

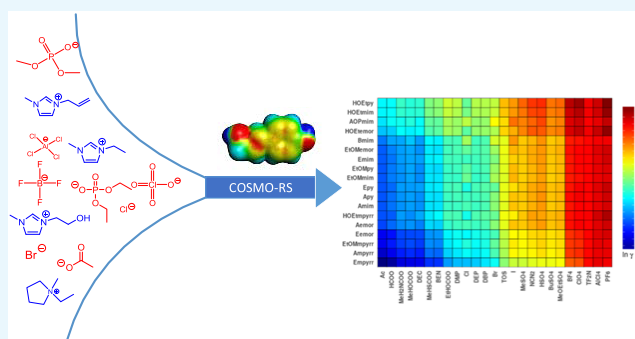
MoDooop: An Automated Computational Approach for COSMO-RS Prediction of Biopolymer Solubilities in Ionic Liquids

Yunhan Chu and Xuezhong He*[✉]

Department of Chemical Engineering, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway

Supporting Information

ABSTRACT: An automated computational framework (MoDooop) was developed to predict the biopolymer solubilities in ionic liquids (ILs) on the basis of conductor-like screening model for real solvents calculations of two thermodynamic properties: logarithmic activity coefficient ($\ln \gamma$) at infinite dilution and excess enthalpy (H^E) of mixture. The calculation was based on the optimized two-dimensional structures of biopolymer models and ILs by searching the lowest-energy conformer and optimizing molecular geometry. Three lignin models together with one IL dataset were used to evaluate the prediction ability of the developed method. The evaluation results show that $\ln \gamma$ is a more reliable property to predict lignin solubilities in ILs and the *p*-coumaryl alcohol model is considered as the best model to represent lignin molecules. The developed MoDooop approach is efficient for rapid in silico screening of suitable ionic liquids to dissolve biopolymers.



INTRODUCTION

Lignocellulosic biomass is the most abundant renewable biomaterial on the earth. It is a composite material with three main biopolymers: cellulose, hemicellulose, and lignin.¹ Cellulose and hemicellulose are typically used for production of textiles, paper, pharmaceutical compounds, etc., whereas lignin is usually converted into liquid biofuels or turned to be feedstocks of chemicals, such as binders, dispersants, surfactants, and emulsifiers.² Due to the growing concern of sustainable development and environmental protection, substantial attention has been put on the conversion of lignocellulosic biomass into biofuels or valuable products through thermochemical/biological conversion.^{3–5} A key step in the utilization of lignocellulosic biomass is to dissolve these contained biopolymers (i.e., cellulose, hemicellulose, and lignin). Among them, lignin is a cross-linked polyphenolic polymer mainly acting as a barrier preventing biological and physical attacks to cellulose and hemicellulose.¹ The crystalline structure of cellulose and the cross-linked structure of lignin make them even more difficult than hemicellulose to be deconstructed. Therefore, proper solvents should be identified to dissolve these biopolymers.

Ionic liquids (ILs) are green solvents and typically consist of a bulky, asymmetric organic cation and an anion that largely adjusts the physical and chemical properties.^{6,7} Compared to conventional solvents, ILs have desirable properties such as high thermal stability, nonvolatility, high solvation ability, and low toxicity.^{8–13} Moreover, ILs can be altered with a wide range of cations and anions to produce new ILs with a wide spectrum of physical, chemical, and biological properties.^{14,15} All of the aforementioned advantages make ILs promising solvents for

dissolution of biopolymers of lignocellulosic biomass.^{16,17} Moreover, due to the large diversity of ILs, experimental screening of ILs with preferred dissolution ability from a vast number of potential ILs to dissolve biopolymers is not practical, which highlights the importance of applying an automated rapid tool to predict their dissolving ability.

Combining statistical thermodynamics and quantum chemistry, conductor-like screening model for real solvents (COSMO-RS)^{18–21} as a well-founded approach has recently received a significant amount of attention. With the large number of segments of the molecular surfaces of the compounds, and the assumption that the segment of one molecule overlaps perfectly with that of another, the charge distribution (σ -profile) on the molecular surface and chemical potential distribution (σ -potential) of the molecule in liquid mixture are computed by COSMO-RS on the basis of quantum chemistry and statistical thermodynamics. The resulting μ turns out to be the foundation for evaluation of other equilibrium thermodynamic properties, e.g., activity coefficient (γ) and excess enthalpy (H^E). Given the ability to predict the thermodynamic data of compounds, COSMO-RS can be used as an in silico tool to screen molecules for a specific problem solely on the basis of the information arising from their molecular structures.

