Solvent Screening for Separation of Lignin-Derived Molecules Using the NIST-UNIFAC Model

Amir Jalalinejad[1†], Jaber Yousefi Seyf[2‡], Axel Funke[*,1†], Nicolaus Dahmen[1†]

[†]Institute of Catalysis Research and Technology (IKFT), Karlsruhe Institute of Technology, 76344 Eggenstein-Leopoldshafen, Germany

[‡]Department of Chemical Engineering, Hamedan University of Technology, P.O. Box: 65155-579, Hamedan, Iran

Corresponding Author's E-mail: axel.funke@kit.edu, phone number: +49 721-608-22391

ABSTRACT:

Liquid-liquid phase equilibria were calculated to investigate the potential separation of value-added components from products obtained by lignin depolymerization. In this study, the ability of the group-contribution model was evaluated in the prediction of mutual solubility, and liquid-liquid phase equilibria of phenolic compounds. The phase equilibria behavior of quaternary systems was evaluated by the NIST-UNIFAC model so that the predicted results were in good agreement with the available experimental data. In sequence, the partition coefficient of 29 lignin-derived molecules with complex and polar functional groups, was predicted by the model with good accuracy (RMSE=0.7424). The ability of binary, ternary, and quaternary solvent systems was evaluated in the counter-current chromatography (CCC) separation of the products reported in the literature that are obtained through lignin depolymerization processes. Based on the empirical solvent selection criteria for CCC measurements, promising solvent systems were found for some of the lignin products. The difficult separation of some products in other cases can be attributed to the very similar chemical structure of the monomers. Finally, it was found that the NIST-UNIFAC model could qualitatively predict the solvent systems from the Arizona series; suitable for the separation of 4-hydroxybenzoic acid, vanillin, acetovanillone, syringaldehyde, acetosyringone, vanillic acid, and syringic acid.

KEYWORDS: *lignin separation, solvent screening, counter-current chromatography, liquid-liquid equilibria, NIST-UNIFAC model*

1. INTRODUCTION

Global warming and environmental concerns encourage most countries to move towards renewable energies.[1] In addition, shortages, and increasing energy consumption of fossil fuels contribute to a worldwide increasing the cost of these fuels. To solve these challenges the majority of stockholders turn to use sustainable energies.[2] Another important aspect is the substitute petroleum-based chemicals.[3] One of the valuable and renewable carbon sources to achieve this aim is lignocellulosic biomass containing cellulose, hemicellulose, and lignin.[4] Lignin, as a second major component of lignocellulose,[5] is the largest source of aromatic molecules on the planet and has the highest great potential to serve as starting material to produce bio-based products.[6, 7] The lignin products typically require upgrading to obtain a valuable product with high purity.[7, 8] To achieve high yield and effective selectivity of aromatic monomers production from lignin, there are several different catalytic techniques including depolymerizing lignin.[9, 10, 11] There are several procedures to separate the desired lignin monomers such as liquid-liquid extraction with different solvents[11], countercurrent chromatography (CCC)[12], and Centrifugal Partition Chromatography (CPC) which are effective techniques for separating the chemicals from the complex product mixture.[13-15]

Although these experimental techniques are applicable, lignin derivates are mainly aromatic and polar complex molecules with similar structures, making these techniques more expensive, time-consuming, and complicated. Therefore, predictive thermodynamic models are vital to be used in phase equilibria calculation to reduce experimental separation costs. Also, the search for cleaner, more sustainable solvents will always be an art of balancing tensions.[16]

Developing a predictive thermodynamic model with reliable results for novel molecules, in particular for systems with no experimental data, is challenging. [15, 17-19] The predictive thermodynamic models are based on the excess Gibbs free energy (G^{ex}) and equation of state (EoS). The most widely used predictive thermodynamic activity coefficient models, in particular for polar components, are based on the G^{ex}. Of the predictive activity coefficients models, group contribution-based models provide almost consistent results compared to the experimental data. The UNIFAC (UNIQUAC Functional-group Activity Coefficients) type models including UNIFAC,[20] UNIFAC-LLE,[21] UNIFAC-LBY,[22] and UNIFAC-DMD[23] have attracted researchers and practitioners in the conceptual design of the process. The UNIFAC-DMD model and its revised and extended group-group interaction parameters[24-28] are a powerful predictive model in all types of phase equilibria calculations, including VLE, LLE, SLE, etc. Despite all these advantages, most of the parameters of the UNIFAC-DMD model are available to consortium members.[29] Also, segment-based activity coefficient models such as the NRTL-SAC[30] and UNIQUAC-SAC[31] have been developed to be used for very polar and complex pharmaceuticals. However, these models require some experimental data to obtain the segment numbers of each molecule. For lignin and bio-oil-based products, predictive models such as COnductor like Screening MOdel for Real Solvents (COSMO-RS) have been used by Shen and Van Lehn.[17] In another study, the COSMO-RS model has been used for solvent selection in the biphasic isolation of Hydroxymethylfurfural (HMF).[18] Wang et al.[15] have used the COSMO-RS as a

prediction model for the extraction of Furfural/HMF. Although the COSMO-RS model has been used by some researchers,[15, 17, 18] it could not predict phase equilibria of even some simple systems. For instance, the phase splitting between water and furfural[15] and also different phase compositions for molecules with multiple conformations are in contradiction with the unique experimental results. Blumenthal et al.[32], also, have used the COSMO-RS model to find the optimal solvent for the extraction of 5-hydroxymethylfurfural. In 2023, König-Mattern et al.[19] used the COSMO-RS model in screening more than 8000 solvent candidates for the dissolving of lignocellulosic biomass. Also, the COSMO-RS model does not satisfies the thermodynamic consistency test for VLE data.[33] To our best knowledge, the UNIFAC-type models perform superior in phase equilibria calculation including VLE,[34] LLE,[35] SLE,[36] and infinite dilution activity coefficient[34, 37] compared to the COSMO-based models such as COSMO-RS and COSMO-SAC. Won Kang et al.[38] have developed the NIST-UNIFAC model and have shown that it performs superior to the most widely previously predictive used models.

Also, some measurements are challenging because of safety regulations or are particularly costly or very difficult in extreme conditions.[39] Besides the limited time, the cost of a large number of measurements is expensive for companies. In addition, due to the unavailable experimental data, correlative models such as the NRTL, and UNIQUAC are inapplicable. So, the predictive models with only a limited number of group interaction parameters can be used in these conditions and in the conceptual design of products and processes to save cost and time. Finally, when there is no reliable thermodynamic choice, the use of predictive thermodynamic models is the only possible solution for the pure component and the mixtures.

This study aims to investigate the capability of the NIST-UNIFAC model to describe liquid-liquid equilibria in comparison with the available experimental LLE, and partition coefficient data in the literature. Several relevant group contribution parameters have been optimized/reoptimized to improve the predictability of the NIST-UNIFAC model for the target compounds of this study. The resulting model was used for solvent screening in the CCC separation of lignin-derived products described in the literature.[40-43] This paper aims to demonstrate how a modified predictive model can better predict the solvent system for a liquid-liquid extraction technique to purify lignin compounds. Using counter-current chromatography (CCC) to fractionate lignin depolymerization products is quite novel; even if it is only at the solvent system modeling stage.

