



OPEN

Multiscale molecular simulations for the solvation of lignin in ionic liquids

Mood Mohan^{1,2}, Blake A. Simmons^{1,3✉}, Kenneth L. Sale^{1,4} & Seema Singh^{1,2}

Lignin, the second most abundant biopolymer found in nature, has emerged as a potential source of sustainable fuels, chemicals, and materials. Finding suitable solvents, as well as technologies for efficient and affordable lignin dissolution and depolymerization, are major obstacles in the conversion of lignin to value-added products. Certain ionic liquids (ILs) are capable of dissolving and depolymerizing lignin but designing and developing an effective IL for lignin dissolution remains quite challenging. To address this issue, the CONductor-like Screening MOdel for Real Solvents (COSMO-RS) model was used to screen 5670 ILs by computing logarithmic activity coefficients ($\ln(\gamma)$) and excess enthalpies (H^E) of lignin, respectively. Based on the COSMO-RS computed thermodynamic properties ($\ln(\gamma)$ and H^E) of lignin, anions such as acetate, methyl carbonate, octanoate, glycinate, alaninate, and lysinate in combination with cations like tetraalkylammonium, tetraalkylphosphonium, and pyridinium are predicted to be suitable solvents for lignin dissolution. The dissolution properties such as interaction energy between anion and cation, viscosity, Hansen solubility parameters, dissociation constants, and Kamlet–Taft parameters of selected ILs were evaluated to assess their propensity for lignin dissolution. Furthermore, molecular dynamics (MD) simulations were performed to understand the structural and dynamic properties of tetrabutylammonium [TBA]⁺-based ILs and lignin mixtures and to shed light on the mechanisms involved in lignin dissolution. MD simulation results suggested [TBA]⁺-based ILs have the potential to dissolve lignin because of their higher contact probability and interaction energies with lignin when compared to cholinium lysinate.

In the race for sustainable energy, lignocellulosic biomass has been estimated to be able to provide 20% of the world's energy demand by 2050¹. A total of 170 billion metric tons of lignocellulosic biomass are produced each year around the world². In order to realize the full potential of lignocellulosic biomass, the lignin (roughly 20–30 wt% of the initial biomass composition) has yet to be efficiently utilized at an industrial scale^{3–5}. Dissolution and depolymerization of lignin during lignocellulosic biomass deconstruction is a critical step in the production of biofuels and lignin-based bioproducts^{6–8}. Not only does lignin prevent enzymatic biomass degradation^{9,10}, but to make biorefineries economically viable, it must be used to produce value-added bioproducts¹¹. After lignin is removed, cellulose and hemicellulose can be further transformed into biofuel precursors^{12–14}, while the dissolved lignin can serve as a valuable precursor for carbon fibers¹⁵, colloids¹⁶, thermoplastics^{17,18}, aromatic chemicals¹⁹, lignin-based nanomaterials^{4,20}, commercial adhesives²¹, and more²².

The aromatic backbone of lignin is composed of three main phenylpropanoid units, p-hydroxy phenol (H), guaiacyl (G), and syringyl (S)^{7,23}. These phenylpropanoid monomers are linked together by ether (C–O–C: β –O–4, α –O–4, and 4–O–5) and C–C (β – β , β –5, β –1, and 5–5) bonds formed during biosynthesis^{23–25}. Lignin is highly resistant to enzymatic breakdown because of its heterogeneity, hydrophobicity, and cross-linked structure²⁶. Additionally, lignin contains an abundant number of intramolecular hydrogen bonds and π – π interactions between the rigid aromatic rings, resulting in a compact three-dimensional H-bond network, adding to its recalcitrance to enzymatic depolymerization. Lignin is also very insoluble in water and most common organic solvents, thus limiting its potential use for high-value chemical applications. As a result, sustainable biorefineries require a solvent that is capable of removing lignin and effectively rendering biomass amenable to deconstruction into upgradable fragments.

¹Deconstruction Division, Joint BioEnergy Institute, 5885 Hollis Street, Emeryville, CA 94608, USA. ²Bioresource and Environmental Security Department, Sandia National Laboratories, 7011 East Avenue, Livermore, CA 94551, USA. ³Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA. ⁴Department of Computational Biology and Biophysics, Sandia National Laboratories, 7011 East Avenue, Livermore, CA 94551, USA. ✉email: BASimmons@lbl.gov

Certain types of ionic liquids (ILs) have been demonstrated to be promising biomass solvents over the past few decades and have opened up new opportunities for the dissolution of cellulose and biomass fractionation^{13,27–31}. Ionic liquids are molten salts with a melting temperature of less than 100 °C and have unique properties including low vapor pressure, nonflammability, non-toxic, high chemical and thermal stability, and high solubility^{32–34}. The most fascinating property of ILs is their diversity in chemical structures resulting from the large number of possible combinations of anions and cations. In the last few years, the dissolution of lignin in ILs has received much attention. Pu et al.³⁵ investigated the influence of imidazolium-based ILs on lignin dissolution and found that ILs containing methyl sulfate and trifluorosulfonate anions were effective for dissolving softwood Kraft lignin. Hou et al.³⁶ measured the solubility of Kraft lignin in a series of carboxylate anion-based ILs with cholinium cation. It has been reported that as the alkyl chain length of the anion increased from formate to octanoate, the ability of the IL to solubilize lignin also increased. In addition, it also has been reported that the anions with basic group(s) (e.g.: lysinate [Lys][−]) are effective ILs for biomass delignification and significantly enhanced the enzymatic hydrolysis yields of glucose and xylose. Other studies claimed that the solubility of lignin is influenced not only by anion properties, but also by the properties of the cation. According to Hart et al.³⁷, the cation has a considerably more subtle effect on lignin solubility, with planar cations being the most effective. Further, Wang et al.³⁸ studied the influence of alkyl chain length of cations on the solubility of lignin in dialkylimidazolium-based ILs. As the alkyl chain length of cations increased from C4 to C8, the solubility of lignin decreased.

Given the potential capabilities of these ILs, it is critical to understand the thermodynamic properties (e.g., solubility, viscosity, interaction between anions and cations, Kamlet–Taft parameters etc.) of neat ILs and their mixtures with lignin. It is, however, impractical to investigate each and every possible IL anion-cation pair experimentally because of the large number of potential combinations of cations and anions exist. Therefore, it is necessary to develop theoretical methodologies that are reliable to screen the enormous number of ILs for lignin dissolution before conducting experiments. Several predictive techniques have recently been developed and adapted for modeling the thermodynamic properties of systems containing ILs, such as the perturbed-chain statistical associating fluid theory (PC-SAFT) method³⁹, group contribution methods⁴⁰, quantitative structure–property relationships (QSPR)⁴¹, Monte Carlo (MC) molecular simulations⁴², molecular dynamics (MD) simulations⁴³, and COnductor-like Screening MOdel for Real Solvents (COSMO-RS)^{44,45}. On the other hand, a potentially useful approach is to develop machine learning (ML) techniques. Using the ML models, the designing of solvents and catalysts for lignin dissolution and depolymerization are possible with higher accuracy. Recently, Ding et al.⁴⁶ developed ML models for screening efficient ionic liquid for cleavage of the lignin ether (β -O-4) linkage by predicting the bond dissociation energies (BDE) of lignin in ILs. Smuga-Kogut et al.^{47,48} have developed the ML models for predicting bioethanol yields from lignocellulosic biomass with the use of ionic liquid pretreatments. These models are developed with a small set of data points; however, to develop a robust ML model for lignin and biomass conversion, a vast number of databases are needed. Therefore, among these predictive techniques, the COSMO-RS model is a powerful tool for rapidly screening large numbers of ILs and has been widely used for the screening of ILs for different applications^{27,49–51}. The ability to screen ILs using predictive methods that only require a minimal number of input parameters would result in significant savings in time, cost, and effort in laboratory studies.

Fundamental understanding and development of a suitable solvent system for lignin dissolution by computational and experimental methods have immense potential, and implementation of the COSMO-RS model in screening of ILs for cellulose, hemicellulose, and lignin dissolution has been successfully demonstrated^{27,51–54}. For the dissolution of lignin, Balaji et al.⁵² and Casas et al.^{53,55} screened hundreds of ILs by predicting the Hildebrand solubility parameter and thermodynamic properties (excess enthalpy and activity coefficient) using the COSMO-RS model. They reported that closer solubility parameter values, lower activity coefficient and exothermic behavior of excess enthalpy are beneficial for higher dissolution of lignin. However, in both studies, only lignin monomer structures were used as model compounds for their computations. Recently, Yu et al.⁵⁶ designed 19 lignin model components (monomers, dimers, and trimers) and computationally screened 3886 ILs using the COSMO-RS calculations. Lignin's monomeric and dimer structures do not accurately depict the lignin molecule because the lack of many different linkages present in lignin. In fact, not only the ILs used for lignin dissolution are limited at present, but also the generalized lignin model structure is not fully explored for molecular simulations. In such a case, no general guidelines can be summarized and used to anticipate the most appropriate solvent to dissolve lignin, which hinders identification and optimization of process for conversion of lignin to and high-value products. Therefore, systematic screening of structurally diverse ILs is essential for developing a comprehensive theoretical understanding of lignin solubilization in IL-based solvent systems and for developing optimal solvents that will facilitate lignin valorization.

For lignin dissolution, ionic liquids have shown some unique properties and have gained much attention in research. Much development work remains to be done, however, as the technical challenges shift from fundamental scientific understanding to process-oriented engineering exploitation, the research priorities are likely expected to evolve. Even though the market for ionic liquids wasn't as big as expected in the last decade, there is an increasing number of encouraging trends. The most impressive aspect is the 57 ionic liquid applications that have been successfully implemented to this day, which represents a significant leap when compared to the 13 applications that were known to exist in 2008⁵⁷. There will be an optimistic future for ionic liquids, and “this is very much the end of the beginning regarding ionic liquid technologies, and not the beginning of the end”. According to a survey that was conducted by Roland Kalb⁵⁷ at Proionic, which includes 25 academic and industrial leaders, it was determined that “the field of ionic liquids will soon reach maturity and enter the megatrend mass markets”.

Thus, the present study attempts to use multiscale molecular simulation strategies to design and develop the fundamental knowledge required to develop robust IL solvent systems for lignin. The following goals are intended to address: (1) develop a representative lignin model structure for the molecular simulations and screen

thousands of ILs for lignin dissolution using the COSMO-RS model. (2) determine the parameters from the COSMO-RS model and quantum chemistry (QC) calculations that influence lignin dissolution. Key parameters such as interaction energy between anion and cation of the IL, viscosity, Hansen solubility parameters, dissociation constants, and Kamlet–Taft parameters were evaluated to assess relationships between IL properties and lignin dissolution. Finally, (3) analyze the structural and dynamic properties (radius of gyration, solvent accessible surface area, hydrogen bonds, radial distribution functions, and mass fractal dimension factor) of lignin in IL-lignin mixtures using molecular dynamics (MD) simulations to shed light on the mechanisms involved in lignin dissolution.

Results and discussions

Benchmarking study. The COSMO-RS model is a powerful computational approach to calculating thermodynamic properties and screening solvents for polymer dissolution. Only structural information of the solvent and solute is typically required for the COSMO-RS to predict the solubility and other thermodynamic properties. It should be worthwhile to mention that the COSMO-RS predicted thermodynamic properties such as logarithmic activity coefficient and excess enthalpies are critical parameters in determining the ability of a solvent on the solute's dissolution. The activity coefficient and excess enthalpies are related to the solubility and interaction between unlike species (e.g., lignin and ILs), respectively. Solute with lower logarithmic activity coefficient $\ln(\gamma)$ and excess enthalpies (H^E) in the solvents indicate that solvent has a stronger tendency to interact and solubilize the solute. In our earlier work on lignocellulosic biomass, we showed the applicability of COSMO-RS for calculating activity coefficients (γ) and predicting carbohydrate solubility in ionic liquids and validated the approach against a set of experimentally studied systems. We then computationally screened thousands of ILs for cellulose and hemicellulose dissolution^{27,32,58}. In addition, we have validated the COSMO-RS model against experimental data for plastic dissolution in ILs and successfully screened 9405 ILs for plastic dissolution⁵⁹.

