from Amine Borane Complexes: An Experimental and Density Functional Theory Investigation. *Ind. Eng. Chem. Res.* **2021**, *60*, 9764–9776, doi:10.1021/acs.iecr.1c01507.
37. Mishra, D.K.; Hussain, R.; Pugazhenth, G.; Banerjee, T. Catalytic Effect of Ionic Liquid Induced H<sub>2</sub>-Release from Morpholine Borane Complex: An Efficient Hydrogen Storage Carrier. *ACS Sustain. Chem. Eng.* **2022**, *10*, 6157–6164, doi:10.1021/acssuschemeng.2c00561.
38. Mukesh, C.; Huang, G.; Qin, H.; Liu, Y.; Ji, X. Insight into lignin oxidative depolymerization in ionic liquids and deep eutectic solvents. *Biomass and Bioenergy* **2024**, *188*, 107305, doi:10.1016/j.biombioe.2024.107305.
39. Singh, S.K. Ionic liquids and lignin interaction: An overview. *Bioresource Technology Reports* **2022**, *17*, 100958, doi:10.1016/j.biteb.2022.100958.
40. Zhang, Y.; Huo, F.; Wang, Y.; Xia, Y.; Tan, X.; Zhang, S.; He, H. Theoretical Elucidation of  $\beta$ -O-4 Bond Cleavage of Lignin Model Compound Promoted by Sulfonic Acid-Functionalized Ionic Liquid. *Front. Chem.* **2019**, *7*, 78, doi:10.3389/fchem.2019.00078.
41. Liu, G.; Lu, Y.; Lu, J.; Wang, Y.; Liang, S.; He, H.; Jiang, L. Ionic liquid-trimetallic electrocatalytic system for C-O bond cleavage in lignin model compounds and lignin under ambient conditions. *Nano Res.* **2024**, *17*, 2420–2428, doi:10.1007/s12274-023-6086-z.
42. Luo, Q.; Li, C.; Zhao, W.; Ding, W.; Liu, Y.; Xiao, W.; Liu, H.; Pang, X.; Sun, J. Lignin Demethylated by Protic Ionic Liquid as a Novel and Sustainable Chrome-Free Tanning Agent for Eco-Leather Production. *ACS Sustain. Chem. Eng.* **2024**, *12*, 9682–9694, doi:10.1021/acssuschemeng.4c01036.
43. Chen, R.; Tang, H.; He, P.; Zhang, W.; Dai, Y.; Zong, W.; Guo, F.; He, G.; Wang, X. Interface Engineering of Biomass-Derived Carbon used as Ultrahigh-Energy-Density and Practical Mass-Loading Supercapacitor Electrodes. *Adv. Funct. Mater.* **2023**, *33*, doi:10.1002/adfm.202212078.
44. Ding, W.-L.; Zhang, T.; Wang, Y.; Xin, J.; Yuan, X.; Ji, L.; He, H. Machine Learning Screening of Efficient Ionic Liquids for Targeted Cleavage of the  $\beta$ -O-4 Bond of Lignin. *J. Phys. Chem. B* **2022**, *126*, 3693–3704, doi:10.1021/acs.jpcc.1c10684.
45. Zhang, T.; Wu, C.; Xing, Z.; Zhang, J.; Wang, S.; Feng, X.; Zhu, J.; Lu, X.; Mu, L. Machine learning prediction of photocatalytic lignin cleavage of C–C bonds based on density functional theory. *Materials Today Sustainability* **2022**, *20*, 100256, doi:10.1016/j.mtsust.2022.100256.
46. Senftle, T.P.; Hong, S.; Islam, M.M.; Kylasa, S.B.; Zheng, Y.; Shin, Y.K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M.J.; Aktulga, H.M.; Verstraelen, T.; Grama, A.; van Duin, A.C.T. The ReaxFF reactive force-field: development, applications and future directions. *npj Comput. Mater.* **2016**, *2*, 15011, doi:10.1038/npjcompumats.2015.11.
47. Beste, A. ReaxFF study of the oxidation of lignin model compounds for the most common linkages in softwood in view of carbon fiber production. *J. Phys. Chem. A* **2014**, *118*, 803–814, doi:10.1021/jp410454q.
48. Beste, A. Reaxff study of the oxidation of softwood lignin in view of carbon fiber production. *Energy Fuels* **2014**, *28*, 7007–7013, doi:10.1021/ef501901p.

