Evaluation of COSMO-SAC Method for the Prediction of the Alcohol-Water Partition Coefficients of the Compounds Encountered in Aqueous Phase Fructose Dehydration

Ruichang Xiong¹, Jacob Miller², Marta León¹, Vladimiro Nikolakis^{1*} and Stanley I. Sandler¹

¹Catalysis Center for Energy Innovation and Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE 19716, USA.

²Department of Chemical and Biological Engineering, Princeton University

* Author to whom correspondence should be addressed (vlad@udel.edu)

Abstract

Liquid-liquid extraction of fructose, 5-hydroxymethylfurfural (HMF), and HMF rehydration products, levulinic acid (LA) and formic acid (FA), from aqueous solutions to C4 to C7 linear and branched alcohols was studied. The alcohol-water partition coefficients for these compounds were measured at 30°C and 1 atm. The COSMO-SAC method was used to predict these partition coefficients and the water/alcohol and HMF/water/alcohol liquid-liquid equilibria. The predictions are in good agreement with experiments, and both show that the partition coefficients decrease with increasing the number of carbon atoms in the alcohols. Furthermore, with the same alcohol, the partition coefficient of HMF is higher than that of fructose and the HMF rehydration products. The agreement between the results of experiments and calculations indicates that the COSMO-SAC method is reasonable for the prediction of such thermodynamic data, and can be used to provide insights into the separations of these systems that is generally lacking in biomass processing. This is an important step towards screening organic solvents for liquid-liquid extraction that may preferentially extract HMF from an aqueous phase into an organic phase, thereby preventing HMF rehydration.

Key words: LLE extraction; COSMO-SAC; Phase diagram; HMF; Alcohol; Fructose dehydration; Biomass

1. Introduction

Petroleum, besides being a major transportation fuel, also serves as a source of ~95% of carbon-containing chemicals currently used in our society(Christensen et al., 2008). Concerns about its availability and the need to minimize CO₂ emissions have motivated research for the development of novel methods for the production of these same chemicals from abundant renewable sources (i.e. from biomass, etc.). Finding new ways to synthesize the intermediates for such chemicals is particularly important. One such pathway involves the Brønsted acid catalyzed dehydration of fructose to 5-hydroxymethylfurfural (HMF), which is a valuable platform chemical that can potentially be converted to a variety of fuels and chemicals(Roman-Leshkov et al., 2006; Saha and Abu-Omar, 2014; van Putten et al., 2013; Wang et al., 2014). Development of an effective procedure to remove HMF from the aqueous reaction medium during reaction can prevent unwanted side reactions that form levulinic acid (LA), formic acid (FA), and insoluble polymers called humins (see Scheme 1). Coupling fructose dehydration with liquid-liquid extraction is a frequently used approach for enhancing HMF yields since HMF is removed from the aqueous solution into an organic phase that prevents its degradation. Thus it is an important step toward improving HMF yields and creating an efficient large-scale process.

A large number of extraction solvents and conditions have been previously examined for aqueous solutions containing HMF. In the 1950s, 1-butanol was first used as an extracting solvent for HMF by Peniston(Peniston, 1956). The first solvent to be widely used to synthesize HMF was methyl isobutyl ketone; ethyl acetate is another common extraction solvent, especially in larger scale processes(van Putten et al., 2013). In an evaluation of the solvent effects on fructose dehydration to

HMF, Román-Leshkov and Dumesic(Román-Leshkov and Dumesic, 2009) found that 1-butanol and tetrahydrofuran (THF) were particularly effective at extracting HMF, especially when the aqueous layer was saturated with salt. They suggested that simultaneously dehydrating fructose and extracting HMF into an organic phase is the most effective procedure for obtaining HMF in the highest yields possible. Also, Dumestic *et al.*(Dumesic et al., 2009) developed a catalytic process for producing furan derivatives in a biphasic reactor in which the sugar is dehydrated in an aqueous phase while the furan derivative product is continuously extracted into an organic phase thus reducing side reactions. They found that one of the preferred organic solvents is 1-butanol and reported the HMF partitioning ratios in 1-butanol (and 1-hexanol) of a 30 wt% fructose solution at ~ 50% conversion. They also found that organic phase modifiers increase the extracting capability and/or selectivity towards furan derivative compounds, and that 2-butanol is the most preferred organic phase modifier. Recently Liu *et al.*(Liu et al., 2014) studied the partitioning of HMF, LA and furfural in single and mixed solvents comprised of 2-butanol, THF, dichloromethane and mixtures thereof. They investigated the effect of pH and salt concentration (up to 20 wt%) and found that the partition of all components was enhanced by the addition of salt, and except for LA was insensitive to pH.