In the current work we first performed a series of benchmark studies for dissolution of lignin in a series of cholinium-based ILs by comparing computed thermodynamic properties to experimentally determined lignin solubility (Fig. 1). Figure 1a shows the COSMO-RS predicted $\ln(\gamma)$ against known experimental lignin solubility data taken from the literature^{36,60,61}. It was observed that the amino acid based ILs had shown lower $\ln(\gamma)$ value resulted in higher solubility of lignin. Figure 1b–d shows the correlations between COSMO-RS calculated H^E , $\ln(\gamma)$, and contact probability against experimental lignin dissolution in cholinium-based ILs. Figure 1b–d also elucidating the effect of the alkyl chain length of anions on lignin solubility. Similar to $\ln(\gamma)$, lower values of H^E was associated with greater dissolution capability of lignin in ILs. In contrast, higher values of contact probability were associated with greater dissolution capability of lignin in ILs. Results of the above validation showed good correlation between COSMO-RS calculated thermodynamic properties and experimental lignin dissolution in ILs, indicating calculated $\ln(\gamma)$, H^E and contact probability should be good predictors of the ability of a specific IL to solubilize lignin. In addition, the solubility of lignin in different classes of ionic liquids was also performed and the results are depicted in Fig. S1. From Fig. S1, the correlation between experimental lignin solubilities and the COSMO-RS predicted activity shows excellent agreement. Therefore, we used the COSMO-RS approach to calculate these thermodynamic properties for an additional 5670 ionic liquids and to predict their propensity toward lignin dissolution. It is important to mention that we have not considered weakly coordinating and halogen-based anions (i.e., $[\text{BF}_4]^-$, $[\text{PF}_6]^-$, Cl^- , Br^- , I^- , F^- etc.) for the IL screening study due to their higher melting points and viscosity⁶². Moreover, the current manuscript is a theoretical work and mainly focusses on the development of ILs using multiscale molecular simulation strategies.

In the reported studies^{52,53,55,56}, monomeric and dimer structures of lignin were the most popular models, however, these structures do not directly represent the lignin molecule due to the absence of many key linkages present in lignin. The authenticity and objectivity of solute in theoretical models have been largely ignored. In this study, one of our purposes is to establish a good model structure for lignin. Therefore, in order to obtain better theoretical results, a lignin structure was built by joining all the major lignin linkages ($\beta\text{-O-4}$, $\beta\text{-}\beta$, 4-O-5 , $\alpha\text{-O-4}$, and $\beta\text{-5}$) present in the native lignin. By performing the quantum chemical calculations, the most stable conformer of lignin structure was obtained and used in the COSMO-RS calculations. The same chemical structure and stable conformer was used in our previous COSMO-RS calculations for lignin in amines, Cyrene, and protic ILs, and the results are in good agreement with experimental lignin solubility^{8,63,64}.

The solubility of lignin in different IL species were investigated using the monomeric (coniferyl alcohol and Coumaryl alcohol), dimer (lignin $\beta\text{-O-4}$ and lignin 4-O-5 linkage), trimer (dibenzodioxocin) and polymeric structure of lignin as model components for the COSMO-RS calculations. In this work, COSMO-RS predicted $\ln(\gamma)$ and H^E values are correlated with the experimental lignin solubility (Fig. S2)⁶¹. Linear regression was performed between the COSMO-RS predicted thermodynamics properties ($\ln(\gamma)$ and H^E) and experimental solubilities, and the performance of each lignin model was measured by R-square (R^2 : measures the goodness of data fitting) and the residual sum of squares (RSS), which reflected the prediction accuracy. The R-square of longer chain lignin polymer and trimer ($R^2 = 0.88$) were higher among the investigated lignin model structures. However, the RSS values for lignin trimer (i.e., dibenzodioxocin) was smaller than other models. Further, we have performed the linear regression between H^E and lignin solubility (Fig. S2b) and observed that the longer chain lignin polymeric model structure has good R^2 value (0.78) than other model structures ($R^2 = 0.39\text{--}0.41$). By considering both regression analysis ($\ln(\gamma)$ vs. experimental and H^E vs. experimental), the structure which has all the major linkages represents a better model compound for lignin with better predictions.

COSMO-RS screening of ionic liquids for lignin dissolution. Ionic liquids can be produced by combining organic cations and organic/inorganic anions, resulting in a wide range of solvents with the potential for

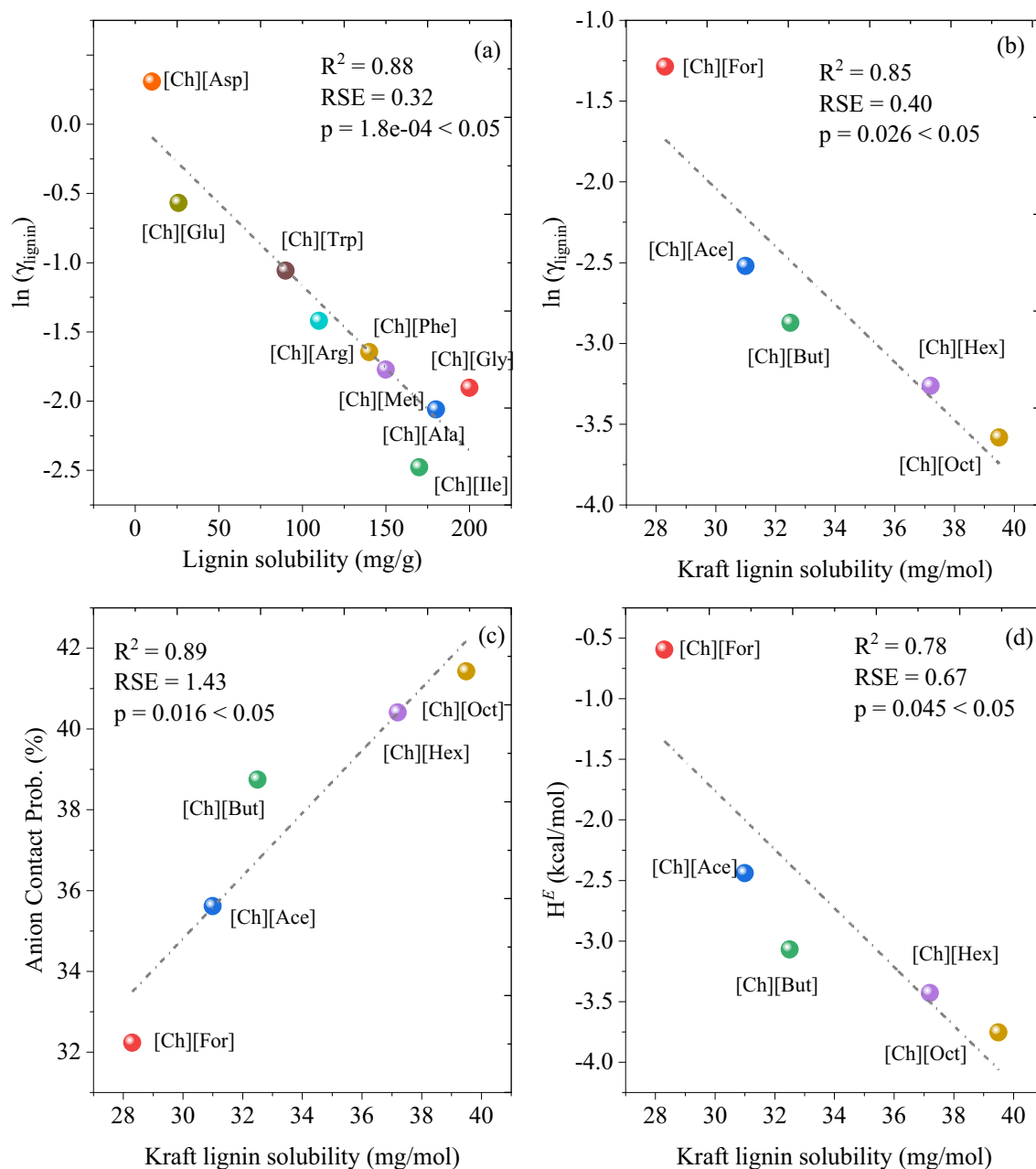


Figure 1. Correlation between the COSMO-RS predicted thermodynamic properties (a, b) logarithmic activity coefficient ($\ln(\gamma)$) versus lignin solubility, (c) excess enthalpy, H^E , versus lignin solubility, and (d) contact probability between lignin-anions versus experimental lignin solubility. COSMOTherm version 19.0.1 was used to calculate the $\ln(\gamma)$, H^E , and contact probabilities (<https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/solvation-chemistry/biovia-cosmotherm/>). (a) Data taken from Lie et al.⁶¹ and (b–d) data taken from Hou et al.⁶⁰.

polymer (i.e., lignin) dissolution. Experimental screening of the vast number of these combinations for their ability to solubilize lignin is expensive and time-consuming; the ability to accurately predict lignin dissolution in ILs will greatly reduce the cost and time to develop effective solvent systems. In the present study, 63 cations and 90 anions comprising 5670 combinations were screened for lignin dissolution by ranking them based on their $\ln(\gamma)$ and H^E values computed using COSMO-RS model (Figs. 2 and 3). The chemical structures of anions and cations of ionic liquids employed in this work can be seen in Tables S1 and S2, respectively. The COSMO files for all cations and anions were generated based on the procedure outlined in the COSMO-RS calculations section. According to our previous study on plastic-IL screening, the anions and cations were selected with different functional groups⁵⁹. The anions such as amino acid-based anions are considered for the screening study based on the Ohno and Fukumoto⁶⁵ study. Note that some of the investigated cation–anion combinations were not liquids at room temperature, thus the ILs were screened at 363.15 K.

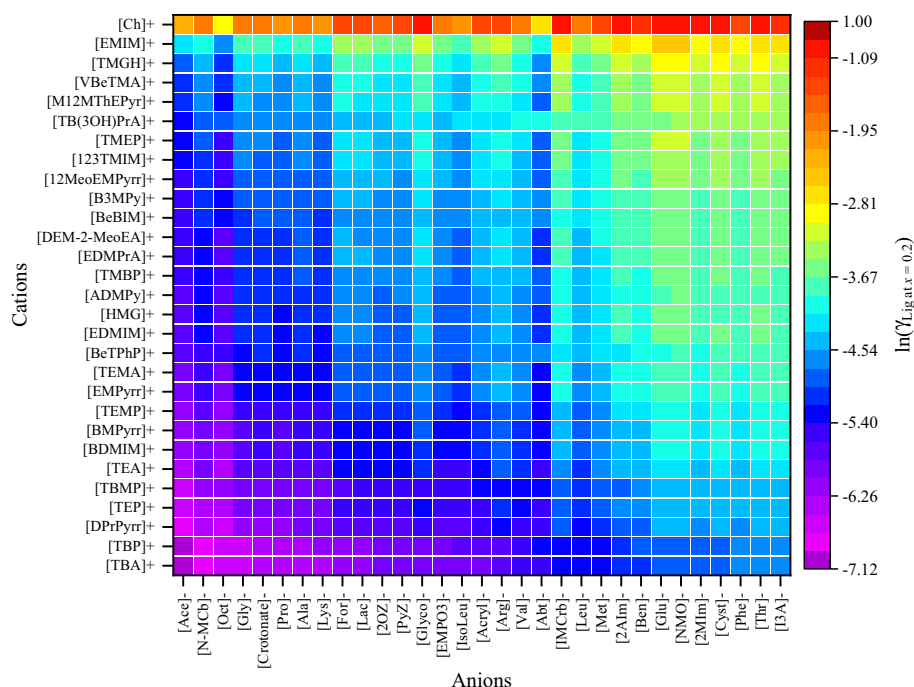


Figure 2. Graphical representation of the logarithmic activity coefficients ($\ln(\gamma)$) of lignin in different ILs at 363.15 K by the COSMO-RS model. Origin2020 SR1 used to draw the heatmaps (<https://www.originlab.com/2020>).

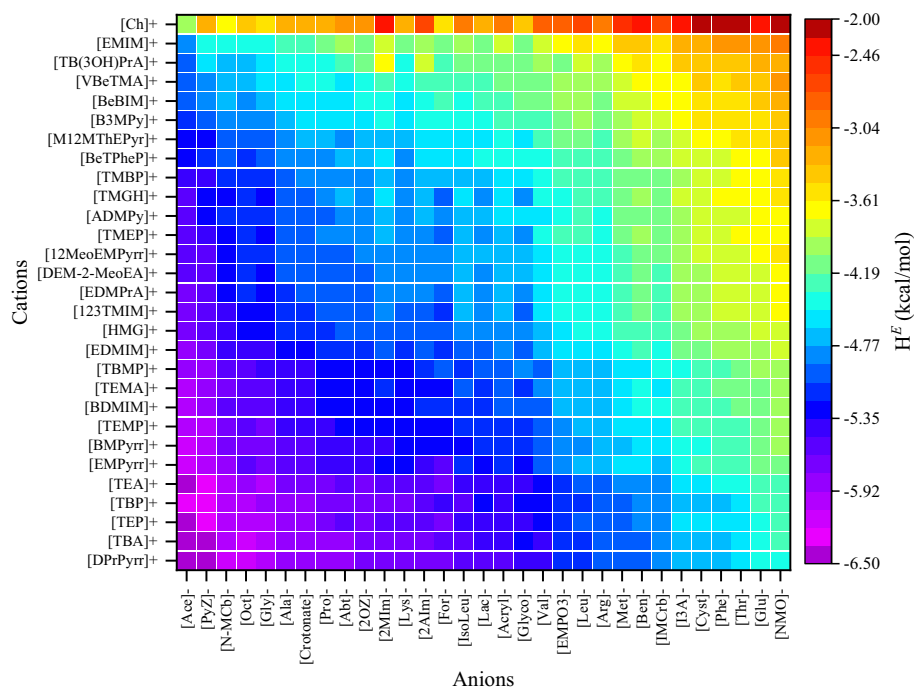


Figure 3. Graphical representation of the excess enthalpy, H^E (kcal/mol) of lignin in different ILs at 363.15 K by COSMO-RS model. Origin2020 SR1 used to draw the heatmaps (<https://www.originlab.com/2020>).