COSMO-RS has been proven to be effective for prediction of properties of ILs.^{22–29} It integrates dominant interactions such

Received: November 21, 2018

Accepted: January 22, 2019

Published: January 30, 2019

as electrostatic misfits, H-bonds, and van der Waals forces to summarize multiple solvation among IL systems; so, mixture calculations can be performed at different temperatures.³⁰ Compared to group contribution methods (e.g., UNIFAC model^{31–34}), COSMO-RS is a priori predictive method, which allows calculations of systems with a qualitative accuracy.³⁵ Some literature has also reported the suitability of using COSMO-RS to predict solubilities of cellulose^{36–38} and drug molecules³⁰ in ILs. On the availability of a database of quantum COSMO calculated compounds, COSMO-RS is adequate for rapid in silico screening of a large number of solutes or solvents on the basis of their selected molecular models. Moreover, the conformations of biopolymers/ILs have a high influence on the prediction results of COSMO-RS in that different predictions of thermodynamic properties can be resulted from different conformations of the same molecule.^{23,39} Therefore, it is essential to use proper molecular models and conformations searched by a stable routine to acquire qualitatively and quantitatively precise predictions.

In this work, we present an automated computational framework that allows COSMO-RS-based prediction of biopolymer solubilities in ILs (MoDooop). The computational framework is developed on the basis of a script calling of different tools: ChemAxon Convert and Cxcalc,⁴⁰ OpenBabel,⁴¹ MOPAC,⁴² and Amsterdam density functional (ADF) COSMO-RS.^{43–45} By selecting an appropriate force field and geometry optimization method, MoDooop generates a single thermodynamically stable conformer for both biopolymers and ILs. The single thermodynamically stable conformer can be used to calculate COSMO result files,⁴⁵ which permits rapid qualitative screening of ILs against selected biopolymer models on the basis of COSMO-RS.

To evaluate the developed MoDooop method, the solubilities of lignin in ILs were predicted. Lignin is represented by three different models as *p*-coumaryl, coniferyl, and sinapyl alcohol. The logarithmic activity coefficient ($\ln \gamma$) of lignin models in ILs at infinite dilution and the H^E of mixtures were calculated by COSMO-RS as qualitative measures of their solubilities in ILs. $\ln \gamma$ is correlated with differences in the strength among molecules due to the dominant interactions, which leads to the affinity between solutes and solvents.³⁹ H^E , as the temperature derivative of Gibbs free energy, is a sensitive measure of the intermolecular interactions within a mixture, which reflects the behavior of the species in solution. Linear regressions are conducted to compare the calculated $\ln \gamma$ and H^E with available experimental solubilities of lignin, and R -squared (R^2) and residual standard error (RSE) are used to measure the goodness of fit of the regression models to reflect their prediction accuracies with respect to lignin solubilities in ILs. On the basis of the evaluation of the two thermodynamic properties, the best lignin model and suitable ILs are identified.

RESULTS AND DISCUSSION

σ -Potentials of Lignin Models. The σ -potential in COSMO-RS measures the affinity between the system *S* and a surface of polarity σ . It can roughly be divided into H-bond acceptor region, the nonpolar region, and the H-bond donor region²³ for the σ -potential distribution on the molecular surface. As shown in Figure 1, the sinapyl alcohol model shows the strongest hydrogen-bonding acceptor capacity due to a more negative σ -potential in the H-bond donor region and a more positive σ -potential in the H-bond acceptor region. The *p*-coumaryl alcohol model shows the strongest hydrogen donor

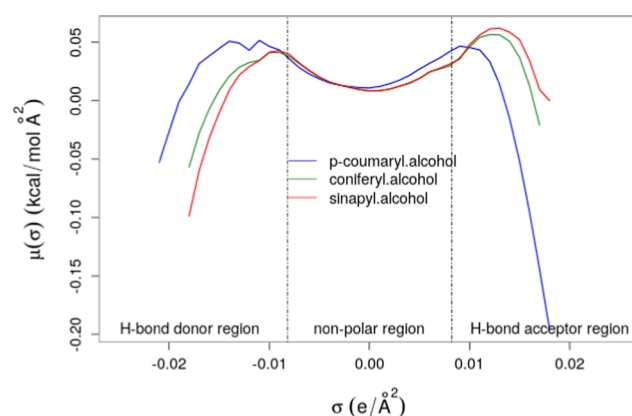


Figure 1. σ -Potentials of the three lignin models: *p*-coumaryl, coniferyl, and sinapyl alcohol predicted by COSMO-RS.

capacity due to a more negative σ -potential in the H-bond acceptor region and a more positive σ -potential in the H-bond donor region. The coniferyl alcohol model is somehow in between. Thus, the solubility ranking of the three lignin models in ILs is *p*-coumaryl alcohol > coniferyl alcohol > sinapyl alcohol, given that the IL dissolution process is anion dominated.

Model Validation. The thermodynamic properties of $\ln \gamma$ and H^E calculated by MoDooop on the basis of the proposed three lignin models along with experimental lignin solubilities in the four selected ILs from the IL dataset are listed in Table 1. The experimental lignin solubilities are compared to the calculated $\ln \gamma$ and H^E by linear regressions, and R^2 and RSE are used to characterize the goodness of fit as listed in Table 2. The lignin solubilities in the selected ILs are predicted by each regression model on the basis of calculated $\ln \gamma$ and H^E , as shown in Table 1. There are deviations between the predicted solubilities and the experimental data. However, the dissolution ability trends can be well predicted on the basis of these models, which can be used for the qualitative screening of suitable ILs for lignin dissolution.