2. METHOD

The thermodynamic calculations are based on the NIST-UNIFAC activity coefficient model. The interaction parameters between some groups of the NIST-UNIFAC model were re-optimized or optimized (section 2.1). The chemical structure of the molecules that have experimental[44] log$P$ data and that were used in the optimization procedure is shown in Figure 1. The obtained interaction parameters in this step are used in the calculation of the involved molecules in this study. The activity coefficient of each species, including solvents and lignin-related compounds, was calculated using the NIST-modified UNIFAC activity coefficient.[38]
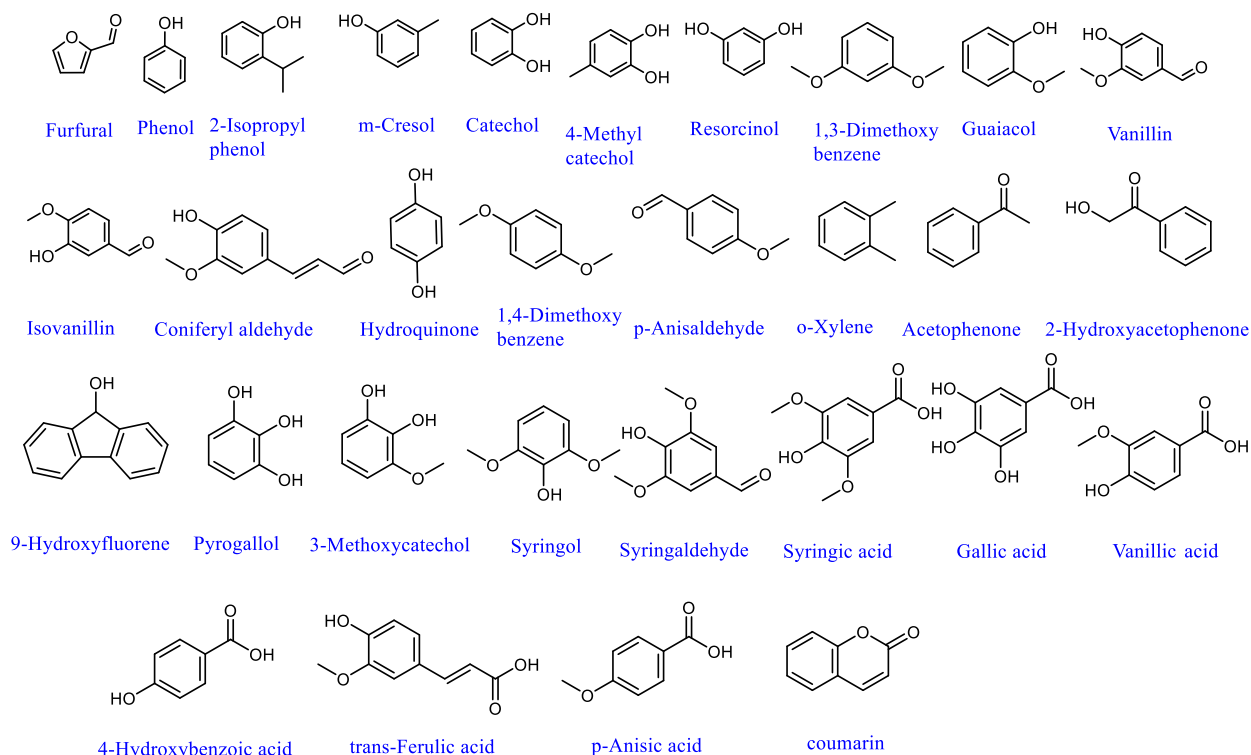
Figure 1. The chemical structure of the lignin-related compounds[44]

2.1. Revision of the NIST-UNIFAC Model

Usually, as the group number increases in the group contribution models, the number of available experimental data sets decreases. So, the accuracy of some of the obtained interaction parameters is questionable. This is true in the case of ACOH, ACCHO, and ACCOOH groups occurring in the lignin-related compounds. Also, some of the needed interaction parameters between the groups were not initially determined by Kang et al.[38] Therefore, the interaction parameters of the NIST-UNIFAC model, between CH2 (G1), ACH (G3), ACCH2 (G4), OH (G5), CCOO (G11), CH2O (G13), CCl2 (G22), c-CH2 (G42) occurring in the solvents and ACOH (G8), ACCHO (G57), ACCOOH (G58) occurring in the lignin-related compounds were revisited. The interaction parameters between $G_{1,58}$, $G_{5,58}$, $G_{11,57}$, $G_{11,58}$, $G_{11,57}$, $G_{11,58}$, $G_{11,57}$, and $G_{11,58}$ were undetermined in the original version of the NIST-UNIFAC model. So, these interaction parameters were tuned using the logP/logD experimental data.[44] The available NIST-UNIFAC model interaction parameters or zero value were used as the initial value in the optimization procedure. The *fminsearch* function of MATLAB software was used in the fitting of the interaction parameters. The details of the original and revised interaction parameters of the NIST-UNIFAC model are provided in Table S1 of Supporting Information 2.

2.2. LLE Calculations

4

The liquid-liquid equilibrium calculations in the present study are divided into four subsections (Table S1 of Supporting Information 1) including the mutual solubility of organic solvents+water at 298.15 K, the mutual solubility of organic solvents+water at different temperatures (usually up to the upper critical solution temperature (UCST)), liquid-liquid phase equilibria of ternary components mainly used in separation of lignin-derived molecules, and finally phase equilibria of quaternary solvent systems such as Arizona (heptane-ethyl acetate-methanol-water), HEMWat (hexane-ethyl acetate-methanol-water), PEMWat (pentane-ethyl acetate-methanol-water), CyEMWat (Cyclohexane-ethyl acetate-methanol-water), and IsooEMWat (Isoocatne-ethyl acetate-methanol-water).[45] The Arizona and HEMWat are standard mixtures; their composition is varied to change the polarity of both the mobile phase (mainly water+methanol), and stationary phase (mainly hexane/heptane+ethyl acetate), so that the volume of the stationary and mobile phase remains constant and equal. 16 different solvent compositions numbered sequentially from -7 to 8, and 23 different solvent compositions labeled from Z to A (except E, I, and O) in the order of increasing polarity are defined in HEMWat and Arizona series, respectively. Having a wide range of polarity, these solvents are used to separate molecules with different polarities.[46] However, because of the limited design space of these series, one can change the composition, and the types of the solvent (PEMWat, CyEMWat, IsooEMWat,[45] HIMWat, TEMWat,[17] and limonene/methanol/water as a green mixture[47]) to find optimal space to separate a mixture containing molecules with different polarities. These solvents are widely used in counter-current chromatography to separate the mixture of molecules with complex structures such as products obtained through the lignin depolymerization process.[17] These mixtures are obtained via the lignin depolymerization process.[40-43] The chemical structure of these products is shown in Figure S1 of Supporting Information 1. The NIST-UNIFAC activity coefficient model and the Rochford-Rice algorithm were used in the LLE calculations to obtain the organic and aqueous phase composition. See Supporting Information 1 (S.1.) for the calculation procedure of the partition coefficient. Although the performance of the COSMO-RS[48] model was evaluated in the liquid-liquid phase equilibria calculation of Arizona, HEMWat, PEMWat, PEMWat, and IsooEMWat systems. The average value of RMSE for the COSMO-RS model is equal to 0.0665 that is approximately twice of the NIST-UNIFAC model RMSE, equal to 0.0334. The details of the comparisons are provided in Supporting Information 1 (S.5.). If experimental data and group interaction parameters are missing, the COSMO-base models are recommended.[49]

2.3. Selection Criteria

Based on Friesen et al.,[46] there is a sweet spot in the K_D value to separate the component of a mixture using the (CCC) technique. Small values of K_D provide low peak resolution, while large values of K_D tend to produce excessive sample band broadening and extended run time. Therefore, $0.4<K_D<2.5$ ($-0.4<\log K_D<0.4$) is the working definition of the sweet spot. Of the processing variables, the type of the solvents and their composition affect the value of K_D.

Another essential factor in liquid-liquid separation is selectivity, S_{ij}; defines the extent of separation of components i and j in a mixture. S_{ij} is expressed as:

$$S_{ij} = K_D^i \big/ K_D^j \tag{1}$$

where K_D^i and K_D^j stand for the partition coefficient of component i, and j, respectively. A larger value of the selectivity and also $K_D^i > K_D^j$ is desired because of provides solutes with a large value in elution time in CCC. $S_{ij} > 1.5$ or log $S_{ij} > 0.17$ is sufficient to separate two solutes from each other.[17, 46, 50] When there are more than two solutes in the feed, first, the values of their K_D are sorted in increasing order. Then, S_{ij} is calculated for two binary solutes, providing n-1 S_{ij} for a feed containing n solute. K_D and S_{ij} values must meet the aforementioned criteria to separate the solutes from each other. Finding a solvent mixture may be difficult to meet these separation criteria. Therefore, a loose criterion (-0.9< logK_D<0.9 and logS_{ij} >0.15) is also acceptable.[17]

3. RESULTS and DISCUSSION

It must be stressed that the phase composition of the constituent solvents in each phase strongly affects the infinite dilution activity coefficient of the solute that must be separated from the mixed depolymerization product (γ_i^∞). The value of γ_i^∞ is the most important factor in the calculation of the partition coefficient (equation 4 of Supporting information 1). As upper and lower phase composition is determined by the NIST-UNIFAC model, evaluating the predictive capability of the model step by step from simple to complex systems is a crucial procedure in this study. The details and results of the mutual solubility, binodal curve, ternary LLE calculations, quaternary LLE calculations, and partition coefficient are given in Supporting Information 1 (S.2. to S.6.). The Root Mean Square Error (RMSE) of the model for the mutual solubility, binodal curve of water+furfural, ternary LLE of water+guaiacol+esters (methyl acetate, ethyl acetate, n-propyl acetate, and n-butyl acetate), ternary LLE of water+guaiacol+alcohols (1-pentanol, 1-hexanol, 1-heptanol, 1-octanol) is equal to 0.0664, 0.0064, 0.3230, 0.1548, respectively. Also, the model provides excellent results for the prediction of the mole fraction of the components in the quaternary systems. For instance, the total RMSE of the mole fractions for Arizona, HEMWat, PEMWat, CyEMWat, and IsooEMWat systems are equal to 0.0281, 0.0026, 0.0546, 0.0449, and 0.0170, respectively. These promising results indicate that the NIST-UNIFAC model can be used to predict the partition coefficient of the lignin-derive molecules in the separation of produced monomers in the lignin depolymerization process.