According to the solid–liquid equilibria assumptions^{66,67}, the reciprocal of the activity coefficient characterizes lignin solubility in the respective IL. Cations and anions were sorted according to their dissolving capacity and arranged in such a way that ILs with high dissolving power (i.e., $\ln(\gamma)$ and $H^E < 0$) of lignin are situated in the left and bottom portion of Figs. 2 and 3, and the weaker ones (i.e., $\ln(\gamma)$ and $H^E > 0$) are situated at the top portion and right side of Figs. 2 and 3. The anions such as acetate, methyl carbonate, octanoate, glycinate,

alaninate, and lysinate are predicted to dissolve lignin more efficiently when in combination with cations like tetraalkylammonium, tetraalkylphosphonium, and pyridinium. This is due to the fact that effective ILs form strong H-bonds, C–H... π , and cation– π interactions with lignin. However, the anions such as triflate, gentisate, histidinate, and bis(trifluoromethylsulfonyl)imide have high positive values of $\ln(\gamma)$. The $\ln(\gamma)$ and H^E values of lignin in 5670 ILs are provided in Tables S3 and S4. As the alkyl chain length of anion or cation increases, the $\ln(\gamma)$ and H^E of lignin were seen to be decreasing. For example, in a comparison between the imidazolium-based cations ([12Dmim]⁺, [Amim]⁺, [Emim]⁺, [Bmim]⁺, [Hmim]⁺, and [Omim]⁺) with all 90 investigated anions, the $\ln(\gamma)$ of lignin decreased (i.e., more negative) as alkyl chain length of the cation increased from [Emim]⁺ to [Omim]⁺ (Fig. S3a). A similar observation was made in the case of anions (Fig. S3b). A contrary observation was reported in the literature, where Wang et al.³⁸ reported that the solubility of lignin decreases with increase in the alkyl chain length of cations. This discrepancy may be due to the viscosity of ionic liquids. As the alkyl chain length of cation increases, the viscosity of IL also increased⁶⁸. This higher viscosity of IL restricts the mass transfer rate of liquids which results in the lowering of lignin solubility. A similar observation was also noticed in our previous studies⁶³, where spermidine and spermine showed stronger interaction with lignin, but results in lower biomass delignification due to their higher viscosity. Therefore, the combination of longer alkyl chain length of both the cation and anion resulted in weaker solvent for lignin solvation as compared to that of the combination of a highly polar and less polar ions. Based on these observations, the ions of IL should obey the following successive criteria: (1) either of the ions should be a good hydrogen bond acceptor or donor, and (b) another ion to be slightly polar (to weakly coordinate with counter ion thereby reduces the cross interactions between anion and cation). According to this thumb rule, the cations such as tetraalkylammonium, tetraalkylphosphonium, and alkylpyridinium are less polar, and the anions such as acetate, methyl carbonate, glycinate, alaninate, and lysinate are highly polar in nature. The interaction between the polar anion and the lignin is energetically much stronger than the interaction between anion and cation, resulting in a high lignin solvation capability.

Casas et al.^{53,55} performed comprehensive COSMO-RS calculations to screen all combinations of 20 cations and 16 anions of ionic liquids for lignin dissolution. Further, Balaji et al. (2012) screened 1156 ILs for extraction of lignin from biomass based on COSMO-RS calculated solubility parameters. Casas et al.⁵³ and Balaji et al.⁵² used lignin monomers as a lignin model compound for the COSMO-RS calculations and reported that 1-ethyl-3-methylimidazolium acetate, tetrabutylammonium fluoride, tetrabutylammonium acetate, and tetrabutylammonium chloride are potential solvents for lignin solvation. However, in our case tetrabutylammonium glycinate, tetrabutylammonium acetate, tetrabutylphosphonium glycinate, and 1,1-dipropylpyrrolidinium acetate are the better ILs for lignin dissolution. The discrepancy between the literature (Casas et al. and Balaji et al.) and results from our study are due to the chemical structure of lignin that is used in the COSMO-RS calculations. Lignin monomers and dimers are rich in hydrogen bond donors (see Fig. S4); thus, they interact with anions (rich in hydrogen bond acceptors). While the polymeric structure of lignin is less polar in nature and exhibits significant hydrophobic sites. Recently, Yu et al.⁵⁶ screened 3886 ILs for the lignin using the COSMO-RS model and reported that 1-propyl-1-methylpyrrolidinium chloride, 1-propyl-1-methylpyrrolidinium acetate, 1-methyl-1,8-diazabicyclo(5.4.0)undec-7-ene chloride, 1-methyl-1,8-diazabicyclo(5.4.0)undec-7-ene acetate, and 1-butyl-1,8-diazabicyclo(5.4.0)undec-7-ene chloride are better ILs for lignin dissolution. In the same study, it was also reported that phosphonium cation-based ILs show lower $\ln(\gamma)$ values of lignin than ammonium-based ILs. However, the present study results demonstrate that ammonium-based ILs are relatively better solvents for lignin than phosphonium cation-based ILs. Moreover, Yu et al.⁵⁶ ranked the solvents only based on the activity coefficient of lignin, however, excess enthalpy of lignin is also a critical parameter in ranking the solvents⁵³. Lignin monomeric and dimeric structures shows the poor linear regression fitting ($R^2 = 0.39\text{--}0.41$) between the H^E and lignin solubility (Fig. S2b). Therefore, the monomeric and dimer structures of lignin do not directly represent the lignin molecule due to the absence of many key linkages present in lignin. It is important to mention that the ILs based on tetraalkylammonium, tetraalkylphosphonium, and pyrrolidinium had a greater capacity for dissolving lignin. This was because these ILs had weaker interactions between their cations and anions, as a result, the ions of ILs had more opportunities to interact with lignin, thereby enhancing the lignin dissolution capability.

Sigma potential of anions and cations. The sigma potentials ($\mu(\sigma)$) of lignin, anions, and cations were calculated to understand the affinity of solvents for the surface polarity of lignin. The σ -potentials are divided into three main types: non-polar (green/yellow: $-0.01 \text{ e}/\text{\AA}^2 < \sigma < +0.01 \text{ e}/\text{\AA}^2$), H-bond acceptor (red: $\sigma < -0.01 \text{ e}/\text{\AA}^2$), and H-bond donor (blue: $\sigma > +0.01 \text{ e}/\text{\AA}^2$) regions (see Fig. 4). The sigma potential, $\mu(\sigma)$, of lignin is negative in both the negative and positive charge density regions ($\sigma < -0.01 \text{ e}/\text{\AA}^2$ and $\sigma > +0.01 \text{ e}/\text{\AA}^2$), indicating that lignin has a tendency to interact with both negative and positive polar surfaces of molecule (i.e., H-bond donors and acceptors in the solvent) (Fig. 4a). On the negative screening charge densities side ($\sigma > -0.01 \text{ e}/\text{\AA}^2$), the σ -potential value of anions is negative (-1.5 to $-2.4 \text{ kcal/mol \AA}^2$), implying that anions have more affinity to interact with the positive surface charge density of lignin or cation (i.e., H-bond donors) (Fig. 4a). In contrast, the σ -potential value of anions is positive ($0.5\text{--}1.0 \text{ kcal/mol \AA}^2$) in the region of positive screening charge densities ($\sigma > +0.01 \text{ e}/\text{\AA}^2$), which reflects that the anions lack electron donor surfaces. For cations such as [Ch]⁺ and [Emim]⁺, the $\mu(\sigma)$ value is negative in the positive screening charge density side ($\sigma > +0.01 \text{ e}/\text{\AA}^2$), indicating that [Ch]⁺ and [Emim]⁺ cations have higher affinity toward negatively charged surfaces of lignin and anions (Fig. 4). For system composed of lignin and [Ch]⁺ or [Emim]⁺-based ILs, the anions are more likely to interact with the cations than lignin since [Ch]⁺ and [Emim]⁺ are polar cations with higher tendency to interact with negatively charged surfaces thereby forming stronger electrostatic and hydrogen bondings. Therefore, the logarithmic activity coefficient and excess enthalpy of lignin in such ILs are weaker. However, in the case of [TBA]⁺, [TBP]⁺, and [DPrPyr]⁺, the cations lack both positive and negative surfaces (positive $\mu(\sigma)$ value), implying these cations have less tendency to interact with anions i.e., the interaction between anion and cation is weaker. Thus, both the

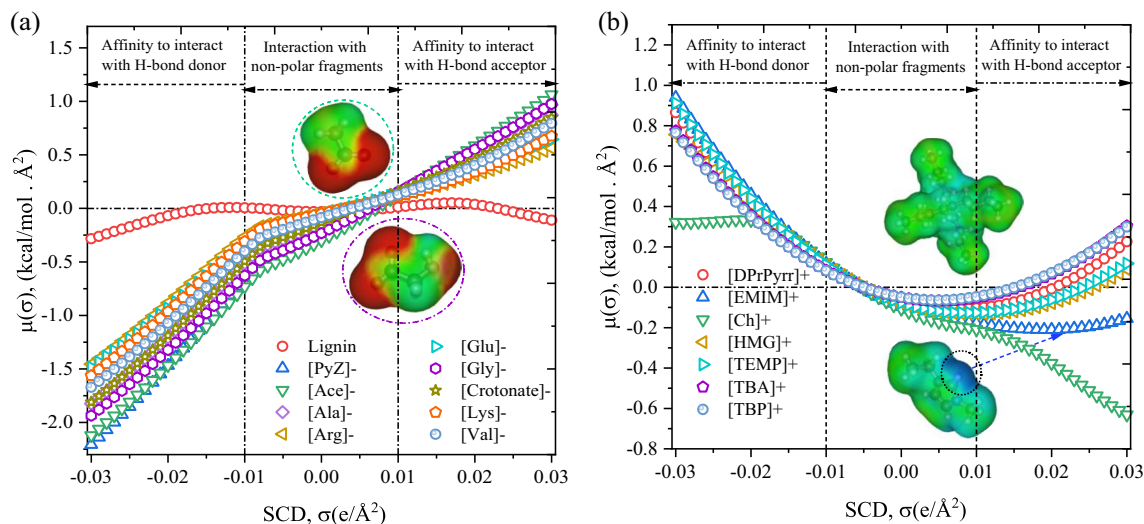


Figure 4. COSMO-RS predicted sigma potential of (a) lignin and anions, and (b) cations of ILs at 363.15 K. COSMOTherm version 19.0.1 was used to calculate the sigma potentials (<https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/solvation-chemistry/biovia-cosmotherm/>).

anions and cations have higher odds of interacting with lignin, thereby having a stronger interaction with lignin. Anions interacting through electrostatic and H-bonding interactions with lignin and the [TBA]⁺, [TBP]⁺, and [DPrPyrr]⁺ cations interact with lignin through vdW and cation- π interactions.

Parameters that influence lignin dissolution

Viscosity and interaction energy between anion and cation of ILs. The viscosity of solvent also plays an important role in the dissolution of a polymer and higher viscosity can restrict the effective mass transfer rate, which is known to have a negative effect on polymer dissolution^{27,33,69}. Ionic liquids present a higher viscosity (around 10–726 mPa s) than typical molecular liquids (around 0.2–10 mPa s). The viscosity of an IL depends on the combination of cation-anion and impurities present, and these relationships are well studied in the literature⁶². The ILs with lower excess enthalpy, H^E , and lower activity coefficient, $\ln(\gamma)$, values of lignin should also be cross-checked for viscosity of a solvent. In our previous publication on plastic dissolution in ILs, we predicted the viscosity of ILs using the COSMO-RS model and showed that the predicted values were highly correlated with the experimental viscosities⁵⁹. Therefore, the viscosity of tetrabutylammonium, tetrabutylphosphonium, cholinium, 1-ethyl-3-methylimidazolium, and 1,1-dipropylpyrrolidinium-based cations in combination with amino acid and carboxylate-based anions were predicted using the COSMO-RS model.