While numerous studies have used the concept of HMF extraction for enhancing reaction yields, little has been published about its partitioning between water and organic extractants (as well as that of LA, FA or fructose compounds that are often found together in the reaction mixture (Kruger et al., 2012)). Such information is extremely important not only for improving overall HMF yields but also for the design and techno-economic evaluation of alternative HMF production processes. The ultimate goal is to find a solvent that preferentially extracts HMF from fructose, LA, and FA for use in a simultaneous fructose dehydration/HMF extraction process. An ideal solvent would isolate HMF from other molecules and stabilize HMF in a phase in which it will not degrade through rehydration and so

improve selectivity to HMF. In addition to the required high HMF partition coefficient and selectivity, the solvent compatibility with downstream processes has also to be considered.

Both experiments and predictions for screening effective extracting solvents are needed. Measurements are not only difficult but also expensive, so the accurate prediction of the thermodynamic properties is useful and can aid in the interpretation of experimental observations. Activity coefficient models such as NRTL(Renon and Prausnitz, 1968) and UNIQUAC(Abrams and Prausnitz, 1975) can be used to correlate, and UNIFAC-LLE(Fredenslund et al., 1975; Magnussen et al., 1981) can be used to predict, the multicomponent liquid-liquid equilibrium (LLE) and the partition coefficient. These models contain the interaction parameters that are usually estimated from binary experimental data. However, model parameters have not yet been established for compounds such as HMF, perhaps due to the lack of sufficient experimental data. Thus the use of a predictive activity coefficient model for screening the extracting solvents in an a priori manner is extremely attractive.

Quantum-based solvation models, such as the conductor-like screening model for real solvents (COSMO-RS) developed by Klamt(Klamt, 1995), for the prediction of fluid phase thermodynamic properties is gaining importance in the prediction of LLE since this method can be used to study any mixture without regressing the model parameters from experiment. Lin and Sandler(Lin and Sandler, 2002) reported a COSMO-SAC model (where SAC denotes segment activity coefficient), a variation of COSMO-RS method(Klamt, 1995) by invoking a necessary thermodynamic consistency criterion and making other changes. Since then, much effort has been made to improve the predictive accuracy of the COSMO-SAC model. Two recent refinements are the COSMO-SAC (2010) model(Hsieh et al., 2010) optimized using mixture phase equilibrium data and the COSMO-SAC (2013) model(Xiong et al., 2014) optimized using data for pure compound and mixtures simultaneously and by including a dispersive interaction.

In this work, we have measured the liquid-liquid partition coefficients of HMF, LA, FA and fructose between the aqueous media and a series of C4 to C7 alcohols. The COSMO-SAC method was also used to predict these partition coefficients and binary and ternary LLE, and also to interpret the experimental observations. Although the alcohols may not be the ideal extracting solvents (four carbon alcohols have been considered in the past (Dumesic et al., 2009; Liu et al., 2014; Peniston, 1956; Román-Leshkov and Dumesic, 2009)), we use them to demonstrate liquid-liquid extraction as a potential separation approach using solvents with different degrees of hydrophilicity, and the utility of the COSMO-SAC for predictions as well as for providing insights for the reasons behind the observed trends. This particular set of compounds (water and oxygenates) often coexist in a variety of biomass processing systems, in which the computational method used here can be applied.

2. Experimental

2.1. Materials

Aqueous solutions of HMF, LA, FA, and fructose were each extracted with 1-butanol, 1-pentanol, 1-hexanol, 1-pentanol, isobutanol and isopentanol. All chemicals were obtained from Sigma-Aldrich, except HMF that was obtained from Acros Organics, and all chemicals were reagent grade, and used without further purification.