In another study, the NIST-UNIFAC model is evaluated for predicting the logK_D of 29 lignin-derived molecules (Figure 1) against the experimental data[44] at 298.15 K. A number of 10 binary solvent systems, including organic solvent+water has been used by Tshepelevitsh et al.[44] in the experimental procedure. Figure S6 A) and B) in supporting information 1 present the experimental and predicted logK_D of lignin-derived molecules, before, and after refitting some of

the NIST-UNIFAC interaction parameters, respectively. The results of the model with RMSE=0.7424 is better compared to RMSE=1.2942 before and after interaction parameter refitting, respectively.

Also, Figure S7 shows the predicted and experimental[51] value of $\log K_P$ for coumarin and vanillin using the HEMWat series. The results are near the diagonal line indicating the predictive power of the model with RMSE equal to 0.5148.

3.1. Solvent Screening

3.1.1. Binary Solvent Screening for Separation of the Products from Barta, Liu, Luterbacher, and Stahl

The results of the previous sections confirm the predictive ability of the NIST-UNIFAC model in describing the composition of the liquid phases in LLE calculations and also the partition coefficient of lignin-derived molecules with complex chemical structures having phenolic, carboxylic, etc. functional groups. This model has been evaluated by V. Jasperson et al.[52] in the liquid-liquid prediction of molecules related to the catalytic fast pyrolysis of biomass, including 2-ethylphenol, 3-ethylphenol, and 4-ethylphenol, 2-methoxyphenol, 3-methoxyphenol, and 4-methoxyphenol, benzofuran, and 1H-indene for the temperature range of 300-360 K. Their results were in quantitative and in some cases in qualitative agreement with the experimental data. Therefore, we are going to evaluate the NIST-UNIFAC model with some revised interaction parameters to find a binary solvent mixture than can separate the components of four sets. These sets are products from Barta[40] (P1=C2-((1,3-dioxolan-2-yl)methyl)phenol, P2=C2-((1,3-dioxolan-2-yl)methyl)-2-methoxyphenol and P3=C2-((1,3-dioxolan-2-yl)methyl)-2,5-dimethoxyphenol), Liu[41] (vanillin, p-hydroxybenzaldehyde and syringaldehyde), Luterbacher[42] (4-ethylsyringol, 4-propylsyringol and 4-propanolsyringol), and Stahl[43] (syringyl diketone, guaiacyl diketone, syringaldehyde and syringic acid) obtained from the lignin catalytic depolymerization. The binary solvent sets were selected from the list of 62 solvents approved by the Food and drug administration (FDA), that form a liquid-liquid system. A number of 1830 binary states is possible, but based on the NIST-UNIFAC model, only 160 binary mixtures meet equation 6 of Supporting Information 1 at T=298.15 K. These binary mixtures, together with the equilibrium mole fractions at T=298.15 K are provided in Supporting Information 3. These binary solvents (160 of 1830) are shown with a black square in Figure S8 of Supporting Information 1. The value of predicted $\log K_D$ versus the number of 160 binary solvent systems is shown in Figure 2. As is shown in this Figure (black rectangles), two binary solvent sets (64, and 153) can meet both separation conditions as -0.4<$\log K_D$<0.4, and $\log S_{ij}$>0.17. These sets are DMSO+diethyl ether, and sulfolane+triethyl amine. If the loose condition (orange region) is considered, the model provides 22 binary solvents (Supporting Information 3). The predicted results for the other three sets of lignin-derived products are provided in Supporting Information 3. For the products published by Liu et al,[41] the model also predicts two binary immiscible solvents that can meet the separation criteria. These are chlorobenzene+formic acid, and sulfolane+triethylamine. If the loose criteria are considered, the model suggests 21 binary solvents. Because of the very similar chemical structures of the products described by Luterbacher et al,[42] the model could not find any binary solvents capable of separating the components. However, the model suggests 7 binary solvents that meet the loose criteria. The product published by Stahl et al[43] is a four components mixture that makes it

difficult to find a binary solvent to meet both the $\log K_D$, and $\log S_{ij}$ criteria. In this case, the model found 3 binary solvents with ideal criteria. The binary solvents that provide loose separation are highlighted with blue color in Supporting Information 3. It must be stressed that these results are predicted by the model, and they require experimental validation in future work. Shen and Van Lehn[17] have tested a limited number of binary solvents using the COSMO-RS model and showed that the tested binary solvent systems are insufficient to achieve the efficient separation of mixtures of lignin-derived products. Using a systematic solvent screening approach suitable binary solvents may be available.
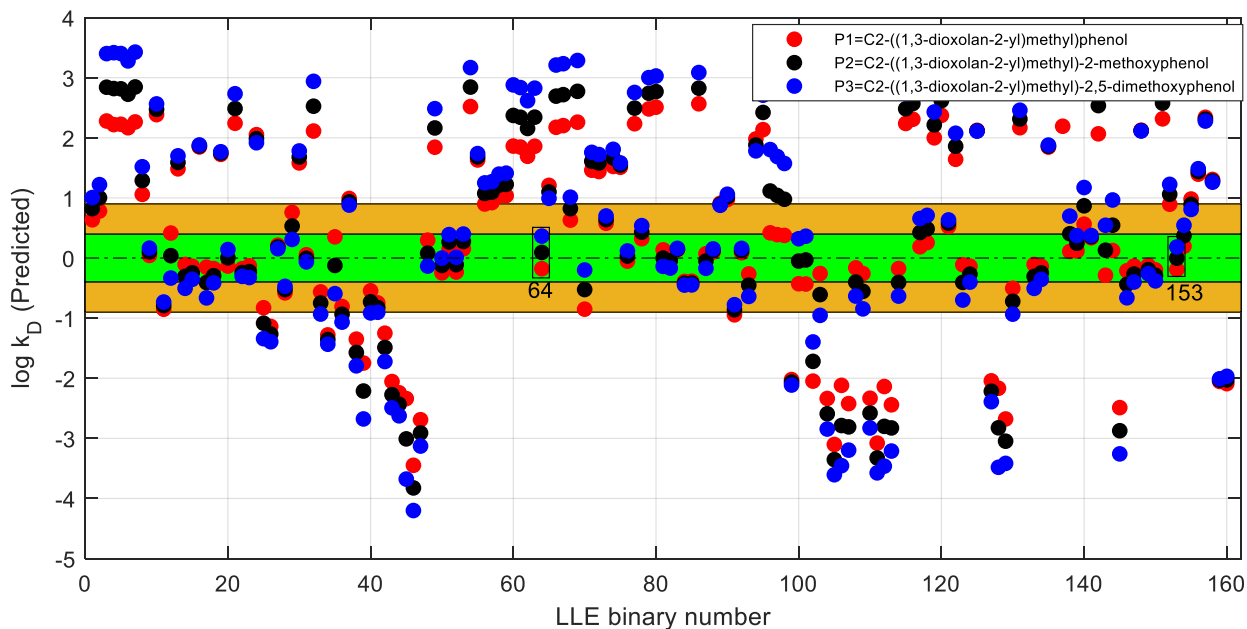

Figure 2. Predicted $\log K_D$ (partition coefficient) for the product from Barta in 160 binary solvent mixtures that form a liquid-liquid system. Possible solvent systems chosen for one-step liquid-liquid extraction are enclosed in rectangles. The green region indicates the ideal $-0.4 < \log K_D < 0.4$ criterion, and the orange region indicates the loose $0.9 < \log K_D < 0.9$ criterion.

3.1.2. Ternary Solvent Screening

One approach to finding a ternary solvent mixture is to start from the binary solvents that were predicted by the model in the previous section (64, and 153). For instance, the separation of the product from Barta using DMSO+diethyl ether, and sulfolane+triethyl amine in combination with the third solvent is considered. For a ternary solvent mixture against a binary solvent mixture, however, the composition of feed forms a degree of freedom. It is suggested to start from the binary solvent mixture and increase the mole fraction of the third solvent up to the plait point (in type I ternary systems) or another binary solvent mixture (in type II ternary systems). In other words, almost all of the ternary systems that form two liquid phases are type I or II. Figure S9 of Supporting Information 1 shows the predicted ternary system by the NIST-UNIFAC model, a) type I (DMSO+DEE+MTBE), and b) type II (DMSO+DEE+isopropyl benzene). Type I systems show the plait point where the two liquid phases become identical.