For calculating viscosity using the COSMO-RS model, the conformers of each ionic liquid were generated using the Turbomole^{70,71} and BIOVIA COSMOconfX2022 package, and the predicted viscosities are summarized in Fig. 5. The viscosity of liquids are calculated from a quantitative-structure property relationship (QSPR) using molecular descriptors (molecular surface area, σ -moments, pure compound entropy, and number of ring atoms in the compound) that are computed by COSMOtherm⁷². From Fig. 5, the viscosity of ammonium, pyridinium, and phosphonium-based ionic liquids had a lower viscosity than cholinium and imidazolium-based ionic liquids. It has been reported that as the alkyl chain lengths of either the cation or anion increased, the viscosity of ionic liquids increased and the H-bond acidity decreased⁷³. A similar pattern was also observed in our study. As the alkyl chain length of the anion increased ([Gly]⁻, [Ala]⁻, [Val]⁻, and [Lys]⁻), the viscosity of amino acid-based ILs (e.g.: [TBA][Ala], [TBA][Gly], [TBA][Val], and [TBA][Lys]) increased. The viscosity of ILs with tetraalkylammonium, tetraalkylphosphonium, and pyridinium-based cations was lower than cholinium and imidazolium-based ILs with the same anion (e.g., [Lys]⁻ and [Ace]⁻). On the other hand, the viscosity of ammonium and phosphonium-based ILs were predicted to be lower than the traditional ionic liquids such as [Ch][Lys], [Ch][Ace], and [Emim][Ace]. Similar trends were also observed for interaction energy between anion and cation of ILs (Fig. S5). Tetrabutylammonium, tetrabutylphosphonium, and 1,1-dipropylpyrrolidinium-based ILs have shown lower interaction energies between cation and anion, which is in line with sigma potential observations (Fig. 4). The ILs [Ch][Lys], [Ch][Ace], [Emim][Lys], and [Emim][Ace] showed stronger and higher interaction energies between anion and cations (Fig. S5), which helps explain why the viscosity of cholinium and imidazolium-based ILs are higher than tetraalkylammonium, tetraalkylphosphonium, and pyridinium-based ILs.

Solubility parameters. The solubility parameter (SP) of a molecular species is a vital property that characterizes polarity and quantifies the ‘like-seeks-like’ principle, which is used in chemistry to screen solvents for dissolution of solutes. There are several literature reports on the delignification of biomass using ionic liquids^{13,38,74–76}. Therefore, we explored Hansen solubility parameters as a measure of the capability of ionic liquids to delignify biomass. Hansen solubility parameters, δ , characterize the affinity of solvents towards polymers dissolution using three parameters δ_d for dispersion or van der Waals, δ_p for polarity, which is related to dipole moments, and δ_h for hydrogen bonding. The solubility parameters of lignin are 13.5 MPa^{1/2} polar (δ_p), 11.3

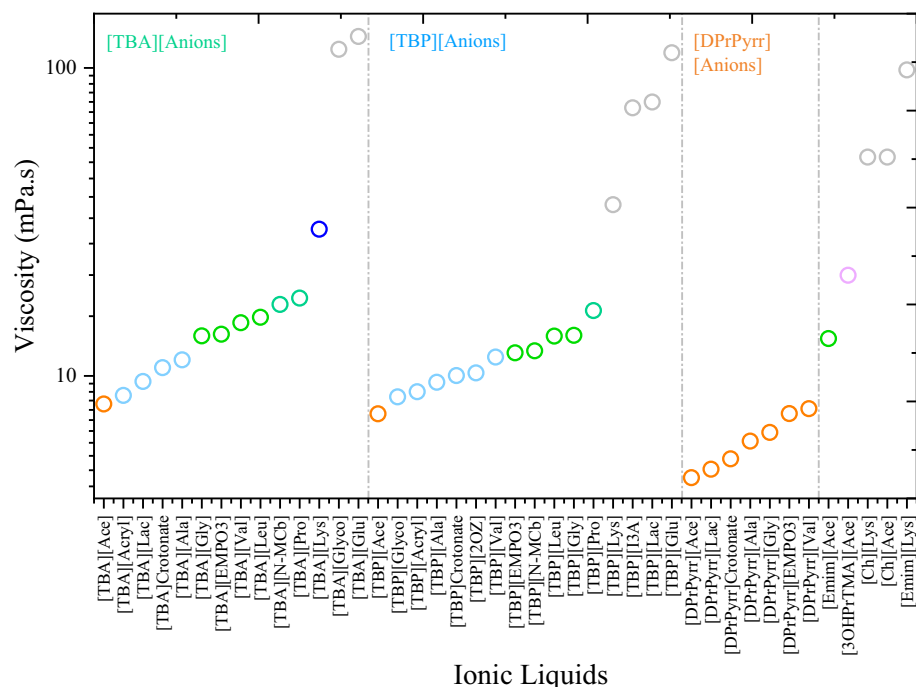


Figure 5. COSMO-RS-based predicted viscosities of pure ionic liquids at 363.15 K. COSMOTerm version 19.0.1 was used to calculate the sigma potential (<https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/solvation-chemistry/biovia-cosmotherm/>).

MPa^{1/2} hydrogen-bonded (δ_h), 16.7 MPa^{1/2} dispersion (δ_d), and 24.3 MPa^{1/2} total (δ_t)^{63,77}. The ‘like-seeks-like’ principle suggests that for cases in which solubility parameters of lignin and IL are closer, the affinity between the lignin polymer and the IL will be higher (i.e., ILs with similar solubility parameters as lignin will have greater dissolution capability). The solubility parameters of tetrabutylammonium, tetrabutylphosphonium, cholinium, 1-ethyl-3-methylimidazolium, and 1,1-dipropylpyrrolidinium-based ionic liquids were computed using the COSMO-RS model, and the results are depicted in Fig. 6. The solubility parameters of ILs were calculated using the COSMOquick⁷⁸. The detailed calculation procedure of HSP using COSMOquick is reported elsewhere^{79,80}. In our earlier publication on the development of solubility parameters for lignin and ILs, multiscale molecular simulation approaches were used to successfully predict the HSPs of lignin and ILs. It has also been demonstrated that the COSMO-RS model accurately predicts the HSPs of IL without any experimental data, hence here we used the COSMO-RS model to predict the solubility parameters of ILs⁸⁰.

Figure 6 reports the HSPs of lignin and ionic liquids along with their RED (relative energy difference) values of lignin in ILs. The RED values are calculated as the ratio of radius of interaction (R_a) to the 3D sphere radius of the solute (R_0)⁸⁰. If RED < 1, then the affinity of the IL towards lignin is said to be high. While if RED > 1, the affinity between the IL and lignin is weaker and leads to lower biomass delignification. From the RED point of view, [TBA]⁺, [TBP]⁺, and [DprPyrr]⁺-based ionic liquids had lower RED values of lignin than [Ch]⁺ and [Emim]⁺-based ILs. The ILs based on [Ch]⁺ and [Emim]⁺ had higher polar and hydrogen-bonded contributions, which is due to the stronger polarity and hydrogen-bonding capability of [Ch]⁺ and [Emim]⁺-based ionic liquids. The larger RED values of [Ch]⁺ and [Emim]⁺-based ILs are due to the higher polarity and hydrogen bonding contributions, which further leads to stronger interactions between anion and cation of the IL. The stronger anion and cation interactions (Fig. S5) result in higher viscosity (Fig. 5) and weaker solvents for lignin dissolution. The RED ranking of anions with [TBA]⁺, [TBP]⁺, and [DprPyrr]⁺-based cations for lignin dissolution was as follows: [Ala]⁻ > [Ace]⁻ > [Gly]⁻ > [Lys]⁻ > [Val]⁻.

Kamlet–Taft parameters and dissociation constant. Kamlet–Taft parameters, first introduced in 1976, are also commonly used solvent scales for predicting the behavior of solvents. The Kamlet–Taft solvent parameter π^* defines the polarizability of a molecule, which is the ability of a solvent to stabilize a charge or a dipole by virtue of its dielectric effect. The α parameter describes the strength of a hydrogen bond donor (i.e., hydrogen bond acidity), while the β parameter is used for describing the strength of a hydrogen bond acceptor (i.e., hydrogen bond basicity). Generally, ILs have high π^* and β values and moderate α values, while β primarily depends on the choice of anion and α depends on the cation. It was reported that hydrogen bond basicity β of IL was the most important parameter for the dissolution of cellulose and lignin, as well as biomass fractionation^{56,81}. Dissolution of biopolymers occurs in ILs when the value of β is greater than approximately 0.8. The Kamlet–Taft ‘ β ’ parameter of tetrabutylammonium, tetrabutylphosphonium, cholinium, and [Emim]-based ILs were taken from the literature and depicted in Fig. 7^{82–84}. From Fig. 7, amino acid-based anions had higher β values than carboxylate-based anions, indicating that amino acid-based anions in combination with tetrabutylammonium

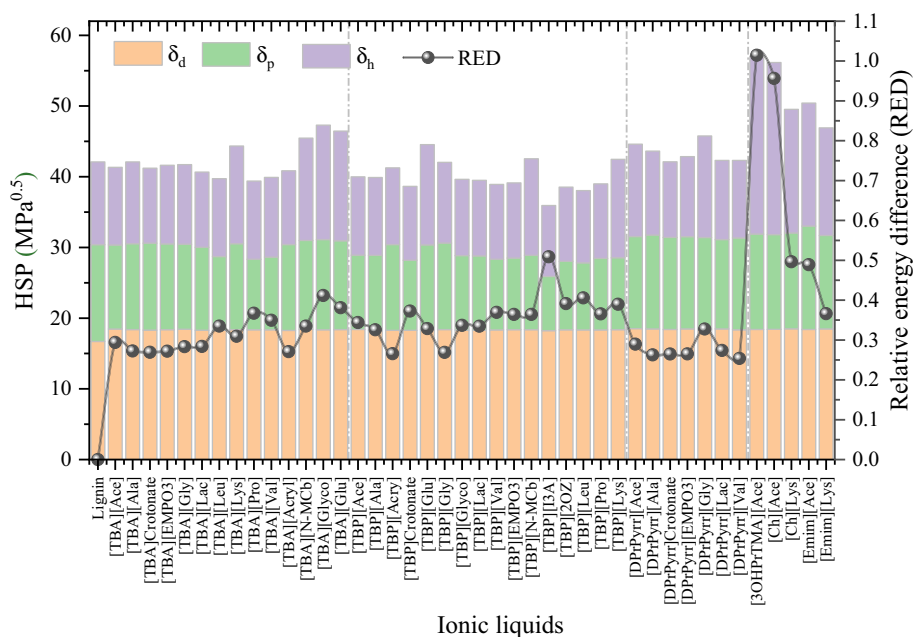


Figure 6. The COSMO-RS-based predicted Hansen Solubility Parameters of ionic liquids. COSMOquick tool version 1.7 is used to predict the HSP of ILs (<https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/solvation-chemistry/biovia-cosmoquick/>).

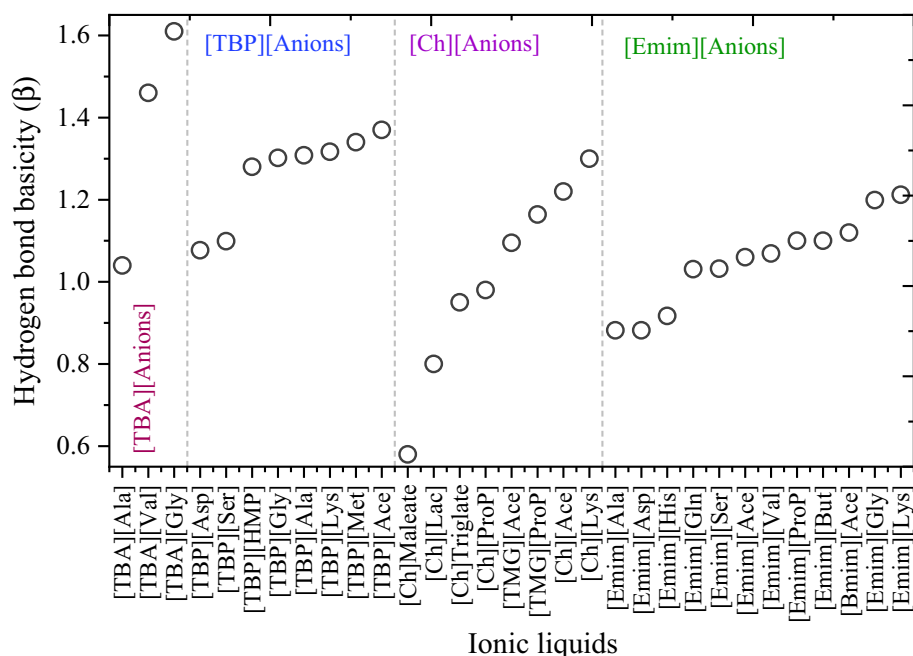


Figure 7. Hydrogen-bond basicity (β) of ionic liquids. The experimental β values are taken from literature.

and tetrabutylphosphonium cations could effectively dissolve the lignin and deconstruct biomass. The higher HB basicity of $[TBA]^+$ and $[TBP]^+$ -based ILs were due to the weaker interaction energy between anion and cation of ILs (Fig. S5). The weaker interaction energy between anion and cation of ILs leads to the presence of more free anions, thus the glycinate, valinate, and lysinate-based anions exhibit higher HB basicity values. The ranking of ILs with high β values was: $[TBA][Gly] > [TBA][Val] > [TBP][Ace] > [TBP][Lys] > [Ch][Lys]$.