2.2. Measurements

Liquid-liquid extractions were performed in a water jacketed glass cell with a 5 wt% aqueous mixture of the species to be extracted and the alcohol phase in a 1:1 volume ratio of 5 mL each. A Neslab Endocal refrigerated circulating water bath held the system at 30°C (± 1). The cells were mixed for 3 hours with a magnetic stir bar. The mixture was then allowed to equilibrate, with temperature held constant, for about 16 hours. The 16-hour period was selected in order to ensure that the two

phases have equilibrated and phase separated after the 3-hour intense mixing procedure. Pre-tests indicated that a time period of 2.5 to 5 hours is sufficient to reach equilibrium, however, a transient measurement was not taken for each sample. For scheduling purposes the 3-hour stirring period was planned to end around the end of the working day so then the samples were equilibrated overnight which corresponds to 16 hours. Finally, samples were then taken of the aqueous and alcohol layers for analysis.

These samples, except those containing both isobutanol and HMF, were then analyzed with high performance liquid chromatography (HPLC) using a Waters e2695 system with a Waters 2414 RI detector and a Biorad HPX878 column. The column temperature was kept at 50°C and the flow rate of the aqueous mobile phase that contained 0.005M H₂SO₄ was 0.5 mL/min. Samples containing both isobutanol and HMF were analyzed using an Agilent Technologies 7890 A gas chromatography system with an HP INNOWax 19091N-233 column. The column diameter was 0.025 mm with a 0.5 micron film thickness. The inlet temperature started at 40°C with a 2.586 mL/min flow and a 70:1 split ratio. Helium was used as carrier gas. The column temperature was held at 40°C for 4 minutes, then ramped to 250 °C at 10 °C/minute, finally holding at 250 °C for four minutes, resulting in a 29-minute analysis run. The flame-ionization detector used had 45 mL/min H₂ flow, a 400 mL/min air flow, and 10 mL/min He makeup flow. The column pressure was about 1.8 bar. This method gave concentrations of solutes in both phases and the concentration of the alcohol solvent in the aqueous phase. A Mettler-Toledo Karl Fischer titrator with Fluka Analytical dry methanol solution and a Fluka Analytical Hydranal Composite 5 reagent was used to determine concentrations of water in the alcohol phases.

The partition coefficient, R_c , of component i is defined as:

$$R_c = \frac{C_i^{organic}}{C_i^{aqueous}} \quad (1)$$

where $C_i^{organic}$ is the solute concentration (g/l) in the phase rich in the organic (here alcohol) solvent and $C_i^{aqueous}$ is that in the phase rich in water. To investigate the reproducibility of the measurements we ran 2 to 5 repeated tests for each solute using one of the solvents. The corresponding standard errors of the partition coefficients are the following: HMF in 1-butanol 0.011, LA in 1-butanol 0.001, FA in i-butanol 0.02, fructose in 1-butanol 0.0075. In all cases the standard error is less than 5% of the measured partitioning ratio except for fructose which as a result of having the smallest concentrations is of ~23% . For the measurements in other solvents we assume that the uncertainties are similar as those mentioned above.

2.3. COSMO-SAC calculations

The partition coefficient from the COSMO-SAC model is calculated as:

$$R_{c,\infty} = \frac{C_{II}}{C_I} \frac{x_i^{II,\infty}}{x_i^{I,\infty}} = \frac{C_{II}}{C_I} \frac{\gamma_i^{I,\infty}}{\gamma_i^{II,\infty}} \quad (2)$$

where C_I is the total molar concentration (mol/l) for the solution phase rich in water (phase I), C_{II} is that for the solution phase rich in the alcohol solvent (phase II), and $\gamma_i^{I,\infty}$ and $\gamma_i^{II,\infty}$ are the infinite-dilution activity coefficients of solute i in phases I and II, respectively. Calculations using eqn. 2 are for the solute concentration at infinite dilution in each phase. However, we demonstrate in the next section that this is valid as long as the solute is of small mole fraction, as at the experimental conditions. The partition coefficient can be also calculated from ternary LLE according to eqn. 1. In these calculations, the activity coefficients and the phase equilibria were predicted using the COSMO-SAC model. The mole fractions, used to calculate the activity coefficients, were converted, assuming no volume change on mixing, to the total concentration in each phase (eqn. 2) and the solute concentrations in each phase (eqn. 1) for calculating the partition coefficient values. The procedure for