Tie lines connect two liquid phases that are in equilibrium. So, the two-phase region was automatically divided into some tie lines from binary solvent to the plait point (in type I systems), and from binary solvent to another binary solvent (in type II systems). The tie line number, hereinafter in the figures, stands for the number of tie line in the two-phase region in ternary systems. Although the partition coefficient can be calculated in the two-phase region, all feed compositions that are located on the specific tie line provide the same partition coefficient. So, one point from each tie line is sufficient. Having the tie lines, the partition coefficient is calculated. Starting from a good binary solvent and adding another third solvent gives two states. For type I systems, the selectivity (S_{ij}) reduces as the system approaches the plait point (Figure 3 a). Also, for the type II system, adding the third solvent not only causes the k_D value to exit from the green region but also decreases selectivity (Figure 3 b). Nevertheless, one can find a wide range of ternary solvent compositions that meet the good criteria of separation.
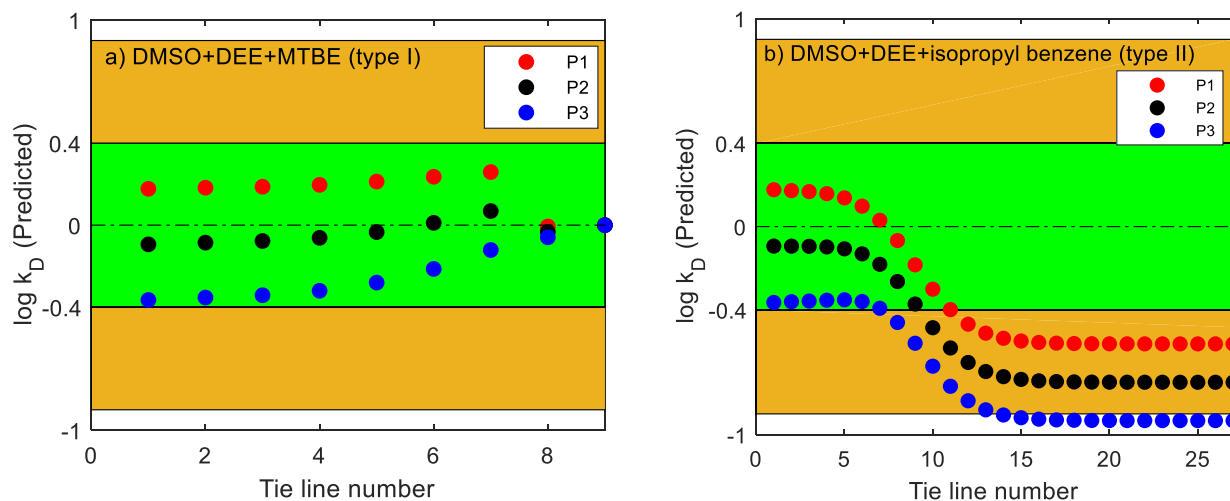

Figure 3. The predicted partition coefficient of product from Barta by the NIST-UNIFAC model, a) type I (DMSO+DEE+MTBE), and b) type II (DMSO+DEE+isopropyl benzene).

In contrast, starting from an unsuitable binary solvent and adding another third solvent to change the composition can provide a ternary mixture that meets the good separation criteria. For instance, water+ethyl acetate could not separate the product from Barta, but the model suggests that adding solvents such as cyclohexane, n-heptane, n-hexane, methylcyclohexane, and pentane (alkanes and cycloalkanes homologs) to this binary mixture can provide a sweet spot that meets the good separation criteria (Figure 4b). These ternary systems also are type II. Although it is possible to find a ternary solvent system where the partition coefficients fall in the green region for Type I ternary systems, our calculations showed that finding such cases is almost unlikely, at least for the water+ethyl acetate system (Figure 4a).
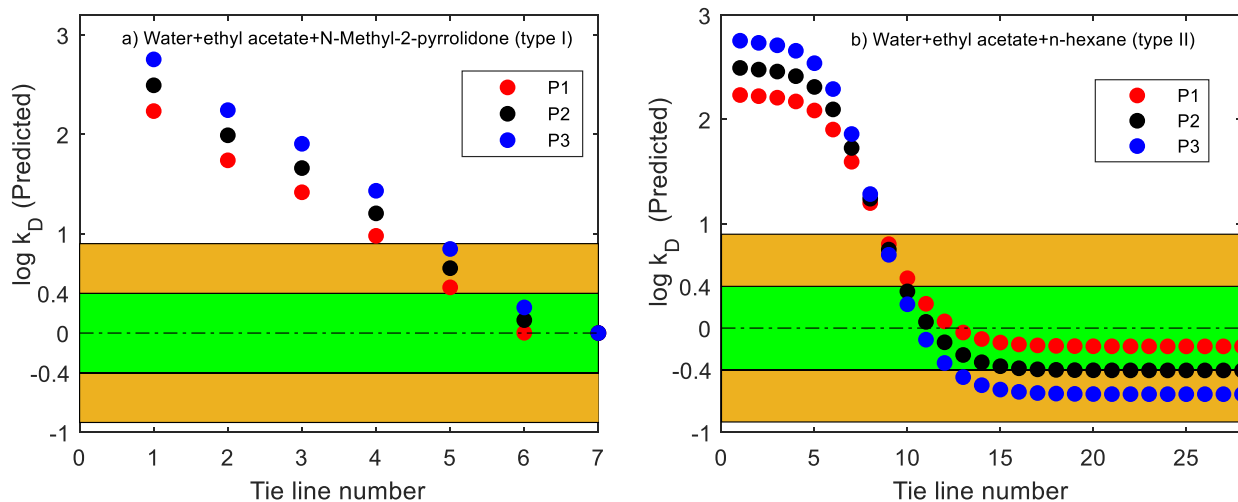
Figure 4. The predicted partition coefficient of product from Barta by the NIST-UNIFAC model, a) type I (water+ethyl acetate+N-methyl-2-pyrrolidone), and b) type II (water+ethyl acetate+n-hexane)

The data of Figure S9 of Supporting Information 1, Figure 3, and Figure 4 are given in Supporting Information 4. It seems that the number of ternary solvent systems that can separate the product from Barta is high, but because of the high computational time; here, we have tested some random sets of the ternary systems. For products published by Liu et al,[41] the model predicts two good binary solvents (Supporting Information 5). So, the same description given for the product from Barta (starting from a good binary solvent) is valid for the product from Liu. However, starting from water+ethyl acetate as an unsuitable binary solvent and adding the third solvent that forms the LLE system, the model predicts 15 ternary systems that the partition coefficient of the product from Liu falls in the green region. Among these 15 ternary systems, only 2 of them met the $\log K_D$, and $\log S_{ij}$ criteria, including water+ethyl acetate+ethylene glycol and water+ethyl acetate+ N-methyl-2-pyrrolidone (Supporting Information 5). Figure 5 presents the predicted partition coefficient of the product from Liu by the NIST-UNIFAC model, a) type I (water+ethyl acetate+ethylene glycol), and b) type II (water+ethyl acetate+N-methyl-2-pyrrolidone). The data of $\log K_D$, and $\log S_{ij}$ are provided in Supporting information 5.
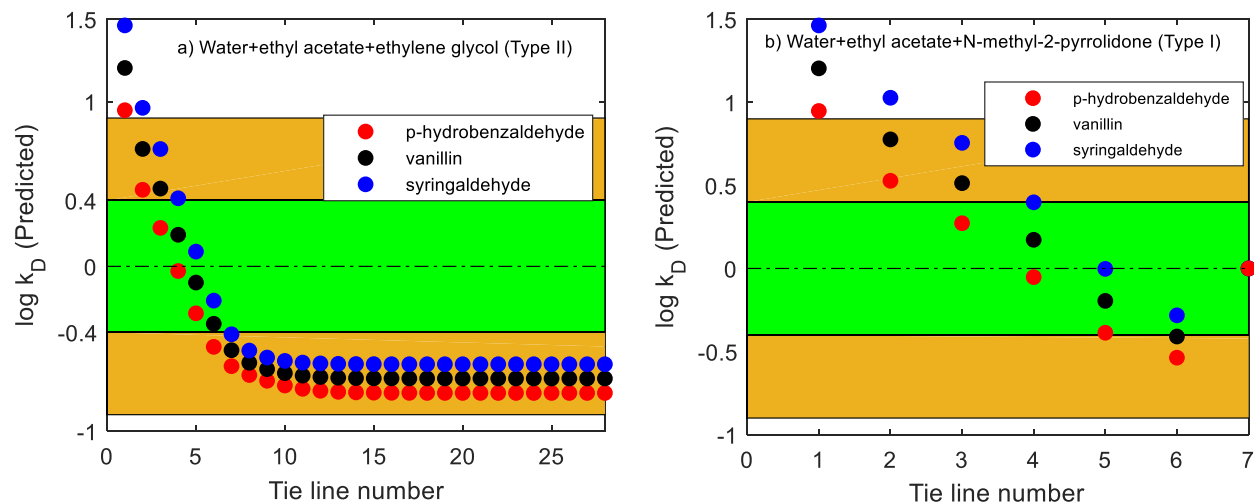
Figure 5. The predicted partition coefficient of product from Liu by the NIST-UNIFAC model, a) type II (water+ethyl acetate+ethylene glycol), and b) type I (water+ethyl acetate+N-methyl-2-pyrrolidone)