In addition to Kamlet–Taft parameters, the dissociation constant (pK_a) of ILs plays a critical role in the deprotonation of lignin. Recently, Mohan et al.⁷ reported the efficacy of $[Ch][Lys]$ over $[Ch][Oct]$ for lignin dissolution. The pK_a values of lignin protons lie in the range of 1–5 (carboxylic protons) and 6–11 (phenolic protons)⁸⁵. This implies that all lignin protons would be deprotonated (hence enhancing the solubility of lignin)

in the presence of a chemical with a similar pK_a value or greater than 11. The pK_a of carboxylate anions are in the range of 4.27–5.5 (acetate to decanoate), while the pK_a 's of amino acid-based anions are 2.74 (COOH)–12.41 (NH₂) (Fig. S6). Amino acid based ILs are therefore able to deprotonate the protons of lignin with a pK_a value up to 10.3, which implies the deprotonation of almost all the available carboxylic and phenolic protons of lignin, leading to a maximum charge density and higher delignification.

Molecular dynamics simulations of lignin-ionic liquids. Molecular dynamics simulations are a prevailing computational technique for investigating the interactions among molecules in binary mixtures and are a useful tool for understanding the delignification of lignocellulosic biomass in ionic liquids^{7,77}. From the COSMO-RS calculations, [TBA]-based ILs stood out as potential solvents for lignin dissolution, and we investigated biomass delignification with [TBA]-based ILs and compared the results to cholinium lysinate. For MD simulations, lignin model polymer structure is composed of 26 units of G-, S-, and H-units with a molecular weight of ~5 kDa (Fig. S7).

Radius of gyration and solvent-accessible surface area analysis. The radius of gyration (R_g) of lignin (DP-26) in different IL environments were measured to study the compactness of the lignin polymer structure. For polymers, a lower R_g corresponds to the collapsed-like structure, whereas a larger value of R_g corresponds to the coil or extended polymer-like structure. Figure 8 depicts the probability distribution of the radius of gyration of lignin in different IL systems. The average R_g values were 15.75 Å, 17.92 Å, 18.45 Å, 18.85 Å, and 19.96 Å for [TBA][Val], [TBA][Ace], [TBA][Ala], [TBA][Gly], and [Ch][Lys], respectively. According to this R_g values (> 14 Å), lignin adopted an extended polymer-like structure, while if the R_g value of lignin is lower than 13 Å, lignin forms collapsed compact conformation which was generally obtained in the water environment⁸⁶. Representative snapshots of lignin (i.e., coil-like structure) from the simulations in ILs were shown in Fig. S8.

The solvent-accessible surface area (SASA) of a polymer is a useful metric for measuring the interface between a solute and its solvent i.e., the surface area of solute that is in contact with the solvent in which it resides. For solutes, hydrophobicity increases with a decreasing SASA. The probability distribution of lignin SASA in the investigated IL systems is illustrated in Fig. 9. From Fig. 9, the SASA of lignin in all investigated ILs was similar and higher than 3200 Å², indicating lignin makes more energetically favorable interactions with the solvent than with itself, resulting in solubilization of lignin. The conclusions drawn from the SASA's probability distribution analysis support those observed from the analysis of R_g 's distribution.

Radial distribution functions and intra and intermolecular hydrogen bonds of lignin. To obtain the structural arrangements and microscopic interactions, radial distribution functions ($g(r)$ or RDF) were computed between lignin and the IL solvent. Radial distribution functions are the probability of finding a molecule at a distance of ' r ' from the reference molecule⁷. The RDF plots are a useful tool for understanding the structural and explicit interactions of lignin and solvents. In our case, the RDFs are plotted between the oxygen (O) atom of lignin molecule with O1 (acceptor) atom of anion and N atom of cations, and the results are depicted in Fig. 10a,b.

Figure 10a shows the RDF plot between the lignin and anions of ILs. The first and largest solvation shell of RDF peak between lignin and anion was attained at 2.65 Å with a $g(r)$ peak height of 5–10, indicates that anion

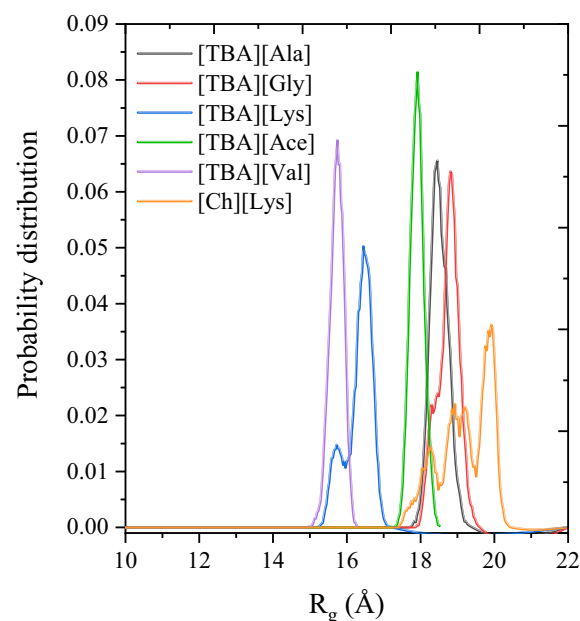


Figure 8. The probability distribution of R_g for the lignin polymer in ionic liquids.

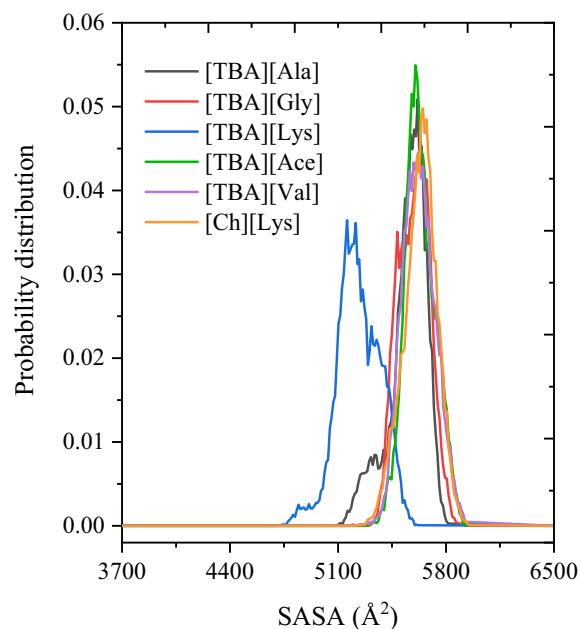


Figure 9. The probability distribution of SASA for the lignin polymer in ionic liquids.

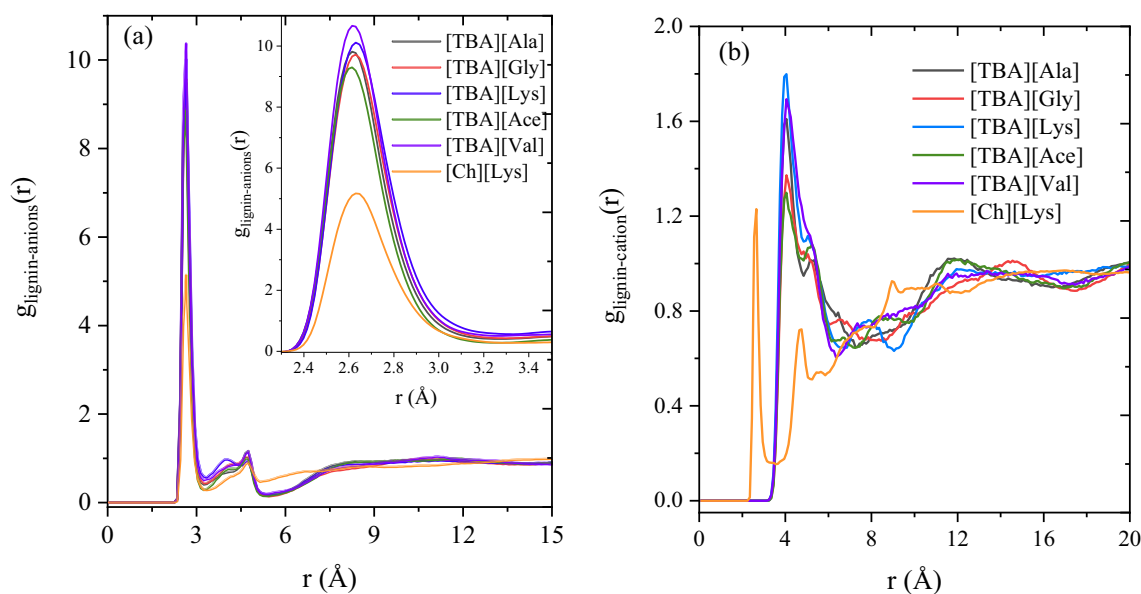


Figure 10. RDF plot between the O atom of lignin with (a) O1 atom of anion, and (b) N1 atom of cation for different lignin-IL systems.

forms regular and definite coordination spheres around lignin moiety at a distance of 2.65 Å and the RDF plot is primarily dominated by the first coordination shell. The secondary solvation shells are less ordered when compared to the first solvation shell (Fig. 10a). It is interesting to note that the height of RDF peak $g(r)$ for lignin-anion from [TBA]-based ILs are ~ 10 , implying that the contact probability between lignin and anion of [TBA] [-ILs] are almost 10 times in their first solvation shell (3.45 Å), which is higher than [Ch][Lys]'s RDF peak. In the case of lignin-cation RDFs, cations also form a regular and well-ordered shell around the lignin molecule, where the [TBA]⁺ cation approaches the lignin moiety at a distance of 4.0 Å, while the [Ch]⁺ approaches lignin moiety at a distance of 2.65 Å. For the [TBA]-based systems, the cation [TBA]⁺ had higher $g(r)$ value of lignin molecule than [Ch]⁺ cation, indicating that the contact probability between [TBA]⁺ cation and lignin is stronger than [Ch]⁺-lignin. On the other hand, the RDF peak between anion and cation of IL were also measured and depicted in Fig. 11. From Fig. 11, the peak of first solvation shell between lysinate and cholinium was obtained at distance of 2.65 Å with a peak height of ~ 8 , indicating that the interactions between anion and cation of [Ch]

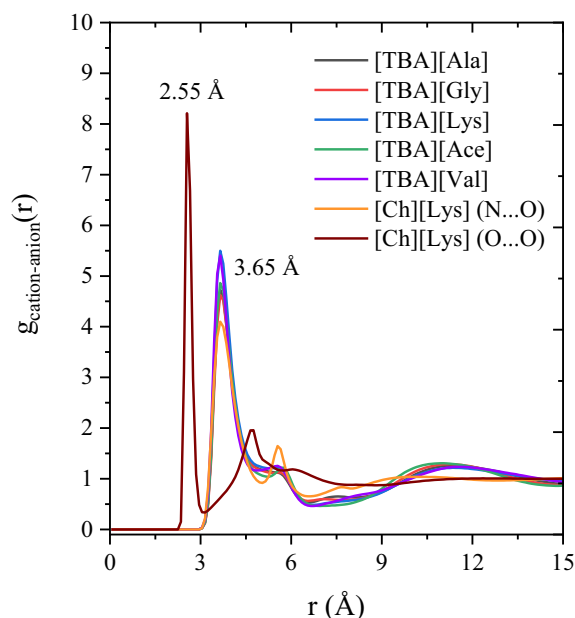


Figure 11. RDF plot between the O atom of anion with N1 atom of cation in different lignin-IL systems.

[Lys] were much stronger than the [TBA]⁺-based ILs ($g(r)$ peak heights are in the range of 4.5–5). The stronger the anion and cation contact probability, the weaker the solvation capability of a solvent because of its higher viscosity. From these findings, it is confirmed that lignin solubility predominantly depends on and can thus be manipulated by the choice of the anion and cation of IL. The less polar cation and high polar anions results in stronger van der Waals interactions and electrostatic interactions with lignin, thereby stabilizing the lignin molecule and enhancing the lignin solubility.