computing activity coefficient in the COSMO-SAC model is as follows. First, the quantum mechanical COSMO computations are performed to obtain the charge density distribution to construct the sigma profile(Mullins et al., 2006). This must be done only once for each molecule. The statistical thermodynamic calculations of segment activity coefficient (SAC) then are performed using this information to determine the electrostatic free energies. Finally, by modeling the cavity formation and dispersion free energies, using the combinational term (the Staverman–Guggenheim model) and mean field theory, respectively, the activity coefficients of each species in a mixture can be calculated. Detailed information on calculating activity coefficient can be found in previous publications(Hsieh et al., 2010; Xiong et al., 2014).

3. Results and Discussion

3.1. LLE of alcohol – water systems

The LLE of binary alcohol + water systems as a function of temperature predicted using the COSMO-SAC (2010 and 2013) models(Hsieh et al., 2010; Xiong et al., 2014) are compared to the experimental data of Stephenson *et al.*(Stephenson and Stuart, 1986; Stephenson et al., 1984) in Figure 1. Both COSMO-SAC models predict the existence of LLE for the alcohol-water binary mixtures in the range of temperatures of interest. The COSMO-SAC (2013)-ADF(Xiong et al., 2014) model (i.e. COSMO-SAC (2013) model with COSMO files generated from ADF quantum software(ADF2012)) has a slightly higher predictive accuracy at low and moderate temperatures than the COSMO-SAC (2010)(Hsieh et al., 2010) model. Overall, the performance of the COSMO-SAC (2013)-ADF(Xiong et al., 2014) model is reasonably good considering that the prediction of LLE for systems with self- and cross-hydrogen bonding association is challenging since it requires an accurate description of the

subtle difference in the hydrogen-bonding interactions among all the pairs of species(Hsieh et al., 2010). Therefore, COSMO-SAC (2013)-ADF(Xiong et al., 2014) model is used hereafter in this work.

The measured concentrations of the alcohol solvents in the water-rich phase and the water concentrations in the alcohol solvent-rich phase after LLE equilibrium at 30°C in the case of binary mixtures (labelled as no solute) and for HMF, LA, FA and fructose single solute systems are shown in Figure 2. The concentrations for the binary alcohol + water system are similar to previously reported literature values(De Santis et al., 1976; Maczynski et al., 2007a, b, c, d; Stephenson and Stuart, 1986; Stephenson et al., 1984). Overall, the concentrations found for the single solute mixtures are close to those of the binary alcohol + water systems. The effect of the individual solutes on these data is not pronounced, and no simple significant effect of the added small amount of solutes on the concentration of the alcohol solvent in the aqueous layer is found. This demonstrates that the addition of a small amount of a single solute (HMF, LA, FA or fructose) does not significantly affect the binary alcohol + water system equilibrium.

3.2. Partition of solutes in the alcohol-rich phase

The measured and predicted alcohol-water partition coefficients of HMF, LA, FA and fructose at 30°C are shown in Figure 3. The measured HMF partition coefficient in 1-butanol is close to literature values(Dumesic et al., 2009; Peniston, 1956). HMF has the highest partition coefficient values in all of the alcohol solvent systems measured, followed by LA and FA, while fructose has the smallest. The larger value of the partition coefficient indicates that the solute is more soluble in the alcohol phase. Fructose shows almost no partitioning into any of the alcohol solvents, which indicates its stronger hydrophilicity. Furthermore, the partition coefficients decrease with an increasing number of carbon atoms in the alcohol. Finally, the partitioning of each solute in the C4 and C5 branched

alcohols is very similar to that in the corresponding normal alcohols indicating that the form of the alkyl chain (linear vs. branched) has a smaller effect than the total number of carbon atoms. The values of the HMF extraction selectivity for each solvent are presented in Table 1. The selectivities of HMF over fructose are high for all the alcohols tested indicating that HMF will be selectively extracted from an aqueous solution that contains both HMF and fructose. This finding is important for the selective extraction of HMF during fructose dehydration reactions. This indicates that the undesired fructose losses to the extracting phase will be small. However, HMF selectivities over the organic acids (FA and LA) are small indicating that once the organic acids are formed it will be difficult to separate them from HMF by the process considered here.