In the case of the products from Luterbacher et al,[41] similar to the binary solvent systems, the model could not find any third solvent in combination with water+ethyl acetate to fall in the green or orange melon region. This can be attributed to the very similar chemical structure of the product components. However, one may find another suitable ternary solvent using screening the fully feasible ternary solvent systems. For instance, the model predicts that the system of water+n-butanol+ethylene glycol can meet the loose criteria of separation.

Finally, for the products published by Stahl et al,[43] no third solvent was found in combination with water+ethyl acetate or water+butanol to meet $-0.4<\log K_D<0.4$, and $\log S_{ij}>0.17$ or $-0.9<\log K_D<0.9$, and $\log S_{ij}>0.15$. Using water+n-butanol+1,2,3,4-tetrahydronaphthalene, although the model predicts that the $\log K_D$ values fall in the green region, the selectivity criterion does not establish. Figure 6 shows the predicted partition coefficient of these solvent systems using the NIST-UNIFAC model. It must be emphasized that the products from Stahl have two types of compounds so that syringic acid with -COOH group differs from syringyl diketone, guaiacyl diketone, and syringaldehyde (similar compounds) with –OMe, and –OH groups. So, finding a ternary mixture that simultaneously separates the four molecules is so difficult and may be impossible.
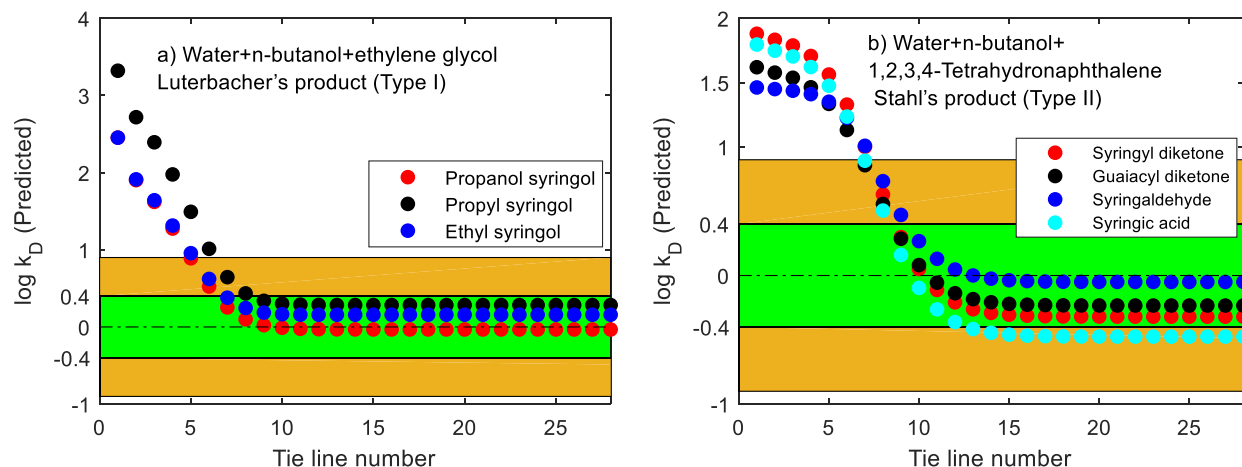
11

Figure 6. The predicted partition coefficient of a) product from Luterbacher[42] by water+n-butanol+ethylene glycol and b) product from Stahl[43] by water+n-butanol+ethylene glycol using the NIST-UNIFAC model

3.1.3. Quaternary Solvent Screening

The performance of quaternary solvent systems, including Arizona and HEMWat, was investigated to separate the components of the four products as presented in Figure 7. The values of $\log K_D$ and $\log S_{ij}$ values were calculated using the NIST-UNIFAC model. None of the Arizona series were able to fall in the ideal range for P1-3, but the R solvent system (one of the Arizona solvent compositions) satisfies both the partition coefficient and selectivity. Also, in the case of the product from Liu, only loose criteria were satisfied by the H, ad J systems in the Arizona series. However, in the case of the products from Luterbacher and Stahl similar to the binary and ternary solvents, none of the Arizona series was suitable to satisfy the ideal and loose condition.
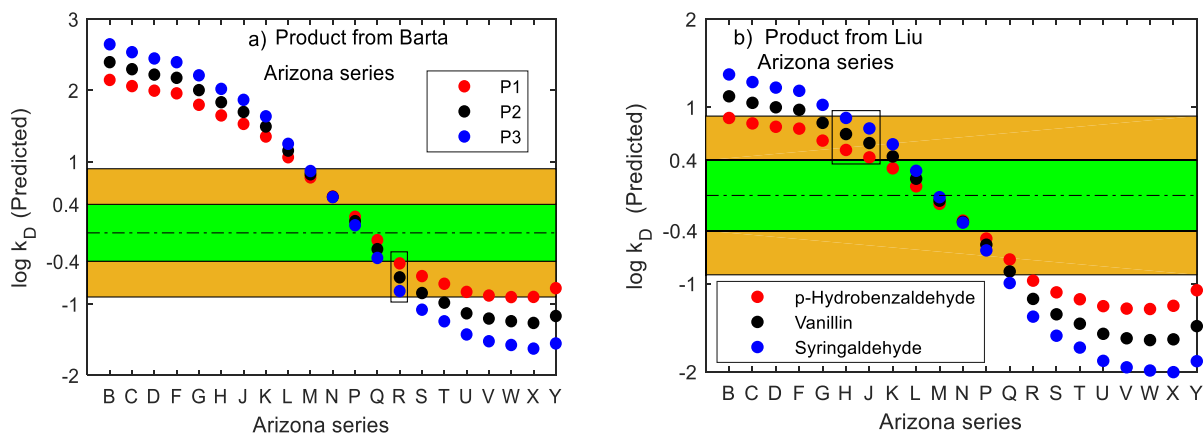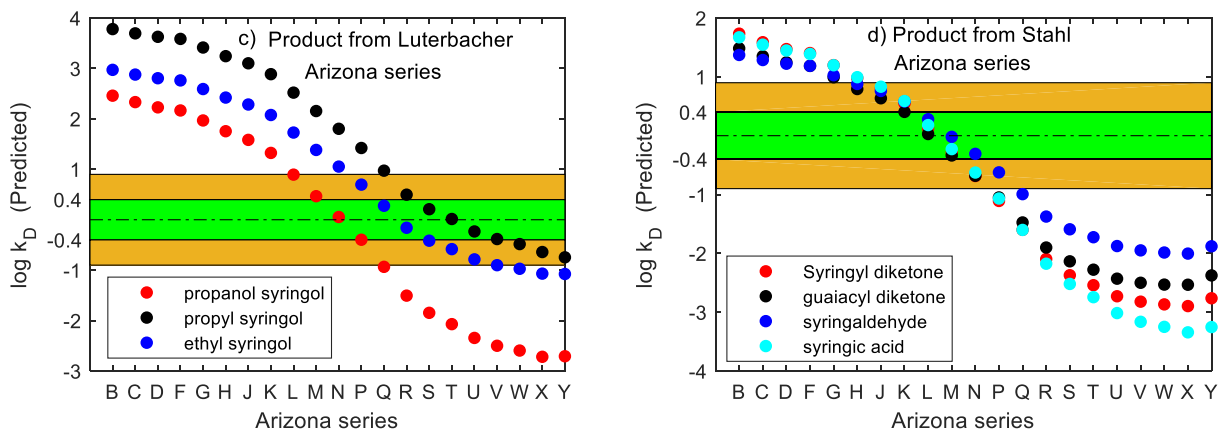
Figure 7. The predicted partition coefficient (using the NIST-UNIFAC model) of products from a) Barta[40], b) Liu[41], c) Luterbacher[42], and d) Stahl[43] in the Arizona series.