The R^2 values of the linear regressions listed in Table 2 based on the three alcohol models show that the coniferyl alcohol model (with medium polarity) gives the best prediction with H^E ($R^2 = 0.96$), whereas the *p*-coumaryl alcohol (with more hydrogen donor capacity) gives the best prediction with $\ln \gamma$ for the lignin solubility ($R^2 = 0.91$). Nevertheless, both models present good predictions regarding the lignin solubilities in ILs as shown in Figure 2.

On the basis of $\ln \gamma$ of the *p*-coumaryl model

$$\text{solubility}^{\text{pred}} = -3.3 \ln \gamma + 1.07 \quad (1)$$

On the basis of H^E of the coniferyl model

$$\text{solubility}^{\text{pred}} = -5.25H^E - 13.64 \quad (2)$$

(See the other prediction models in Appendix A: Figures S1–S4.) It should be noted that more experimental solubilities are probably needed to further validate the robustness of the developed models in the future work.

Screening ILs. The predicted $\ln \gamma$ based on the *p*-coumaryl alcohol model and H^E based on the coniferyl alcohol model in 450 ILs are depicted in Figure 3a,b, respectively. (The detailed values are given in Appendix B: Tables S1 and S2.) The cations and anions are mapped according to scaled values of $\ln \gamma$ and H^E . The ILs with a higher dissolution capacity (highly negative values of $\ln \gamma$ and H^E) are shown in the down-left corner (blue

Table 1. Experimental Solubilities of Lignin along with $\ln \gamma$ and H^E (kJ mol^{-1}) Calculated by MoDooop at 90°C

IL	lignin solubility (wt %) ^{49,50}	$\ln \gamma$ (predicted solubility)			H^E (predicted solubility)		
		<i>p</i> -coumaryl	coniferyl	sinapyl	<i>p</i> -coumaryl	coniferyl	sinapyl
[Emim]Ac	30	-8.04 (27.60)	-3.20 (27.09)	-3.22 (26.39)	-9.94 (28.29)	-8.06 (28.68)	-7.75 (28.38)
[Bmim]Cl	10	-4.13 (14.70)	-1.45 (15.33)	-1.58 (16.12)	-6.16 (13.74)	-5.11 (13.19)	-5.09 (13.70)
[Bmim]BF ₄	4	-0.30 (2.06)	0.64 (1.29)	0.81 (1.16)	-3.24 (2.49)	-3.13 (2.79)	-3.11 (2.77)
[Bmim]PF ₆	1	0.15 (0.58)	0.64 (1.29)	0.78 (1.35)	-2.70 (0.42)	-2.67 (0.38)	-2.65 (0.23)

Table 2. Goodness of Fit (R^2 and RSE) Reflected by Linear Regressions Conducted between Experimental Lignin Solubilities and Thermodynamic Properties ($\ln \gamma$ and H^E) Calculated by MoDooop on the Basis of Three Different Lignin Models

goodness of fit	<i>p</i> -coumaryl		coniferyl		sinapyl	
	$\ln \gamma$	H^E	$\ln \gamma$	H^E	$\ln \gamma$	H^E
R^2	0.91	0.94	0.87	0.96	0.83	0.95
RSE	3.99	3.12	4.71	2.62	5.42	3.03

region), whereas those with a lower dissolution capacity (highly positive $\ln \gamma$ and H^E values) are shown in the upper-right corner (red region). Both thermodynamic properties, $\ln \gamma$ and H^E , vary significantly with anions, but are less dependent on cations, which indicates that the dissolution power is strongly dependent on anions. The ionic liquids containing the anions of Ac^- , HCOO^- , $\text{MeH}_2\text{NCOO}^-$, MeHOCOO^- , DEC^- , MeHSO_3^- , and BEN^- are found to have a high dissolution power for lignin. On the other hand, the H^E calculated on the basis of the coniferyl alcohol model shows a small difference with cations, which may indicate the challenges in distinguishing the dissolution power of ILs containing different cations. Thus, $\ln \gamma$ is regarded as a more reliable property and the *p*-coumaryl alcohol is considered as the optimal model to predict lignin solubility in ILs.