49. Ahmed, S.; Eder, S.J.; Dörr, N.; Martini, A. Tracking Thermo-Oxidation Reaction Products and Pathways of Modified Lignin Structures from Reactive Molecular Dynamics Simulations. *J. Phys. Chem. A* **2024**, *128*, 5398–5407, doi:10.1021/acs.jpca.4c00964.
50. Lee, C.H.; Kim, J.; Ryu, J.; Won, W.; Yoo, C.G.; Kwon, J.S.-I. Lignin structure dynamics: Advanced real-time molecular sensing strategies. *Chemical Engineering Journal* **2024**, *487*, 150680, doi:10.1016/j.cej.2024.150680.
51. Wang, Y.; Kalscheur, J.; Ebikade, E.; Li, Q.; Vlachos, D.G. LigninGraphs: lignin structure determination with multiscale graph modeling. *J. Cheminform.* **2022**, *14*, 43, doi:10.1186/s13321-022-00627-2.
52. Bodo, E. Perspectives in the computational modeling of new generation, biocompatible ionic liquids. *J. Phys. Chem. B* **2022**, *126*, 3–13, doi:10.1021/acs.jpcc.1c09476.
53. Borodin, O. Polarizable force field development and molecular dynamics simulations of ionic liquids. *J. Phys. Chem. B* **2009**, *113*, 11463–11478, doi:10.1021/jp905220k.
54. Wei, Z.; Chen, F.; Liu, H.; Huang, R.; Pan, K.; Ji, W.; Wang, J. Mapping the application research on machine learning in the field of ionic liquids: A bibliometric analysis. *Fluid Phase Equilib.* **2024**, *583*, 114117, doi:10.1016/j.fluid.2024.114117.
55. Koutsoukos, S.; Philippi, F.; Malaret, F.; Welton, T. A review on machine learning algorithms for the ionic liquid chemical space. *Chem. Sci.* **2021**, *12*, 6820–6843, doi:10.1039/d1sc01000j.
56. Sun, J.; Sato, Y.; Sakai, Y.; Kansha, Y. A review of ionic liquids and deep eutectic solvents design for CO<sub>2</sub> capture with machine learning. *J. Clean. Prod.* **2023**, *414*, 137695, doi:10.1016/j.jclepro.2023.137695.
57. Yusuf, F.; Olayiwola, T.; Afagwu, C. Application of Artificial Intelligence-based predictive methods in Ionic liquid studies: A review. *Fluid Phase Equilib.* **2021**, *531*, 112898, doi:10.1016/j.fluid.2020.112898.
58. Ge, H.; Bai, Y.; Zhou, R.; Liu, Y.; Wei, J.; Wang, S.; Li, B.; Xu, H. Explicable Machine Learning for Predicting High-Efficiency Lignocellulose Pretreatment Solvents Based on Kamlet–Taft and Polarity Parameters. *ACS Sustain. Chem. Eng.* **2024**, *12*, 7578–7590, doi:10.1021/acssuschemeng.4c01563.
59. Dias, A.H.S.; Cao, Y.; Skaf, M.S.; de Visser, S.P. Machine learning-aided engineering of a cytochrome P450 for optimal bioconversion of lignin fragments. *Phys. Chem. Chem. Phys.* **2024**, *26*, 17577–17587, doi:10.1039/d4cp01282h.
60. Castro Garcia, A.; Shuo, C.; Cross, J.S. Machine learning based analysis of reaction phenomena in catalytic lignin depolymerization. *Bioresour. Technol.* **2022**, *345*, 126503, doi:10.1016/j.biortech.2021.126503.
61. Liu, Y.; Cheng, S.; Cross, J.S. Machine learning assisted chemical process parameter mapping on lignin hydrogenolysis. *Energies* **2022**, *16*, 256, doi:10.3390/en16010256.
62. Löfgren, J.; Tarasov, D.; Koitto, T.; Rinke, P.; Balakshin, M.; Todorović, M. Machine learning optimization of lignin properties in green biorefineries. *ACS Sustain. Chem. Eng.* **2022**, *10*, 9469–9479, doi:10.1021/acssuschemeng.2c01895.
63. Makarov, D.M.; Fadeeva, Yu.A.; Shmukler, L.E.; Tetko, I.V. Beware of proper validation of models for ionic Liquids! *J. Mol. Liq.* **2021**, *344*, 117722, doi:10.1016/j.molliq.2021.117722.
64. Gao, W.; Zhou, L.; Liu, S.; Guan, Y.; Gao, H.; Hui, B. Machine learning prediction of lignin content in poplar with Raman spectroscopy. *Bioresour. Technol.* **2022**, *348*, 126812, doi:10.1016/j.biortech.2022.126812.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.