Using the HEMWat series for the product from Barta, only one composition, namely -5, was found to be suitable to loosely separate the mixture components. Also, the model predicts that +4, and +5 HEMWat systems provide loose separation of the product from Liu. Similar to the binary, and ternary solvent systems, none of the HEMWat series (-7 to +8) could separate the products of Luterbacher and Stahl, even with loose conditions.

In addition to the conventional four solvent systems, namely Arizona and HEMWat, the system of HIMWat (Hexane+Isopropyl acetate Methanol+Water) was also investigated this study. One composition of this system (-6) could separate the product from Barta monomers based on the loose criteria. Also, the product from Liu could be separated into its individual components by three predicted compositions by the NIST-UNIFAC model, including +2, +4, and +5. Similar to the previous systems the HIMWat was not useful for the products from Luterbacher and Stahl. From this series, the system of HMMWat (Hexane+Methyl isobutyl ketone+Methanol+Water) was also checked. Results showed that the products of Barta and Liu, both can be loosely separated by the -6 composition of the HMMWat series. No option was suggested for the products from Luterbacher and Stahl.

In the second step, n-Hexane was replaced by the toluene to investigate the TXWate series (where X is ethyl acetate, isopropyl acetate, and methyl isobutyl ketone).

In the case of the TEMWat series, the model predicts that the products from Barta, Luterbacher, and Stahl can not be separated products into their monomers, even inside the orange region. While the model suggests that the product from Liu can be loosely separated by the +4, and +5 solvent systems of the TEMWat series. Further, the system of TIMWat (Toluene+Isopropyl acetate+Methanol+Water) was checked by the model. No option was suggested for the product from Barta, but the model predicted that +2, and +4 satisfy only the loose criteria. Also, no composition in the TIMWat series was found to be useful for the products from Luterbacher and Stahl. To investigate the third alternative, ethyl acetate was exchanged in TEMWat by methyl isobutyl ketone (MIBK) as a ketone solvent (TMMWat=Toluene+Methyl isobutyl ketone+Methanol+Water). No composition in the TMMWat series was found for all products. It

must be emphasized that the used feed composition for HIMWat, HMMWat, TEMWat, TIMWat, and TMMWat series are the same as the HEMWat series, namely from -7 to +8.

Table 1. The Summary of Quaternary Solvent Systems, Namely Arizona, HEMWat, HIMWat, HMMWat, TEMWat, TIMWat, and TMMwate for the Separation of the Products from Barta, Liu, Luterbacher, and Stahl.

		Product			
		Barta	Liu	Luterbacher	Stahl
	Arizona	R	H, J	-	-
	HEMWat	-5	+4, +5	-	-
Solvent System series	HIMWat	-6	+2, +4, +5	-	-
	HMMWat	-6	-6	-	-
	TEMWat	-	+4, +5	-	-
	TIMWat	-	+2, +5	-	-
	TMMWat	-	-	-	-

In this section, seven quaternary series including Arizona, HEMWat, HIMWat, HMMWat, TEMWat, TIMWat, and TMMWat were checked whether they are able to separate the components of the products from Barta, Liu, Luterbacher, and Stahl by the NIST-UNIFAC model. The results are summarized in Table 1. Although the results for Barta's, and Liu's are promising, more feasible quaternary series should be checked for the products from Luterbacher, and Stahl. This procedure is time-consuming and can be computerized using in-house developed software.

3.1.4. Separation of Lignin Depolymerization Monomers

Alderich et al. [13] have obtained 4-hydroxybenzoic acid, vanillin, acetovanillone, syringaldehyde, acetosyringone, vanillic acid, and syringic acid from depolymerization of lignin in the presence of NaOH, and Cu^{2+} as the catalyst. They used the Arizona series followed by dichloromethane/methanol/water (10:6:4) to separate the vanillin, syringic acid, syringaldehyde, vanillic acid, and p-hydroxybenzoic acid from the crude mixture. Their experimental data showed that the $\log K_P$ of the vanillin, acetovanillone, and syringaldehyde, acetosyringone are almost the same. So, the separation of these components from each other is difficult or even impossible by the centrifugal partition chromatography technique. They showed that the L as a solvent system in the Arizona series (2:3:2:3 of pentane/ethyl acetate/methanol/water) can only separate vanillin and syringic acid ($\Delta \log K_P > 0.2$). In sequence, the 4-hydroxybenzoic acid, vanillic acid, and syringaldehyde have been separated from each other using a composition of halogenated solvents (CH_3Cl, CH_2Cl_2)$+CH_3OH+H_2O$.[53] In the present study, the prediction power of the NIST-UNIFAC model was tested on this mixture. The result is shown in Figure 8. Although the model could not predict the exact value of experimental $\log K_P$ values, the model qualitatively suggests that the systems of K to P in the Arizona series are suitable for the separation of these monomers. The result of the model is consistent with the J to N solvent systems tested by Alderich et al.[13] Therefore, the model can significantly reduce the cost and time of the experimental procedure.
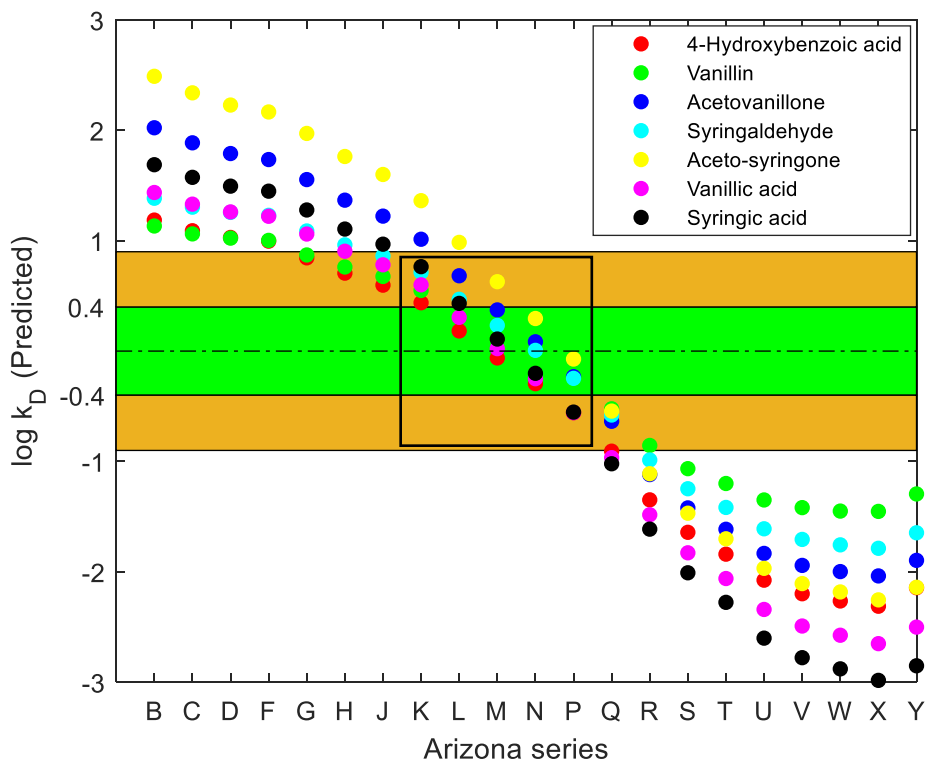
Figure 8. The predicted partition coefficient of the lignin depolymerization monomers by Arizona series and the NIST-UNIFAC model.

4. CONCLUSION

In summary, the results showed that the NIST-modified UNIFAC model could be successfully used for the calculation of liquid-liquid phase equilibria of conventional solvent systems and also complex lignin-derived molecules containing multiple functional groups. The model can be tuned to be used for solvent screening with the aim to separate value-added depolymerization monomers of lignin-derived products. Using the given procedure in the study, the feasible solvent mixture region including two, three, four, and also higher order solvents can be constrained into the tailor-made solvent systems. The partition coefficient of the lignin-derived entities not only depends on the solvent types but also the solvent compositions as another degree of freedom, considering 62 FDA-approved solvents lead to a more complex solvent-composition region. This project serves as a base for future studies and provides an important opportunity to advance the development of predictive models to be used in biomass processing. To find more reliable predicted results it is recommended that the NIST-UNIFAC model be refitted to the partition coefficient data.

■ ASSOCIATED CONTENT

Supporting Information

This information is available free of charge via the Internet at ???.

15

Supporting Information S1: Additional Figures, Tables, and descriptions (DOC).

Supporting Information S2: The original and revised interaction parameters of the NIST-UNIFAC model (XLSX).