In addition to the R_g , SASA, and RDF plots, the average number of hydrogen bonds (HBs) per lignin molecule were calculated to further understand the structural rigidity of lignin polymer. The intra and inter-molecular hydrogen-bonding patterns in the lignin polymer and lignin-ILs were monitored as a function of simulation time and the results are shown in Fig. 12. Figure 12a shows the intermolecular hydrogen bonding between lignin and anions of the various ILs. From Fig. 12a, amino acid-based anions ([Gly]⁻, [Ala]⁻, [Val]⁻, and [Lys]⁻) had a higher

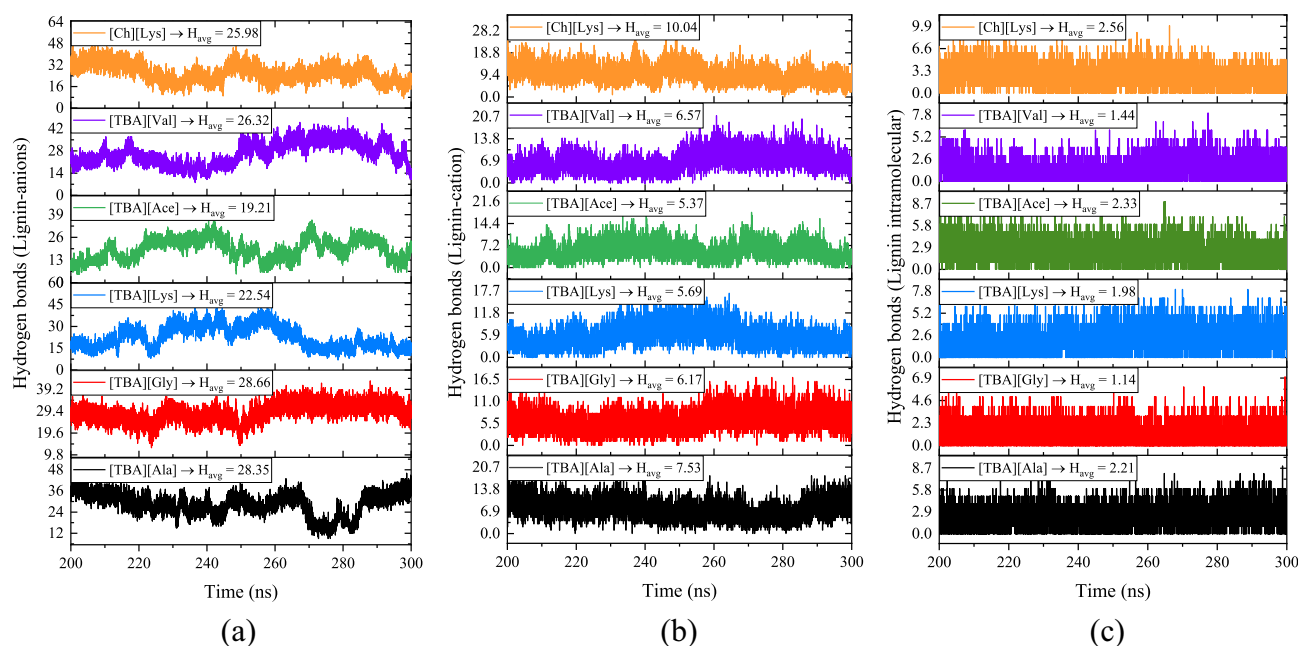


Figure 12. MD calculated average number of hydrogen bonds between the (a) lignin-anion, (b) lignin-cation, and (c) within the lignin molecules in different IL-lignin systems. Origin 2020 SR1 software is used to draw this plot (<https://www.originlab.com/2020>).

number of hydrogen bonds with lignin than carboxylate-based anion (e.g. [Ace]⁻), indicating that amino acid anion forms multiple hydrogen bonds with lignin than acetate. However, [TBA][Ala], [TBA][Gly], and [TBA][Val] formed a marginally higher number of hydrogen bonds with lignin than [Ch][Lys]. This may be due to the formation of additional hydrogen bonds between cholinium and lysinate ions. On the other hand, HBs between lignin and cation of ILs were also characterized and are shown in Fig. 12b. [TBA]⁺ cation had a lower number of H-bonds with lignin than [Ch]⁺ which is due to the higher polarity of the cholinium cation. The additional hydroxyl group of cholinium forms HBs with lignin, resulting in a higher number of lignin-[Ch]⁺ H-bonds. The average number of intramolecular H-bonds within the lignin molecule was also measured and [Ch][Lys] had a higher number of intrachain H-bonds (Fig. 12c). The higher number of intramolecular H-bonds of lignin in [Ch][Lys] is due to the weaker hydrogen bonding dynamics of lignin-anion and lignin-cation, which is further evidenced by their RDF plots (lower contact probability). The significant loss of intrachain H-bonds in [TBA][Gly] and [TBA][Val] ILs revealed effective lignin dissolution in these solvent environments. The breakdown of intramolecular hydrogen bonds of lignin could be due to the favorable interactions between the lignin polymer and [TBA]-based ILs. On the other hand, the interaction within the lignin molecule is less favorable in [TBA]-based ILs, thus could lead to a better solvent for lignin.

Mass fractal dimension factor. The mass fractal dimension (α) factor, which is related to the solubility of a polymer (as α below 2 indicates a good solvent, while above indicates a poor solvent). The mass fractal dimension factor was obtained by fitting the relationships between lignin size and structural observables such as the radius of gyration (R_g) and solvent accessible surface area (SASA). The summary of molecular dynamics simulations for different lignin polymers in [TBA][Gly] was presented in Table S5 and the chemical structures and linkages of lignin are reported in Tables S6–S16. Within the (bio)polymer physics, these observables are typically related through a power law and the governing equation of scaling properties are as follows.^{87,88}

$$O = AM^{\nu} \quad (3)$$

$$\alpha = \frac{2}{5\nu - 1} \quad (4)$$

The power law described in Eq. 3 is used to fit the relationships between structural observables, O , such as R_g and SASA, and the molecular weight, M , of lignin polymers. The fitting parameters A and particularly ' ν ' have meaning when interfaced with polymer theory and used to compute the mass fractal dimension (α) factor (Eq. 4).

Figure 13 shows the relationship between the structural observables, R_g /SASA and molecular weights of lignin polymers. From 13a, the R_g is sensitive to lignin polymer molecular weight and the scaling parameter ' ν ' is >0.334 , reflective a good solvent in polymer theory⁸⁸. On the other hand, SASA for a given lignin polymer is largely determined by its molecular weight and variation within SASA values within a trajectory is comparatively smaller than it was for R_g . For a sphere of uniform density, the SASA should increase with the $M^{2/3}$ _{sphere}. The scaling parameter $\nu > 0.67$ (i.e., 0.94), indicating that there is an excess surface area for real lignin polymers relative to an ideal spherical globule^{87,88}. Mass-fractal dimensions α , closer to 1 (i.e., below 2) are “good” solvents while fractal dimensions near 3 or above are “bad/poor” solvents. From the literature, in water, the α value of lignin is greater than 3 and thus water is a poor solvent for lignin dissolution^{87,89}.

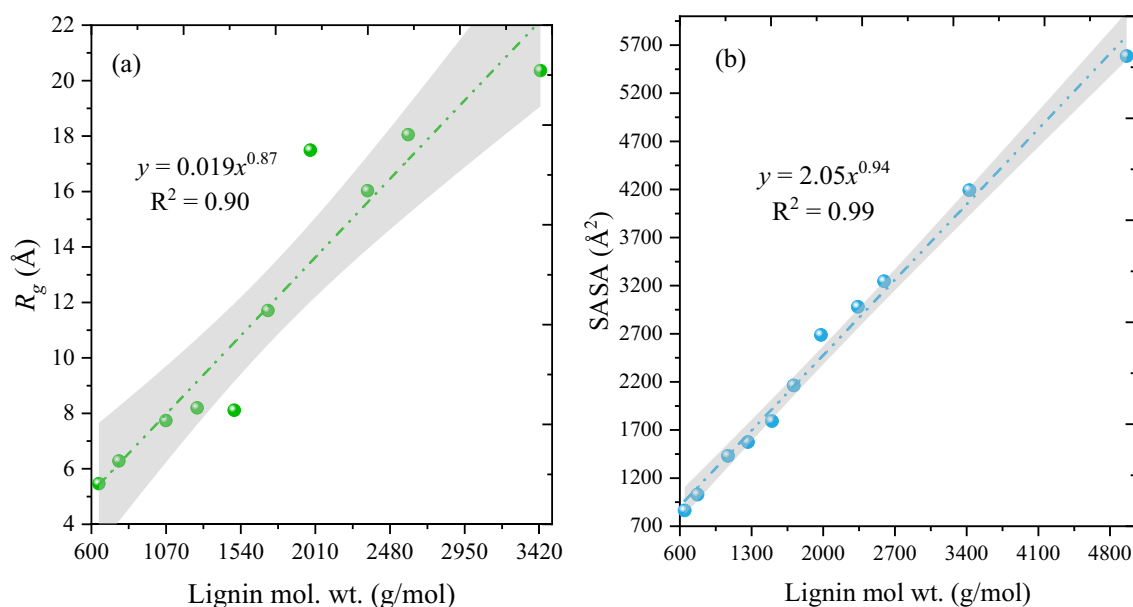


Figure 13. Relationship between (a) R_g and (b) SASA of lignin and the molecular weight of the lignin polymer.

Conclusions

The present study reports the screening of large number of ionic liquids for the dissolution of lignin using the COSMO-RS model. A total of 5670 ILs (combination of 63 cations and 90 anions) were screened for lignin by predicting the logarithmic activity coefficient ($\ln(\gamma)$) and excess enthalpies (H^E) of lignin in ILs. Based on the COSMO-RS predictions, it can be concluded that anions such as acetate, methyl carbonate, glycinate, alaninate, and lysinate are predicted to dissolve the lignin more efficiently in combination with the cations like tetraalkylammonium, tetraalkylphosphonium, and pyridinium. After screening 5670 ILs, the dissolution properties such as interaction energy between anion and cation, viscosity, Hansen solubility parameters, dissociation constants, and Kamlet-Taft parameters have been evaluated to assess the lignin dissolution. It has been noticed that [TBA]⁺ and [TBP]⁺-based ILs were seen to be potential solvents for lignin dissolution as they show weaker interactions with their counter ion, possess lower viscosity, closer HSP values, and higher hydrogen bonding basicity. Furthermore, MD simulations were performed for lignin and [TBA]-based ILs and the results were compared with [Ch][Lys]. MD simulation results show that the lignin adopts an extended polymer-like structure (higher R_g and SASA) in both [TBA]⁺ and [Ch]-based ILs, however, the contact probability and the intermolecular hydrogen bonds between lignin and [TBA]-based ILs were stronger than that of lignin and [Ch][Lys], revealing the potential capabilities of [TBA]-based ILs for lignin dissolution. It is worth mentioning that the cation and anion of IL play a crucial role in the dissolution of lignin. In addition, the mass fractal dimensional (α) factor was also evaluated for [TBA][Gly] to assess the potential capability of [TBA]-based ILs and observed that the value of ' α ' factor below 1 results in a good solvent for lignin. This work provides an effective strategy for fast screening of appropriate IL for biomass delignification and the analysis of obtained results help to comprehend the dissolution mechanism of lignin in ionic liquids.

Computational details

COSMO-RS calculations. The COSMO-RS calculations were carried out to measure the capability of ILs for lignin dissolution based on the activity coefficient and excess enthalpy of lignin in ILs. The geometries of all the investigated lignin polymer and ions (anions and cations) of ionic liquids were generated in the Avogadro software version 1.2.0⁹⁰. The geometries of investigated molecules were fully optimized using *Gaussian09* package at B3LYP theory and 6-311 + G(d, p) basis set⁹¹⁻⁹³. COSMO files were generated at the BVP86/TZVP/DGA1 level of theory and basis set using the keyword "scrf=COSMORS"^{51,58,94}. The generated COSMO files were then used as an input to the COSMOtherm (version 19.0.1, COSMOlogic, Leverkusen, Germany) package with BP_TZVP_19 parametrization, which was used to calculate the sigma potentials, viscosity, solubility parameters of ILs, and the logarithmic activity coefficient ($\ln(\gamma)$) and excess enthalpy (H^E) of binary mixtures (lignin and IL)^{95,96}.

The logarithmic activity coefficient ($\ln(\gamma)$) of component i is related to the chemical potential μ_i is given as following Equations^{7,97}:

$$\ln(\gamma_i) = \left(\frac{\mu_i - \mu_i^0}{RT} \right) \quad (1)$$

where μ_i^0 is the chemical potential of the pure component i , R and T are the real gas constant and absolute temperature. Excess enthalpy of a binary mixture was calculated using Eq. (2)⁹⁴

$$H_M^E = \sum x_i H_i^E = \sum x_i [H_{(i,mixture)} - H_{(i,pure)}] \quad (2)$$

where H_m^E is the excess enthalpy of each molecule in the mixture, defined as the enthalpy difference between component i in the mixture and in the pure state, and the total excess enthalpy of a mixture is the algebraic sum of electrostatic (misfit), hydrogen bonding, and van der Waals interactions. To calculate H^E and $\ln(\gamma)$ in the COSMO-RS model, the mole fractions of lignin, cation, and anion was set to 0.2, 0.4, and 0.4, respectively. More details of COSMO-RS calculation in predicting the excess enthalpies and activity coefficients are provided in the COSMOtherm's user Manual⁹⁸.

In general, ILs might either be modelled by different 'conformers' with respect to the relative arrangement of their ions in the COSMO calculation, or simply as two independent, equimolar, ionic species, whose sigma profiles can be combined freely. The second approach is much efficient and was successfully implemented in earlier works^{27,54,59,66}. Similar to that, it is carried out in this study. After screening thousands of ILs using the COSMO-RS model, separate quantum chemical (QC) calculations were performed for ILs (combined equimolar ionic species) to predict the structural properties such as viscosity, solubility parameters, and interaction energies between anion and cation of ILs. We performed a search for conformations of ionic liquids using the Turbomole^{70,71} and BIOVIA COSMOconfX2022 program (version 22.0.0, COSMOlogic, Leverkusen, Germany), which automatically identifies conformers relevant for subsequent COSMO-RS calculations. COSMO calculations within the COSMOConf were performed using the BP-TZVP method and basis set.