CONCLUSIONS

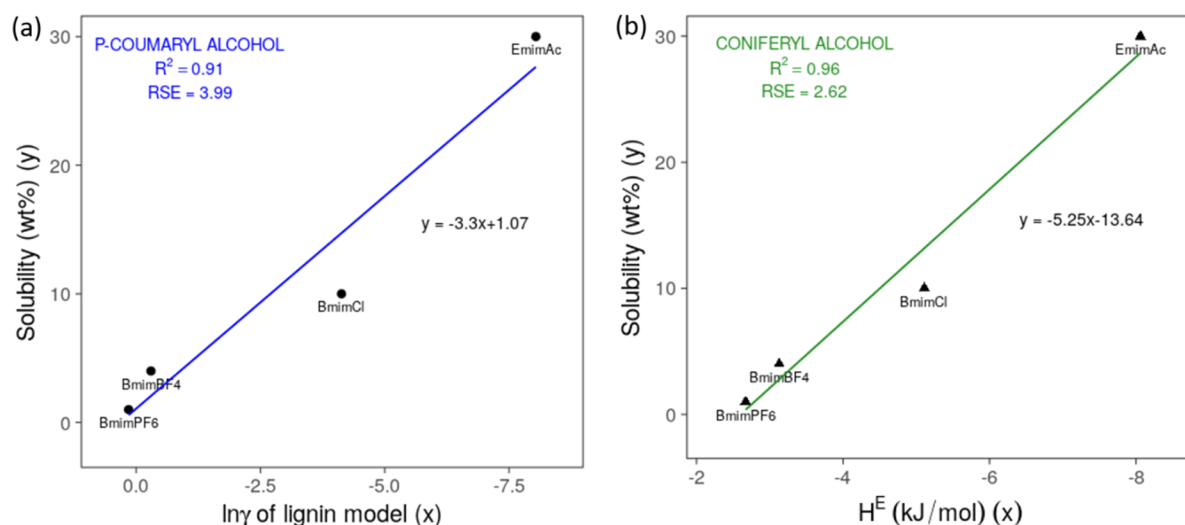
The automated computational framework of MoDooop is used for COSMO-RS-based prediction of biopolymer solubilities in ILs. To conduct the COSMO-RS calculations, the COSMO result files are generated from the two-dimensional (2D) structures of biopolymers and ILs based on the conformers

searched by specific force fields and the geometries optimized by empirical and density functional theory (DFT) methods. The method allows the use of a single thermodynamically stable conformer to represent biopolymers and ILs and thus enables rapid qualitative screening of ILs to dissolve biopolymers. Three selected lignin models have been used to predict the solubilities of lignin in 450 ILs at 90°C following the developed MoDooop method. $\ln \gamma$ is found to be a reliable reference property as it can reflect the variation of the dissolution power of ILs along with both cations and anions. The *p*-coumaryl alcohol model is selected as the best model to predict lignin solubility on the basis of $\ln \gamma$ with the high R^2 of 0.91. The ionic liquids containing the anions of Ac^- and HCOO^- show a high dissolution power for lignin. The developed MoDooop approach is efficient for the large-scale screening of suitable ILs for dissolution of lignin and potentially other biopolymers.

METHODS

Computational Framework. In the MoDooop framework, the COSMO-RS calculations of thermodynamic properties are based on ADF COSMO result files from quantum mechanical calculations of different molecular structures generated by a specific geometry optimization route. The overview of the computational workflow of MoDooop is shown in Figure 4.

- (1) Sketching the 2D structures of biopolymer models (e.g., lignin) and ILs by MarvinSketch;
- (2) converting the 2D structures of biopolymer models to 3D by Molconvert;
- (3) conducting the lowest-energy conformer search for biopolymers by Cxcalc with Dreiding force field;⁴⁶

**Figure 2.** Linear regressions of experimental solubilities of lignin measured in four ILs at 90°C against (a) $\ln \gamma$ calculated on the basis of the *p*-coumaryl alcohol model and (b) H^E of mixture calculated on the basis of the coniferyl alcohol model.

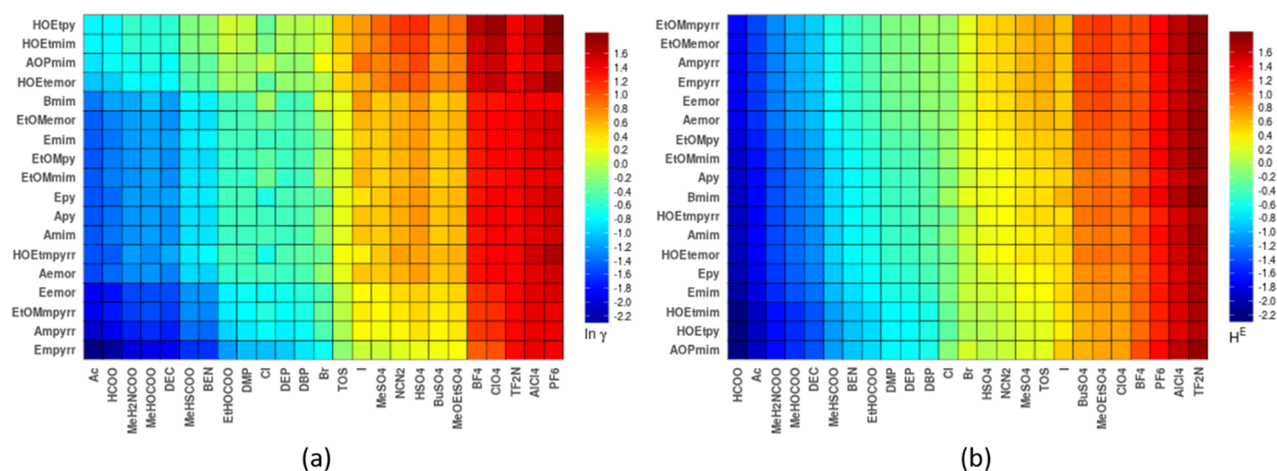


Figure 3. (a) $\ln \gamma$ of lignin in 450 ILs at infinite dilution estimated on the basis of the *p*-coumaryl alcohol model and (b) H^E of mixture calculated on the basis of the coniferyl alcohol model at 90 °C by COSMO-RS. The $\ln \gamma$ and H^E values were scaled.