Supporting Information S3: The predicted mole fractions of binary FDA-approved solvents, also the predicted values of $\log K_D$, and $\log S_{ij}$ in these binary solvents for the products from Barta, Liu, Luterbacher, and Stahl (XLSX).

Supporting Information S4: The LLE phase composition of DMSO+DEE+MTBE, DMSO+DEE+isopropyl benzene, water+ethyl acetate+pyrrolidone, and water+ethyl acetate+n-hexane systems, and also the $\log K_D$ and $\log S_{ij}$ for the product from Barta (XLSX).

Supporting Information S5: The LLE phase composition of water+ethyl acetate+ethylene glycol, and water+ethyl acetate+ pyrrolidone systems, and also the $\log K_D$ and $\log S_{ij}$ for the product from Liu (XLSX).

Supporting Information S6: The predicted mole fractions (mutual solubilities) by the NIST-UNIFAC model (XLSX).

Supporting Information S7: The predicted mole fractions of Arizona, HEMWat, PEMWat, CyEMWat, and IsooEMWat series by the NIST-UNIFAC model (XLSX).

Supporting Information S8: Experimental and predicted $\log K_D$ of these lignin-derived molecules

■ **AUTHOR INFORMATION**

Corresponding Authors

Axel Funke - *Institute of Catalysis Research and Technology (IKFT), Karlsruhe Institute of Technology (KIT), 76344 Eggenstein-Leopoldshafen, Germany;* orcid.org/0000- 0002-8425-1346*; Email:* axel.funke@kit.edu

Authors

Amir Jalalinejad *- Institute of Catalysis Research and Technology (IKFT), Hermann-von-Helmholtz-Platz 1, D-76344 Eggenstein-Leopoldshafen, Germany.* orcid.org/0000-0002-8188-2054

Jaber Yousefi Seyf - *Department of Chemical Engineering, Hamedan University of Technology, P.O. Box: 65155-579, Hamedan, Iran.* orcid.org/0000-0001-5624-1127

Nicolaus Dahmen - *Institute of Catalysis Research and Technology (IKFT), Hermann-von-Helmholtz-Platz 1, D-76344 Eggenstein-Leopoldshafen, Germany.*

Complete contact information is available at:

Author Contributions

The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript. Conceptualization: A. J., J. Y. S., Methodology: A. J., J. Y. S., Investigation: A. J., J. Y. S., Visualization: A. F., Project administration: A. F., Supervision: A. F., N. D., Writing-original draft: A. J., J. Y. S., Writing-review and editing: A. F., N. D.

Notes

The authors declare no competing financial interest.

(1) Figueiredo, M. B.; Hita, I.; Deuss, P. J.; Venderbosch, R.; Heeres, H. Pyrolytic lignin: a promising biorefinery feedstock for the production of fuels and valuable chemicals. *Green Chemistry* **2022**.
(2) Appunn, K. What's New in Germany's Renewable Energy Act 2021. *Clean Energy Wire* **2021**, *23*.
(3) Gunnarsdóttir, I.; Davidsdottir, B.; Worrell, E.; Sigurgeirsdóttir, S. Sustainable energy development: History of the concept and emerging themes. *Renewable and Sustainable Energy Reviews* **2021**, *141*, 110770.
(4) Clark, J. H.; Deswarte, F. *Introduction to chemicals from biomass*; John Wiley & Sons, 2014.
(5) Zhang, Y.; Naebe, M. Lignin: A review on structure, properties, and applications as a light-colored UV absorber. *ACS Sustainable Chemistry & Engineering* **2021**, *9*, 1427-1442.
(6) Huang, J.; Fu, S.; Gan, L. *Lignin chemistry and applications*; Elsevier, 2019.
(7) Rinaldi, R.; Jastrzebski, R.; Clough, M. T.; Ralph, J.; Kennema, M.; Bruijnincx, P. C.; Weckhuysen, B. M. Paving the way for lignin valorisation: recent advances in bioengineering, biorefining and catalysis. *Angewandte Chemie International Edition* **2016**, *55*, 8164-8215.
(8) Wang, H.; Pu, Y.; Ragauskas, A.; Yang, B. From lignin to valuable products–strategies, challenges, and prospects. *Bioresource technology* **2019**, *271*, 449-461.
(9) Sun, Z.; Barta, K. Cleave and couple: toward fully sustainable catalytic conversion of lignocellulose to value added building blocks and fuels. *Chemical communications* **2018**, *54*, 7725-7745.
(10) Sun, Z.; Fridrich, B.; De Santi, A.; Elangovan, S.; Barta, K. Bright side of lignin depolymerization: toward new platform chemicals. *Chemical reviews* **2018**, *118*, 614-678.
(11) Han, Y.; Paiva Pinheiro Pires, A.; Denson, M.; McDonald, A. G.; Garcia-Perez, M. Ternary Phase Diagram of Water/Bio-Oil/Organic Solvent for Bio-Oil Fractionation. *Energy & Fuels* **2020**, *34*, 16250-16264.
(12) Berthod, A.; Faure, K. Separations with a liquid stationary phase: countercurrent chromatography or centrifugal partition chromatography. *Analytical Separation Science* **2015**, 1177-1206.

(13) Alherech, M.; Omolabake, S.; Holland, C. M.; Klinger, G. E.; Hegg, E. L.; Stahl, S. S. From lignin to valuable aromatic chemicals: Lignin depolymerization and monomer separation via centrifugal partition chromatography. *ACS central science* **2021**, *7*, 1831-1837.

(14) Rocha, I. s. L.; da Costa Lopes, A. M.; Ventura, S. P.; Coutinho, J. o. A. Selective Separation of Vanillic Acid from Other Lignin-Derived Monomers Using Centrifugal Partition Chromatography: The Effect of pH. *ACS Sustainable Chemistry & Engineering* **2022**, *10*, 4913-4921.

(15) Wang, Z.; Bhattacharyya, S.; Vlachos, D. G. Extraction of furfural and furfural/5-hydroxymethylfurfural from mixed lignocellulosic biomass-derived feedstocks. *ACS Sustainable Chemistry & Engineering* **2021**, *9*, 7489-7498.

(16) Allen, D. T.; Gathergood, N.; Licence, P.; Subramaniam, B. Expectations for Manuscripts Contributing to the Field of Solvents in ACS Sustainable Chemistry & Engineering. *ACS Sustainable Chemistry & Engineering* **2020**, *8*, 14627-14629. DOI: 10.1021/acssuschemeng.0c06901.

(17) Shen, Z.; Van Lehn, R. C. Solvent selection for the separation of lignin-derived monomers using the conductor-like screening model for real solvents. *Industrial & Engineering Chemistry Research* **2020**, *59*, 7755-7764.

(18) Wang, Z.; Bhattacharyya, S.; Vlachos, D. G. Solvent selection for biphasic extraction of 5-hydroxymethylfurfural via multiscale modeling and experiments. *Green Chemistry* **2020**, *22*, 8699-8712.

(19) König-Mattern, L.; Komarova, A. O.; Ghosh, A.; Linke, S.; Rihko-Struckmann, L. K.; Luterbacher, J.; Sundmacher, K. High-throughput computational solvent screening for lignocellulosic biomass processing. *Chemical Engineering Journal* **2023**, *452*, 139476. DOI: https://doi.org/10.1016/j.cej.2022.139476.

(20) Fredenslund, A.; Jones, R. L.; Prausnitz, J. M. Group-contribution estimation of activity coefficients in nonideal liquid mixtures. *AIChE Journal* **1975**, *21*, 1086-1099.

(21) Magnussen, T.; Rasmussen, P.; Fredenslund, A. UNIFAC parameter table for prediction of liquid-liquid equilibriums. *Industrial & Engineering Chemistry Process Design and Development* **1981**, *20*, 331-339. DOI: 10.1021/i200013a024.

(22) Larsen, B. L.; Rasmussen, P.; Fredenslund, A. A modified UNIFAC group-contribution model for prediction of phase equilibria and heats of mixing. *Industrial & Engineering Chemistry Research* **1987**, *26*, 2274-2286. DOI: 10.1021/ie00071a018.

(23) Weidlich, U.; Gmehling, J. A modified UNIFAC model. 1. Prediction of VLE, hE, and .gamma..infin. *Industrial & Engineering Chemistry Research* **1987**, *26*, 1372-1381. DOI: 10.1021/ie00067a018.

(24) Gmehling, J.; Li, J.; Schiller, M. A modified UNIFAC model. 2. Present parameter matrix and results for different thermodynamic properties. *Industrial & Engineering Chemistry Research* **1993**, *32*, 178-193.

(25) Gmehling, J.; Lohmann, J.; Jakob, A.; Li, J.; Joh, R. A modified UNIFAC (Dortmund) model. 3. Revision and extension. *Industrial & engineering chemistry research* **1998**, *37*, 4876-4882.