Molecular dynamics simulations. Molecular dynamics (MD) simulations of lignin in tetrabutylammonium glycinate [TBA][Gly], tetrabutylammonium alaninate [TBA][Ala], tetrabutylammonium acetate [TBA][Ace], tetrabutylammonium lysinate [TBA][Lys], tetrabutylammonium valinate [TBA][Val], and cholinium lysinate [Ch][Lys] were performed to understand the effect of ILs on the delignification of lignocellulosic biomass. The ILs [TBA][Gly], [TBA][Ala], [TBA][Ace], [TBA][Lys], [TBA][Val], and [Ch][Lys] were chosen based on the COSMO-RS predictions. The lignin model polymer structure is taken from our previous study⁷⁷ and composed of 26 units of G-, S-, and H-units with a molecular weight of ~ 5 kDa (Fig. S7). To build the lignin

structure, all feasible lignin linkages were joined together, and the sequence of the polymer is shown in Table S6. All MD simulations were carried out using the NAMD package⁹⁹ and the CHARMM force field parameters were used for all the investigated molecules. The force field parameters for lignin was taken from Vermaas et al.¹⁰⁰ and for ILs, the force field parameters were developed by following CHARMM-GUI tool^{101,102}.

The initial configuration for all the investigated systems was prepared using PACKMOL¹⁰³. The lignin molecule was solvated in five different [TBA]-based ILs and the results were compared with [Ch][Lys]. The simulation details such as the number of solvent molecules, lignin molecules, and final box size are summarized in Table S17. The potential energy of the system was first minimized for 200,000 steps using a steepest-descent algorithm. After minimization, the system was heated and equilibrated for 15 ns under the NPT ensemble using the Langevin thermostat and Nose–Hoover Langevin barostat^{104,105}. The SHAKE algorithm was implemented to constrain all bonds involving hydrogen atoms¹⁰⁶. The Particle Mesh Ewald (PME) method was implemented to treat long-range electrostatic interactions at a cut-off distance of 12 Å with an accuracy of 10^{-6} (PME tolerance)¹⁰⁷. Three separate production runs for 300 ns were performed on each simulated system, starting with a different velocity distribution. A 2 femtoseconds (fs) time step was used to integrate the equations of motion and the coordinates of the system were saved every 5 ps for structural and dynamics analysis. All MD simulation trajectories were visualized and analyzed by using TCL scripts, TRAVIS package, and Visual Molecular Dynamics (VMD) tool^{108–110}. The non-bonded interaction energies and the number of hydrogen bonds between lignin and ILs were calculated per mole of lignin.

Data availability

All data generated or analyzed during this study are included in this published article and its Supplementary Information file.

Received: 25 July 2022; Accepted: 29 November 2022

Published online: 06 January 2023

References

1. Yang, L., Seshan, K. & Li, Y. A review on thermal chemical reactions of lignin model compounds. *Catal. Today* **298**, 276–297 (2017).
2. Wang, H., Yang, B., Zhang, Q. & Zhu, W. Catalytic routes for the conversion of lignocellulosic biomass to aviation fuel range hydrocarbons. *Renew. Sust. Energ. Rev.* **120**, 109612 (2020).
3. Patri, A. S. et al. A multifunctional cosolvent pair reveals molecular principles of biomass deconstruction. *J. Am. Chem. Soc.* **141**, 12545–12557. <https://doi.org/10.1021/jacs.8b10242> (2019).
4. Grossman, A. & Vermerris, W. Lignin-based polymers and nanomaterials. *Curr. Opin. Biotechnol.* **56**, 112–120. <https://doi.org/10.1016/j.copbio.2018.10.009> (2019).
5. Mohan, M. et al. Optimization and hydrolysis of cellulose under subcritical water treatment for the production of total reducing sugars. *RSC Adv.* **5**, 103265–103275. <https://doi.org/10.1039/C5RA20319H> (2015).
6. Huijgen, W., Smit, A., Grisel, R. & van Hal, J. *Recent Advances in Organosolv Fractionation of Lignocellulosic Biomass*. (Petten: ECN, 2015).
7. Mohan, M. et al. Towards understanding of delignification of grassy and woody biomass in cholinium-based ionic liquids. *Green Chem.* **23**, 6020–6035 (2021).
8. Huang, K. et al. Integration of acetic acid catalysis with one-pot protic ionic liquid configuration to achieve high-efficient biorefinery of poplar biomass. *Green Chem.* **23**, 6036–6049 (2021).
9. McCann, M. C. & Carpita, N. C. Biomass recalcitrance: A multi-scale, multi-factor, and conversion-specific property. *J. Exp. Bot.* **66**, 4109–4118. <https://doi.org/10.1093/jxb/erv267> (2015).
10. Vermaas, J. V. et al. Mechanism of lignin inhibition of enzymatic biomass deconstruction. *Biotechnol. Biofuels* **8**, 1–16 (2015).
11. van Rijn, R., Nieves, I. U., Shanmugam, K., Ingram, L. O. & Vermerris, W. Techno-economic evaluation of cellulosic ethanol production based on pilot biorefinery data: A case study of sweet sorghum bagasse processed via L+ SScF. *BioEnergy Res.* **11**, 414–425 (2018).
12. Yang, B. & Wyman, C. E. Pretreatment: The key to unlocking low-cost cellulosic ethanol. *Biofuels Bioprod. Bioref. Innov. Sustain. Econ.* **2**, 26–40 (2008).
13. Mohan, M., Deshavath, N. N., Banerjee, T., Goud, V. V. & Dasu, V. V. Ionic liquid and sulfuric acid-based pretreatment of bamboo: Biomass delignification and enzymatic hydrolysis for the production of reducing sugars. *Ind. Eng. Chem. Res.* **57**, 10105–10117. <https://doi.org/10.1021/acs.iecr.8b00914> (2018).
14. Mohan, M., Banerjee, T. & Goud, V. V. Hydrolysis of bamboo biomass by subcritical water treatment. *Biores. Technol.* **191**, 244–252 (2015).
15. Baker, D. A., Gallego, N. C. & Baker, F. S. On the characterization and spinning of an organic-purified lignin toward the manufacture of low-cost carbon fiber. *J. Appl. Polym. Sci.* **124**, 227–234 (2012).
16. Sipponen, M. H., Smyth, M., Leskinen, T., Johansson, L.-S. & Österberg, M. All-lignin approach to prepare cationic colloidal lignin particles: Stabilization of durable Pickering emulsions. *Green Chem.* **19**, 5831–5840 (2017).
17. Kubo, S. & Kadla, J. F. Poly (ethylene oxide)/organosolv lignin blends: Relationship between thermal properties, chemical structure, and blend behavior. *Macromolecules* **37**, 6904–6911 (2004).
18. Hilburg, S. L. et al. A universal route towards thermoplastic lignin composites with improved mechanical properties. *Polymer* **55**, 995–1003 (2014).
19. Gao, R., Li, Y., Kim, H., Mobley, J. K. & Ralph, J. Selective oxidation of lignin model compounds. *Chemsuschem* **11**, 2045–2050 (2018).
20. Sipponen, M. H., Lange, H., Crestini, C., Henn, A. & Österberg, M. Lignin for Nano- and Microscaled Carrier Systems. (2019).
21. Johnson, C. W. et al. Enhancing muconic acid production from glucose and lignin-derived aromatic compounds via increased protocatechuate decarboxylase activity. *Metab. Eng. Commun.* **3**, 111–119 (2016).
22. Rinaldi, R. et al. Paving the way for lignin valorisation: Recent advances in bioengineering, biorefining and catalysis. *Angew. Chem. Int. Ed.* **55**, 8164–8215 (2016).
23. Vanholme, R., Demedts, B., Morreel, K., Ralph, J. & Boerjan, W. Lignin biosynthesis and structure. *Plant Physiol.* **153**, 895–905. <https://doi.org/10.1104/pp.110.155119> (2010).
24. Liu, Z.-H. & Yuan, J. S. Systems biology analyses of lignin conversion. In *Lignin Valorization* (ed. Beckham, G. T.) 314–332 (The Royal Society of Chemistry, 2018).
25. Xu, C. & Ferdosian, F. *Conversion of Lignin into Bio-Based Chemicals and Materials* 13–33 (Springer, 2017).

26. Longe, L., Garnier, G. & Saito, K. *Production of Biofuels and Chemicals from Lignin* 147–179 (Springer, 2016).
27. Mohan, M., Viswanath, P., Banerjee, T. & Goud, V. V. Multiscale modelling strategies and experimental insights for the solvation of cellulose and hemicellulose in ionic liquids. *Mol. Phys.* **116**, 2108–2128. <https://doi.org/10.1080/00268976.2018.1447152> (2018).
28. Swatloski, R. P., Spear, S. K., Holbrey, J. D. & Rogers, R. D. Dissolution of cellulose with ionic liquids. *J. Am. Chem. Soc.* **124**, 4974–4975 (2002).
29. Brandt, A., Gråsvik, J., Hallett, J. P. & Welton, T. Deconstruction of lignocellulosic biomass with ionic liquids. *Green Chem.* **15**, 550–583 (2013).
30. da Costa Lopes, A. M., João, K. G., Bogel-Lukasik, E., Roseiro, L. B. & Bogel-Lukasik, R. Pretreatment and fractionation of wheat straw using various ionic liquids. *J. Agric. Food Chem.* **61**, 7874–7882 (2013).
31. Domínguez de María, P. Recent trends in (ligno) cellulose dissolution using neoteric solvents: Switchable, distillable and bio-based ionic liquids. *J. Chem. Technol. Biotechnol.* **89**, 11–18 (2014).
32. Mohan, M., Balaji, C., Goud, V. V. & Banerjee, T. Thermodynamic insights in the separation of cellulose/hemicellulose components from lignocellulosic biomass using ionic liquids. *J. Solution Chem.* **44**, 538–557. <https://doi.org/10.1007/s10953-015-0295-3> (2015).
33. Mohan, M., Banerjee, T. & Goud, V. V. Effect of protic and aprotic solvents on the mechanism of cellulose dissolution in ionic liquids: A combined molecular dynamics and experimental insight. *ChemistrySelect* **1**, 4823–4832. <https://doi.org/10.1002/slct.201601094> (2016).
34. Hallett, J. P. & Welton, T. Room-temperature ionic liquids: Solvents for synthesis and catalysis. *Chem. Rev.* **111**, 3508–3576 (2011).
35. Pu, Y., Jiang, N. & Ragauskas, A. J. Ionic liquid as a green solvent for lignin. *J. Wood Chem. Technol.* **27**, 23–33 (2007).
36. Hou, X. D., Xu, J., Li, N. & Zong, M. H. Effect of anion structures on cholinium ionic liquids pretreatment of rice straw and the subsequent enzymatic hydrolysis. *Biotechnol. Bioeng.* **112**, 65–73. <https://doi.org/10.1002/bit.25335> (2015).
37. Hart, W. E., Harper, J. B. & Aldous, L. The effect of changing the components of an ionic liquid upon the solubility of lignin. *Green Chem.* **17**, 214–218 (2015).
38. Wang, Y. *et al.* Lignin dissolution in dialkylimidazolium-based ionic liquid–water mixtures. *Bioresour. Technol.* **170**, 499–505 (2014).
39. Padaszynski, K. & Domanska, U. Thermodynamic modeling of ionic liquid systems: Development and detailed overview of novel methodology based on the PC-SAFT. *J. Phys. Chem. B* **116**, 5002–5018 (2012).
40. Valderrama, J. & Robles, P. Critical properties, normal boiling temperatures, and acentric factors of fifty ionic liquids. *Ind. Eng. Chem. Res.* **46**, 1338–1344 (2007).
41. Eike, D. M., Brennecke, J. F. & Maginn, E. J. Predicting infinite-dilution activity coefficients of organic solutes in ionic liquids. *Ind. Eng. Chem. Res.* **43**, 1039–1048 (2004).
42. 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* **109**, 10395–10405 (2005).
43. 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* **112**, 16710–16720 (2008).
44. Klamt, A. Conductor-like screening model for real solvents: A new approach to the quantitative calculation of solvation phenomena. *J. Phys. Chem.* **99**, 2224–2235 (1995).
45. Klamt, A. & Eckert, F. COSMO-RS: A novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equilib.* **172**, 43–72 (2000).
46. Ding, W.-L. *et al.* Machine learning screening of efficient ionic liquids for targeted cleavage of the β–O–4 bond of lignin. *J. Phys. Chem. B* <https://doi.org/10.1021/acs.jpcc.1c10684> (2022).
47. Smuga-Kogut, M., Szymanowska-Powalowska, D., Markiewicz, R., Piskier, T. & Kogut, T. Ionic liquid pretreatment of stinging nettle stems and giant miscanthus for bioethanol production. *Sci. Rep.* **11**, 1–11 (2021).
48. Smuga-Kogut, M., Kogut, T., Markiewicz, R. & Słowik, A. Use of machine learning methods for predicting amount of bioethanol obtained from lignocellulosic biomass with the use of ionic liquids for pretreatment. *Energies* **14**, 243 (2021).
49. Hadji-Kali, M. K. *et al.* Screening of ionic liquids for gas separation using COSMO-RS and comparison between performances of ionic liquids and aqueous alkanolamine solutions. *Chem. Eng. Commun.* **207**, 1264–1277 (2020).
50. Khan, H. W. *et al.* Screening of ionic liquids for the extraction of biologically active compounds using emulsion liquid membrane: COSMO-RS prediction and experiments. *J. Mol. Liq.* **309**, 113122 (2020).
51. Liu, Y.-R., Thomsen, K., Nie, Y., Zhang, S.-J. & Meyer, A. S. Predictive screening of ionic liquids for dissolving cellulose and experimental verification. *Green Chem.* **18**, 6246–6254. <https://doi.org/10.1039/C6GC01827K> (2016).
52. Balaji, C., Banerjee, T. & Goud, V. V. COSMO-RS based predictions for the extraction of lignin from lignocellulosic biomass using ionic liquids: Effect of cation and anion combination. *J. Solution Chem.* **41**, 1610–1630 (2012).
53. Casas, A. *et al.* Relation between differential solubility of cellulose and lignin in ionic liquids and activity coefficients. *RSC Adv.* **3**, 3453–3460 (2013).
54. Kahlen, J., Masuch, K. & Leonhard, K. Modelling cellulose solubilities in ionic liquids using COSMO-RS. *Green Chem.* **12**, 2172–2181 (2010).
55. Casas, A. *et al.* Comparison of lignin and cellulose solubilities in ionic liquids by COSMO-RS analysis and experimental validation. *Ind. Crops Prod.* **37**, 155–163 (2012).
56. Yu, K. *et al.* Ionic liquids screening for lignin dissolution: COSMO-RS simulations and experimental characterization. *J. Mol. Liq.* **348**, 118007. <https://doi.org/10.1016/j.molliq.2021.118007> (2021).
57. Kalb, R. S. *Commercial Applications of Ionic Liquids* 261–282 (Springer, 2020).
58. Mohan, M., Banerjee, T. & Goud, V. V. Solid liquid equilibrium of cellobiose, sucrose, and maltose monohydrate in ionic liquids: Experimental and quantum chemical insights. *J. Chem. Eng. Data* **61**, 2923–2932. <https://doi.org/10.1021/acs.jced.5b00914> (2016).
59. Mohan, M., Keasling, J. D., Simmons, B. A. & Singh, S. In-silico COSMO-RS predictive screening of ionic liquids for the dissolution of plastic. *Green Chem.* **24**, 4140–4152. <https://doi.org/10.1039/d1gc03464b> (2022).
60. Hou, X. D., Smith, T. J., Li, N. & Zong, M. H. Novel renewable ionic liquids as highly effective solvents for pretreatment of rice straw biomass by selective removal of lignin. *Biotechnol. Bioeng.* **109**, 2484–2493. <https://doi.org/10.1002/bit.24522> (2012).
61. Liu, Q.-P., Hou, X.-D., Li, N. & Zong, M.-H. Ionic liquids from renewable biomaterials: Synthesis, characterization and application in the pretreatment of biomass. *Green Chem.* **14**, 304–307 (2012).
62. van Osch, D. J. *et al.* Ionic liquids and deep eutectic solvents for lignocellulosic biomass fractionation. *Phys. Chem. Chem. Phys.* **19**, 2636–2665 (2017).
63. Achinivu, E. C. *et al.* A predictive toolset for the identification of effective lignocellulosic pretreatment solvents: A case study of solvents tailored for lignin extraction. *Green Chem.* **23**, 7269–7289 (2021).
64. Mohan, M. *et al.* Multiscale molecular simulation strategies for understanding the delignification mechanism of biomass in cyrene. *ACS Sustain. Chem. Eng.* **10**, 11016–11029 (2022).
65. Ohno, H. & Fukumoto, K. Amino acid ionic liquids. *Acc. Chem. Res.* **40**, 1122–1129 (2007).