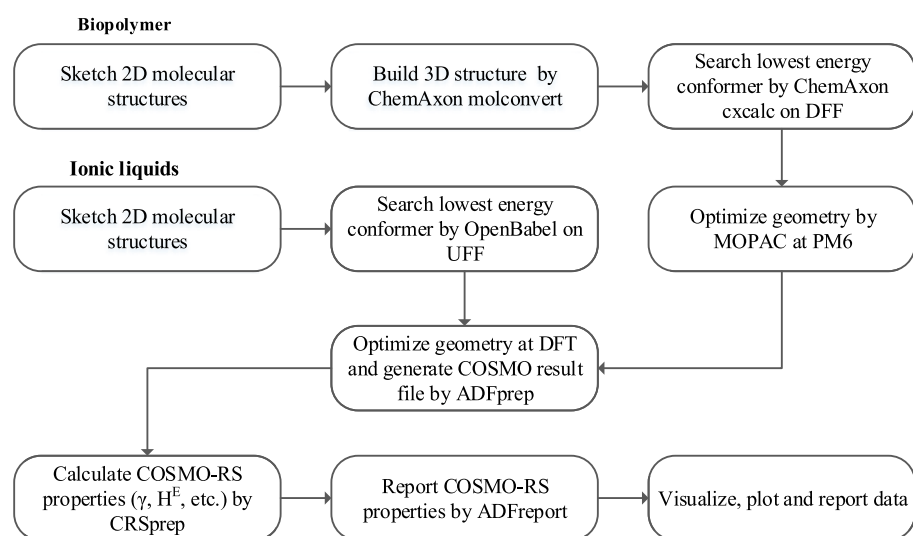


Figure 4. Schematic workflow of MoDoop.

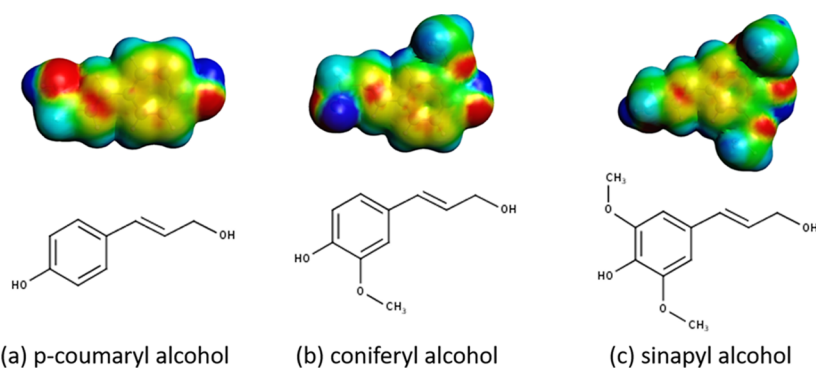


Figure 5. COSMO molecular surfaces and 2D structures of lignin models: (a) *p*-coumaryl alcohol, (b) coniferyl alcohol, and (c) sinapyl alcohol. On the molecular surface, the red area with the underlying molecular charge as negative marks positive COSMO charge density, and the blue area with the underlying molecular charge as positive marks negative COSMO charge density, whereas the yellow and green area marks nearly neutral charges.

- (4) optimizing the geometries of the obtained lowest-energy conformers of biopolymers on the basis of PM6 method by MOPAC software;
- (5) searching the lowest-energy IL conformers for isolated cations/anions by OpenBabel on the basis of the universal force field;
- (6) optimizing the resulted geometries of biopolymers and ILs at DFT level to generate COSMO result files by ADFprep⁴³ on the basis of the main parameterization GGA:BP/TZP;
- (7) calculating the σ -profiles, σ -potentials, and thermodynamic properties (e.g., γ and H^E) at a given temperature

Table 3. Cations of the IL Dataset^{38,47,48}

No.	Cation	Structure	R
1	Emim		—
2	Amim		
3	Bmim		
4	HOEtMim		
5	EtOMmim		
6	AOPmim		
7	Epy		—
8	Apy		
9	HOEtpy		
10	EtOMpy		
11	Eemor		—
12	Aemor		
13	HOEtemor		
14	EtOMemor		
15	Empyrr		—
16	Ampyrr		
17	HOEtmpyrr		
18	EtOMpyrr		

- on the basis of the generated COSMO result files by COSMO-RS implemented in CRSprep,^{43–45}
- (8) reporting the calculated COSMO-RS properties by ADFreport;
 - (9) conducting data visualization, plotting, and reporting to create the overall analysis report.