(26) Gmehling, J.; Wittig, R.; Lohmann, J.; Joh, R. A modified UNIFAC (Dortmund) model. 4. Revision and extension. *Industrial & engineering chemistry research* **2002**, *41*, 1678-1688.

(27) Jakob, A.; Grensemann, H.; Lohmann, J.; Gmehling, J. Further development of modified UNIFAC (Dortmund): revision and extension 5. *Industrial & Engineering Chemistry Research* **2006**, *45*, 7924-7933.

(28) Constantinescu, D.; Gmehling, J. r. Further development of modified UNIFAC (Dortmund): revision and extension 6. *Journal of Chemical & Engineering Data* **2016**, *61*, 2738-2748.

(29) Seyf, J. Y.; Flasafi, S. M.; Babaei, A. H. Development of the NRTL functional activity coefficient (NRTL-FAC) model using high quality and critically evaluated phase equilibria data. 1. *Fluid Phase Equilibria* **2021**, *541*, 113088.

(30) Chen, C.-C.; Song, Y. Solubility modeling with a nonrandom two-liquid segment activity coefficient model. *Industrial & engineering chemistry research* **2004**, *43*, 8354-8362.

(31) Haghtalab, A.; Yousefi Seyf, J. Vapor–liquid and solid–liquid modeling with a universal quasichemical segment-based activity coefficient model. *Industrial & Engineering Chemistry Research* **2015**, *54*, 8611-8623.

(32) Blumenthal, L. C.; Jens, C. M.; Ulbrich, J.; Schwering, F.; Langrehr, V.; Turek, T.; Kunz, U.; Leonhard, K.; Palkovits, R. Systematic Identification of Solvents Optimal for the Extraction of 5-Hydroxymethylfurfural from Aqueous Reactive Solutions. *ACS Sustainable Chemistry & Engineering* **2016**, *4*, 228-235. DOI: 10.1021/acssuschemeng.5b01036.

(33) Lin, S.-T.; Sandler, S. I. A priori phase equilibrium prediction from a segment contribution solvation model. *Industrial & engineering chemistry research* **2002**, *41*, 899-913.

(34) Xue, Z.; Mu, T.; Gmehling, J. r. Comparison of the a priori COSMO-RS models and group contribution methods: Original UNIFAC, modified UNIFAC (Do), and modified UNIFAC (Do) consortium. *Industrial & engineering chemistry research* **2012**, *51*, 11809-11817.

(35) Birajdar, S. D.; Padmanabhan, S.; Rajagopalan, S. Rapid solvent screening using thermodynamic models for recovery of 2, 3-butanediol from fermentation by liquid–liquid extraction. *Journal of Chemical & Engineering Data* **2014**, *59*, 2456-2463.

(36) Bouillot, B.; Teychené, S.; Biscans, B. An evaluation of thermodynamic models for the prediction of drug and drug-like molecule solubility in organic solvents. *Fluid Phase Equilibria* **2011**, *309*, 36-52.

(37) Gerber, R. P.; Soares, R. d. P. Prediction of infinite-dilution activity coefficients using UNIFAC and COSMO-SAC variants. *Industrial & Engineering Chemistry Research* **2010**, *49*, 7488-7496.

(38) Kang, J. W.; Diky, V.; Frenkel, M. New modified UNIFAC parameters using critically evaluated phase equilibrium data. *Fluid Phase Equilibria* **2015**, *388*, 128-141.

(39) Kontogeorgis, G. M.; Dohrn, R.; Economou, I. G.; de Hemptinne, J.-C.; Ten Kate, A.; Kuitunen, S.; Mooijer, M.; Žilnik, L. F.; Vesovic, V. Industrial requirements for thermodynamic and transport properties: 2020. *Industrial & engineering chemistry research* **2021**, *60*, 4987-5013.

(40) Deuss, P. J.; Lancefield, C. S.; Narani, A.; De Vries, J. G.; Westwood, N. J.; Barta, K. Phenolic acetals from lignins of varying compositions via iron (iii) triflate catalysed depolymerisation. *Green Chemistry* **2017**, *19*, 2774-2782.

(41) Liu, S.; Shi, Z.; Li, L.; Yu, S.; Xie, C.; Song, Z. Process of lignin oxidation in an ionic liquid coupled with separation. *RSC advances* **2013**, *3*, 5789-5793.

(42) Shuai, L.; Amiri, M. T.; Questell-Santiago, Y. M.; Héroguel, F.; Li, Y.; Kim, H.; Meilan, R.; Chapple, C.; Ralph, J.; Luterbacher, J. S. Formaldehyde stabilization facilitates lignin monomer production during biomass depolymerization. *Science* **2016**, *354*, 329-333.

(43) Rahimi, A.; Ulbrich, A.; Coon, J. J.; Stahl, S. S. Formic-acid-induced depolymerization of oxidized lignin to aromatics. *Nature* **2014**, *515*, 249-252.

(44) Tshepelevitsh, S.; Hernits, K.; Jenčo, J.; Hawkins, J. M.; Muteki, K.; Solich, P.; Leito,   I. Systematic optimization of liquid–liquid extraction for isolation of unidentified components. *ACS omega* **2017**, *2*, 7772-7776.

(45) Berthod, A.; Hassoun, M.; Ruiz-Angel, M. Alkane effect in the Arizona liquid systems used in countercurrent chromatography. *Analytical and bioanalytical chemistry* **2005**, *383*, 327-340.

(46) Brent Friesen, J.; Pauli, G. F. GUESS—a generally useful estimate of solvent systems for CCC. *Journal of liquid chromatography & related technologies* **2005**, *28*, 2777-2806.

(47) Faure, K.; Bouju, E.; Suchet, P.; Berthod, A. Use of limonene in countercurrent chromatography: a green alkane substitute. *Analytical chemistry* **2013**, *85*, 4644-4650.

(48) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *The Journal of Physical Chemistry* **1995**, *99*, 2224-2235. DOI: 10.1021/j100007a062.

(49) Gmehling, J.; Xue, Z.; Mu, T. Reply to "Comments on 'Comparison of the a Priori COSMO-RS Models and Group Contribution Methods: Original UNIFAC, Modified UNIFAC(Do), and Modified UNIFAC(Do) Consortium'". *Industrial & Engineering Chemistry Research* **2012**, *51*, 13541-13543. DOI: 10.1021/ie3024142.

(50) Zhang, H.; Brace, E. C.; Engelberth, A. S. Selection of a nonaqueous two-phase solvent system for fractionation of xylo-oligosaccharide prebiotics using a conductor-like screening model for real solvents. *Journal of Liquid Chromatography & Related Technologies* **2016**, *39*, 666-673.

(51) Friesen, J. B.; Pauli, G. F. Performance characteristics of countercurrent separation in analysis of natural products of agricultural significance. *Journal of agricultural and food chemistry* **2008**, *56*, 19-28.

(52) Jasperson, L. V.; McDougal, R. J.; Diky, V.; Paulechka, E.; Chirico, R. D.; Kroenlein, K.; Iisa, K.; Dutta, A. Liquid–Liquid Equilibrium Measurements for Model Systems Related to Catalytic Fast Pyrolysis of Biomass. *Journal of Chemical & Engineering Data* **2017**, *62*, 243-252.

(53) Gong, C.; Chen, T.; Chen, H.; Zhang, S.; Wang, X.; Wang, W.; Sun, J.; Li, Y.; Liao, Z. First separation of four aromatic acids and two analogues with similar structures and polarities from Clematis akebioides by high-speed counter-current chromatography. *Journal of separation science* **2016**, *39*, 4660-4666.

For Table of Contents Use Only
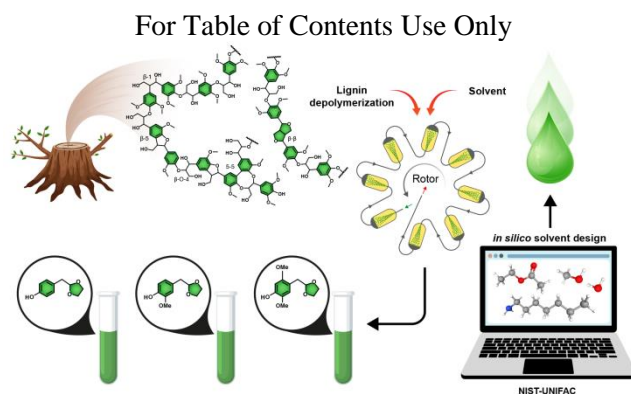
This paper deals with the separation and purification technologies as one of the twelve

principles of green chemistry