66. Mohan, M., Goud, V. V. & Banerjee, T. Solubility of glucose, xylose, fructose and galactose in ionic liquids: Experimental and theoretical studies using a continuum solvation model. *Fluid Phase Equilib.* **395**, 33–43. <https://doi.org/10.1016/j.fluid.2015.03.020> (2015).
67. Prausnitz, J. M., Lichtenthaler, R. N. & De Azevedo, E. G. *Molecular Thermodynamics of Fluid-Phase Equilibria* (Pearson Education, 1998).
68. Rocha, M. A. *et al.* Alkylimidazolium based ionic liquids: Impact of cation symmetry on their nanoscale structural organization. *J. Phys. Chem. B* **117**, 10889–10897 (2013).
69. Andanson, J.-M. *et al.* Understanding the role of co-solvents in the dissolution of cellulose in ionic liquids. *Green Chem.* **16**, 2528–2538 (2014).
70. Furche, F. *et al.* Turbomole. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4**, 91–100 (2014).
71. 722009, T. V. *University of Karlsruhe and Forschungszentrum Karlsruhe GmbH: Karlsruhe, Germany* <http://www.turbomole.com/>.
72. COSMOtherm, version C3.0 release 19.0.1 v. C3.0 release 19.0.1 (COSMOlogic GmbH & Co KG, Leverkusen, Germany, 2019).
73. Li, Y., Wang, J., Liu, X. & Zhang, S. Towards a molecular understanding of cellulose dissolution in ionic liquids: Anion/cation effect, synergistic mechanism and physicochemical aspects. *Chem. Sci.* **9**, 4027–4043 (2018).
74. Sun, J. *et al.* One-pot integrated biofuel production using low-cost biocompatible protic ionic liquids. *Green Chem.* **19**, 3152–3163 (2017).
75. Smink, D., Juan, A., Schuur, B. & Kersten, S. R. Understanding the role of choline chloride in deep eutectic solvents used for biomass delignification. *Ind. Eng. Chem. Res.* **58**, 16348–16357 (2019).
76. Wang, W. & Lee, D.-J. Lignocellulosic biomass pretreatment by deep eutectic solvents on lignin extraction and saccharification enhancement: A review. *Bioresour. Technol.* **339**, 125587 (2021).
77. Mohan, M. *et al.* Prediction of solubility parameters of lignin and ionic liquids using multi-resolution simulation approaches. *Green Chem.* **24**, 1165–1176. <https://doi.org/10.1039/d1gc03798f> (2022).
78. Diedenhofen, M. *et al.* COSMOquick, Version 1.7 (COSMOlogic GmbH & Co. KG, 2018).
79. Niederquell, A., Wyttenbach, N. & Kuentz, M. New prediction methods for solubility parameters based on molecular sigma profiles using pharmaceutical materials. *Int. J. Pharm.* **546**, 137–144 (2018).
80. Mohan, M. *et al.* Prediction of solubility parameters of lignin and ionic liquids using multi-resolution simulation approaches. *Green Chem.* **24**, 1165–1176. <https://doi.org/10.1039/d1gc03798f> (2022).
81. Doherty, T. V., Mora-Pale, M., Foley, S. E., Linhardt, R. J. & Dordick, J. S. Ionic liquid solvent properties as predictors of lignocellulose pretreatment efficacy. *Green Chem.* **12**, 1967–1975 (2010).
82. Xu, A. & Wang, F. Carboxylate ionic liquid solvent systems from 2006 to 2020: Thermal properties and application in cellulose processing. *Green Chem.* **22**, 7622–7664 (2020).
83. Liu, L., Sun, D., Li, F., Ma, S. & Tan, Z. Enantioselective liquid-liquid extraction of valine enantiomers in the aqueous two-phase system formed by the cholinium amino acid ionic liquid copper complexes and salt. *J. Mol. Liq.* **294**, 111599 (2019).
84. Aryafard, M., Jahanshahi, M., Harifi-Mood, A. R., Minofar, B. & Smatanova, I. K. Experimental and theoretical studies of preferential solvation of 4-nitroaniline and 4-nitroanisole in an amino acid ionic liquid with molecular solvents. *J. Chem. Eng. Data* **64**, 5755–5764 (2019).
85. Mishra, P. K. & Ekielski, A. The self-assembly of lignin and its application in nanoparticle synthesis: A short review. *Nanomaterials* **9**, 243 (2019).
86. Manna, B., Datta, S. & Ghosh, A. Understanding the dissolution of softwood lignin in ionic liquid and water mixed solvents. *Int. J. Biol. Macromol.* **182**, 402–412 (2021).
87. Vermaas, J. V., Dellon, L. D., Broadbelt, L. J., Beckham, G. T. & Crowley, M. F. Automated transformation of lignin topologies into atomic structures with LigninBuilder. *ACS Sustain. Chem. Eng.* **7**, 3443–3453 (2018).
88. Huda, M. M., Jahan, N. & Rai, N. Effect of water models on structure and dynamics of lignin in solution. *AIP Adv.* **11**, 065024 (2021).
89. Smith, M. D. *et al.* Cosolvent pretreatment in cellulosic biofuel production: Effect of tetrahydrofuran-water on lignin structure and dynamics. *Green Chem.* **18**, 1268–1277 (2016).
90. Hanwell, M. D. *et al.* Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **4**, 17. <https://doi.org/10.1186/1758-2946-4-17> (2012).
91. Gaussian 09 Rev. D.01 (Gaussian, Inc., Wallingford, CT, 2009).
92. Mohan, M., Naik, P. K., Banerjee, T., Goud, V. V. & Paul, S. Solubility of glucose in tetrabutylammonium bromide based deep eutectic solvents: Experimental and molecular dynamic simulations. *Fluid Phase Equilib.* **448**, 168–177. <https://doi.org/10.1016/j.fluid.2017.05.024> (2017).
93. Mohan, M., Banerjee, T. & Goud, V. V. COSMO-RS-based screening of antisolvents for the separation of sugars from ionic liquids: Experimental and molecular dynamic simulations. *ACS Omega* **3**, 7358–7370. <https://doi.org/10.1021/acsomega.8b00253> (2018).
94. Gonzalez-Miquel, M. *et al.* Excess enthalpy of monoethanolamine+ ionic liquid mixtures: How good are COSMO-RS predictions? *J. Phys. Chem. B* **118**, 11512–11522. <https://doi.org/10.1021/jp507547q> (2014).
95. Li, Y. Y. & Jin, Y. Y. Effects of thermal pretreatment on acidification phase during two-phase batch anaerobic digestion of kitchen waste. *Renew Energy* **77**, 550–557. <https://doi.org/10.1016/j.renene.2014.12.056> (2015).
96. Eckert, F. & Klamt, A. Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE J.* **48**, 369–385. <https://doi.org/10.1002/aic.690480220> (2002).
97. Kurnia, K. A., Pinho, S. O. P. & Coutinho, J. O. A. Evaluation of the conductor-like screening model for real solvents for the prediction of the water activity coefficient at infinite dilution in ionic liquids. *Ind. Eng. Chem. Res.* **53**, 12466–12475. <https://doi.org/10.1021/ie5021415> (2014).
98. Eckert, F. & Klamt, A. COSMOtherm, Version C3.0 Release 19.0.1 (COSMOlogic GmbH & Co KG, 2019).
99. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802. <https://doi.org/10.1002/jcc.20289> (2005).
100. Vermaas, J. V., Petridis, L., Ralph, J., Crowley, M. F. & Beckham, G. T. Systematic parameterization of lignin for the CHARMM force field. *Green Chem.* **21**, 109–122. <https://doi.org/10.1039/C8GC03209B> (2019).
101. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865. <https://doi.org/10.1002/jcc.20945> (2008).
102. Kim, S. *et al.* CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. *J. Comput. Chem.* **38**, 1879–1886. <https://doi.org/10.1002/jcc.24829> (2017).
103. Martínez, L., Andrade, R., Birgin, E. G. & Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **30**, 2157–2164. <https://doi.org/10.1002/jcc.21224> (2009).
104. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **103**, 4613–4621. <https://doi.org/10.1063/1.470648> (1995).
105. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101**, 4177–4189. <https://doi.org/10.1063/1.467468> (1994).
106. Allen, M. P. & Tildesley, D. J. *Computer Simulation of Liquids* 385 (Oxford University Press, 1987).

107. Deleeuw, S. W., Perram, J. W. & Smith, E. R. Simulation of electrostatic systems in periodic boundary-conditions: 3: Future theory and applications. *Proc. R. Soc. London Ser. A* **388**, 177–193 (1983).
108. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) (1996).
109. Brehm, M., Thomas, M., Gehrke, S. & Kirchner, B. TRAVIS: A free analyzer for trajectories from molecular simulation. *J. Chem. Phys.* **152**, 164105. <https://doi.org/10.1063/5.0005078> (2020).
110. Brehm, M. & Kirchner, B. TRAVIS-a free analyzer and visualizer for Monte Carlo and molecular dynamics trajectories. *J. Chem. Inf. Model.* **51**, 2007–2023. <https://doi.org/10.1021/ci200217w> (2011).

Acknowledgements

This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

Author contributions

S.S. conceived, designed and supervised the study. M.M. conceived, performed the simulations, and analyzed the data. All authors wrote and revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-25372-2>.

Correspondence and requests for materials should be addressed to B.A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023