The in-house MoDooP script allows the whole computational workflow to be automated, which outputs the analysis results on the basis of given 2D structures of biopolymer and ILs. It should be noted that the step of COSMO result file generation is most time-consuming, which mainly depends on the sizes of the models; however, it only requires to be performed once per molecule, and the generated COSMO result files are reusable in the subsequent COSMO-RS computations. Moreover, in accordance with our previous calculation of $\ln \gamma$ of cellulose in ionic liquids,⁴⁷ the ILs containing halogen ions (e.g., Cl^- , Br^- , and I^-) are totally underestimated with hydrogen-bond strengths compared to those not containing halogen elements. Therefore, the values of some key parameters of the COSMO-RS model, such as the subkey CHB (chb) and SIGMAHBOND (σ_{hb}) of the key CRSPARAMETERS, are adjusted according to our previous work.⁴⁷

Lignin Models. The quantum COSMO calculation is time-consuming; thus, it is impractical to conduct computation on the whole biopolymer. A feasible way is to represent the biopolymer by a unit part, which is not only compact enough for efficient quantum mechanical calculations but also remains the main characteristic of the molecule. Lignin is usually biosynthesized

Table 4. Anions of the IL Dataset^{38,47,48}

No.	Anion	Structure	No.	Anion	Structure
1	Ac		14	I	I^-
2	AlCl_4		15	MeOEtSO_4	
3	Br	Br^-	16	MeSO_4	
4	BF_4		17	$\text{N}(\text{CN})_2$	
5	BuSO_4		18	PF_6	
6	BEN		19	TF_2N	
7	Cl	Cl^-	20	TOS	
8	ClO_4		21	DEC	
9	DMP		22	MeHOCOO	
10	DEP		23	MeHSOCOO	
11	DBP		24	MeH_2NCOO	
12	HCOO		25	EtHOCOO	
13	HSO_4				

from up to three monomers: *p*-coumaryl, coniferyl, and sinapyl alcohols. However, their compositions in lignin vary due to different material resources (e.g., softwood, hardwood, and grasses). *p*-Coumaryl alcohol is the substructure of the coniferyl alcohol, whereas coniferyl alcohol is the substructure of sinapyl alcohol. Therefore, these three alcohol structures were chosen to represent lignin molecules, and their 2D structures and COSMO molecular surfaces are shown in Figure 5a–c.

IL Dataset. A set of 450 ILs was extracted from the literature,^{38,47,48} which includes 18 cations (Table 3) of

methylimidazolium⁺, ethylmorpholinium⁺, methylpyrrolidinium⁺, and pyridinium⁺, with functional groups of allyl, ethyl, butyl, acryloyloxypropyl, 2-methoxyethyl, or 2-hydroxyethyl, and 2S anions (Table 4). The selected IL dataset was used for COSMO-RS calculations. In addition, four ILs with experimental solubilities for Kraft lignin (Indulin AT)^{49,50} at 90 °C were used to validate the prediction ability of the MoDoop approach.

■ ASSOCIATED CONTENT

● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.8b03255.

In γ based on the coniferyl alcohol model (Figure S1); In γ based on the sinapyl alcohol model (Figure S2); H^E based on the *p*-coumaryl alcohol model (Figure S3); H^E calculated on the basis of the sinapyl alcohol model (Figure S4); COSMO-RS-calculated In γ and predicted solubility of lignin in ILs based on the *p*-coumaryl model (Table S1); COSMO-RS-calculated H^E of mixture and predicted solubilities based on the coniferyl alcohol model (Table S2) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: xuezhong.he@ntnu.no. Tel: (+47) 73593942.

ORCID

Xuezhong He: 0000-0002-3232-4945

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by the Petromaks2 programme of the Research Council of Norway (CO2Hing, 267615). We gratefully thank SCM (<https://www.scm.com>) for the license of ADF Modeling Suite. We also thank ChemAxon (<https://chemaxon.com>) for the academic license of JChem package.

■ REFERENCES

- (1) Brandt, A.; Grasvik, J.; Hallett, J. P.; Welton, T. Deconstruction of lignocellulosic biomass with ionic liquids. *Green Chem.* **2013**, *15*, 550–583.
- (2) Sun, N.; Rahman, M.; Qin, Y.; Maxim, M. L.; Rodriguez, H.; Rogers, R. D. Complete dissolution and partial delignification of wood in the ionic liquid 1-ethyl-3-methylimidazolium acetate. *Green Chem.* **2009**, *11*, 646–655.
- (3) Ragauskas, A. J.; Williams, C. K.; Davison, B. H.; Britovsek, G.; Cairney, J.; Eckert, C. A.; Frederick, W. J.; Hallett, J.; Leak, D.; Liotta, C. L.; Mielenz, J. R.; Murphy, R. J.; Templer, R. H.; Tschaplinski, T. The Path Forward for Biofuels and Biomaterials. *Science* **2006**, *311*, 484–489.
- (4) Pickett, J. *Sustainable Biofuels: Prospects and Challenges*; The Royal Society, 2008.
- (5) Clark, J. H.; Deswarte, F. E. I.; Farmer, T. J. The integration of green chemistry into future biorefineries. *Biofuels, Bioprod. Biorefin.* **2009**, *3*, 72–90.
- (6) Gharagheizi, F.; Sattari, M.; Ilani-Kashkouli, P.; Mohammadi, A. H.; Ramjugernath, D.; Richon, D. Quantitative structure–property relationship for thermal decomposition temperature of ionic liquids. *Chem. Eng. Sci.* **2012**, *84*, 557–563.
- (7) Moosavi, F. The Structure of Supported Ionic Liquids at the Interface. In *Ionic Liquids - New Aspects for the Future*; Kadokawa, J.-i., Ed.; InTech: Rijeka, 2013; Chapter 8.

(8) McCormick, C. L.; Dawsey, T. R. Preparation of cellulose derivatives via ring-opening reactions with cyclic reagents in lithium chloride/*N,N*-dimethylacetamide. *Macromolecules* **1990**, *23*, 3606–3610.

(9) Fischer, S.; Voigt, W.; Fischer, K. The behaviour of cellulose in hydrated melts of the composition $\text{LiX} \cdot n\text{H}_2\text{O}$ ($\text{X} = \text{I}^-, \text{NO}_3^-, \text{CH}_3\text{COO}^-, \text{ClO}_4^-$). *Cellulose* **1999**, *6*, 213–219.

(10) Saalwächter, K.; Burchard, W.; Klüfers, P.; Kettenbach, G.; Mayer, P.; Klemm, D.; Dugarmaa, S. Cellulose Solutions in Water Containing Metal Complexes. *Macromolecules* **2000**, *33*, 4094–4107.

(11) Heinze, T.; Liebert, T. Unconventional methods in cellulose functionalization. *Prog. Polym. Sci.* **2001**, *26*, 1689–1762.

(12) Ciacco, G. T.; Liebert, T.; Frollini, E.; Heinze, T. J. Application of the solvent dimethyl sulfoxide/tetrabutyl-ammonium fluoride trihydrate as reaction medium for the homogeneous acylation of Sisal cellulose. *Cellulose* **2003**, *10*, 125–132.

(13) Fischer, S.; Leipner, H.; Thümmel, K.; Brendler, E.; Peters, J. Inorganic Molten Salts as Solvents for Cellulose. *Cellulose* **2003**, *10*, 227–236.

(14) Deetlefs, M.; Seddon, K. R.; Shara, M. Predicting physical properties of ionic liquids. *Phys. Chem. Chem. Phys.* **2006**, *8*, 642–649.

(15) Abe, M.; Kuroda, K.; Sato, D.; Kunimura, H.; Ohno, H. Effects of polarity, hydrophobicity, and density of ionic liquids on cellulose solubility. *Phys. Chem. Chem. Phys.* **2015**, *17*, 32276–32282.

(16) Swatloski, R. P.; Spear, S. K.; Holbrey, J. D.; Rogers, R. D. Dissolution of Cellulose with Ionic Liquids. *J. Am. Chem. Soc.* **2002**, *124*, 4974–4975.

(17) Wang, H.; Gurau, G.; Rogers, R. D. Ionic liquid processing of cellulose. *Chem. Soc. Rev.* **2012**, *41*, 1519–1537.

(18) 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.

(19) 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.

(20) Eckert, F.; Klamt, A. Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE J.* **2002**, *48*, 369–385.

(21) Klamt, A. *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*; Elsevier, 2005.

(22) Diedenhofen, M.; Klamt, A. COSMO-RS as a tool for property prediction of IL mixtures—A review. *Fluid Phase Equilib.* **2010**, *294*, 31–38.

(23) Gonzalez-Miquel, M.; Massel, M.; DeSilva, A.; Palomar, J.; Rodriguez, F.; Brennecke, J. F. Excess Enthalpy of Monoethanolamine + Ionic Liquid Mixtures: How Good are COSMO-RS Predictions? *J. Phys. Chem. B* **2014**, *118*, 11512–11522.

(24) Lei, Z.; Dai, C.; Chen, B. Gas Solubility in Ionic Liquids. *Chem. Rev.* **2014**, *114*, 1289–1326.

(25) Matheswaran, P.; Wilfred, C. D.; Kurnia, K. A.; Ramli, A. Overview of Activity Coefficient of Thiophene at Infinite Dilution in Ionic Liquids and their Modeling Using COSMO-RS. *Ind. Eng. Chem. Res.* **2016**, *55*, 788–797.

(26) Hernández-Bravo, R.; Miranda, A. D.; Martínez-Mora, O.; Domínguez, Z.; Martínez-Magadán, J. M.; García-Chávez, R.; Domínguez-Esquivel, J. M. Calculation of the Solubility Parameter by COSMO-RS Methods and Its Influence on Asphaltene–Ionic Liquid Interactions. *Ind. Eng. Chem. Res.* **2017**, *56*, 5107–5115.

(27) Gonfa, G.; Bustam, M. A.; Muhammad, N.; Khan, A. S. Evaluation of Thermophysical Properties of Functionalized Imidazolium Thiocyanate Based Ionic Liquids. *Ind. Eng. Chem. Res.* **2015**, *54*, 12428–12437.

(28) Muhammad, N.; Gonfa, G.; Rahim, A.; Ahmad, P.; Iqbal, F.; Sharif, F.; Khan, A. S.; Khan, F. U.; Khan, Z. U. H.; Rehman, F.; Rehman, I. U. Investigation of ionic liquids as a pretreatment solvent for extraction of collagen biopolymer from waste fish scales using COSMO-RS and experiment. *J. Mol. Liq.* **2017**, *232*, 258–264.

(29) Iqbal, J.; Muhammad, N.; Rahim, A.; Khan, A. S.; Ullah, Z.; Gonfa, G.; Ahmad, P. COSMO-RS predictions, hydrogen bond basicity

values and experimental evaluation of amino acid-based ionic liquids for lignocellulosic biomass dissolution. *J. Mol. Liq.* **2019**, *273*, 215–221.

(30) Guo, Z.; Lue, B.-M.; Thomasen, K.; Meyer, A. S.; Xu, X. Predictions of flavonoid solubility in ionic liquids by COSMO-RS: experimental verification, structural elucidation, and solvation characterization. *Green Chem.* **2007**, *9*, 1362–1373.

(31) Lei, Z.; Zhang, J.; Li, Q.; Chen, B. UNIFAC Model for Ionic Liquids. *Ind. Eng. Chem. Res.* **2009**, *48*, 2697–2704.

(32) Lei, Z.; Dai, C.; Wang, W.; Chen, B. UNIFAC model for ionic liquid-CO₂ systems. *AIChE J.* **2014**, *60*, 716–729.

(33) Lei, Z.; Dai, C.; Liu, X.; Xiao, L.; Chen, B. Extension of the UNIFAC Model for Ionic Liquids. *Ind. Eng. Chem. Res.* **2012**, *51*, 12135–12144.

(34) Han, J.; Dai, C.; Lei, Z.; Chen, B. Gas drying with ionic liquids. *AIChE J.* **2018**, *64*, 606–619.

(35) Han, J.; Dai, C.; Yu, G.; Lei, Z. Parameterization of COSMO-RS model for ionic liquids. *Green Energy Environ.* **2018**, *3*, 247–265.

(36) Kahlen, J.; Masuch, K.; Leonhard, K. Modelling cellulose solubilities in ionic liquids using COSMO-RS. *Green Chem.* **2010**, *12*, 2172–2181.

(37) Casas, A.; Omar, S.; Palomar, J.; Oliet, M.; Alonso, M. V.; Rodriguez, F. Relation between differential solubility of cellulose and lignin in ionic liquids and activity coefficients. *RSC Adv.* **2013**, *3*, 3453–3460.

(38) 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.* **2016**, *18*, 6246–6254.

(39) Padaszynski, K. An overview of the performance of the COSMO-RS approach in predicting the activity coefficients of molecular solutes in ionic liquids and derived properties at infinite dilution. *Phys. Chem. Chem. Phys.* **2017**, *19*, 11835–11850.

(40) ChemAxon, JChem 17.13. <https://chemaxon.com/>, 2017.

(41) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *J. Cheminf.* **2011**, *3*, 33.

(42) Stewart, J. J. P. MOPAC: A General Molecular Orbital Package. *Quantum Chem. Prog. Exch.* **1990**, *10*, 86.

(43) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.

(44) Pye, C. C.; Ziegler, T.; van Lenthe, E.; Louwen, J. N. An implementation of the conductor-like screening model of solvation within the Amsterdam density functional package — Part II. COSMO for real solvents. *Can. J. Chem.* **2009**, *87*, 790–797.

(45) ADF COSMO-RS Tutorial. https://www.scm.com/doc/Tutorials/COSMO-RS/Ionic_Liquids.html (accessed Feb 5, 2018).

(46) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: a generic force field for molecular simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.

(47) Chu, Y.; Zhang, X.; Hillestad, M.; He, X. Computational prediction of cellulose solubilities in ionic liquids based on COSMO-RS. *Fluid Phase Equilib.* **2018**, *475*, 25–36.

(48) Xu, A.; Wang, J.; Wang, H. Effects of anionic structure and lithium salts addition on the dissolution of cellulose in 1-butyl-3-methylimidazolium-based ionic liquid solvent systems. *Green Chem.* **2010**, *12*, 268–275.

(49) Fu, D.; Mazza, G.; Tamaki, Y. Lignin Extraction from Straw by Ionic Liquids and Enzymatic Hydrolysis of the Cellulosic Residues. *J. Agric. Food Chem.* **2010**, *58*, 2915–2922.

(50) Lee, S. H.; Doherty, T. V.; Linhardt, R. J.; Dordick, J. S. Ionic liquid-mediated selective extraction of lignin from wood leading to enhanced enzymatic cellulose hydrolysis. *Biotechnol. Bioeng.* **2009**, *102*, 1368–1376.