The predictions from the COSMO-SAC (2013) model(Xiong et al., 2014) are in reasonably good agreement with the measured data. In particular, the quantitative partition coefficient comparisons for HMF and fructose in all these alcohol solvents are quite good, while the predictions for the LA and FA partition coefficients have slightly larger errors. In general, the least accurate predictions of the current COSMO-SAC model among different classes of compounds are for the amines, ketones and acids(Xiong et al., 2014), which could be due to the strong hydrogen-bonding associations of such compounds and/or the possibility of their partial dissociation(Kim et al., 1996), which is not considered in the model. Overall, considering that the model parameters were not optimized specifically for partitioning in biphasic systems, and no adjustment has been made here to any parameters of the model, these complete predictions for these systems can be considered reasonable.

The decrease in the partition coefficients with increasing the number of carbon atoms of the alcohols is interesting and can be analyzed using the COSMO-SAC model. The HMF and fructose activity coefficients calculated for the water-rich and alcohol-rich phases and the resulting partition

coefficients are summarized in Table 2 (A detailed table of all the experimental and computational data is provided as supporting information). The alcohol used did not significantly affect the activity coefficients in the water-rich phase as expected since the mole fraction of water is almost 100% in the aqueous phase for all alcohols examined here. So the difference in partition coefficients arises primarily from the activity coefficients in the alcohol-rich phase. The activity coefficients were calculated with the model as the combination of the electrostatic, dispersive, and cavity terms. These values (for the alcohol-rich phase) are summarized in Table 3. It is clear that in all cases the electrostatic interaction is the dominant contribution. The cavity effect becomes less important as the solvent is changed from 1-butanol to 1-heptanol. Also the dispersion becomes more important as the solute is changed from HMF to fructose. One descriptor for the electrostatic contribution is the charge density distribution profile (σ -profile), which can be separated into two components: non-hydrogen-bonding (non-HB) density and the hydrogen-bonding (HB) density by weighting the probability of the HB segment forming a hydrogen bond(Xiong et al., 2014). The non-HB and HB profiles for the different alcohol molecules are shown in Figure 4. By comparing the σ -profiles, we conclude that the decrease of partitioning in the alcohol phase with increasing carbon number is due to the increase in the non-HB density for these solvents rather than as a result of the increase of HB density, which is almost unchanged.

The COSMO-SAC predicted ternary LLE phase diagrams at 30°C for the HMF-alcohol-water systems are shown in Figure 5. The two-phase regions are under the curves in Figure 5(a). The dashed lines are the tie lines connecting equilibrium concentrations in the two phases. This diagram shows that as the HMF solute concentration increases, the LLE two-phase region shrinks for all the alcohols. Also the two-phase region is smaller for the alcohols with smaller carbon number and disappears at lower HMF concentrations. This means that while HMF partition coefficient between 1-butanol and water is

higher than that between other alcohols and water, the presence of significant HMF quantities will make butanol-water LLE easier to disappear, though this may be unlikely to occur under the most commonly-used experimental conditions. For example, even in the case of 30 wt% initial fructose concentration, which is in the range of the highest initial concentration used[9], the mole fraction of HMF will be lower than ~4 mol% (~20 wt% if it is in pure water) even if fructose is completely dehydrated and converted to HMF ($1 \text{ Fructose} \rightarrow 1 \text{ HMF} + 3 \text{ H}_2\text{O}$). The alcohol-water partition coefficients of HMF as a function of the length of the tie lines are shown in Figure 5 (b). [The slope of tie-line determining the partition coefficients is similar for each of the alcohol solvents studied here, therefore the length of the tie lines (i.e. the size of the miscibility gap) is examined.] The longest tie line for each binary alcohol-water system occurs at infinite dilution of HMF. When tie line length is zero, only a single phase system is present and no partitioning occurs. Figure 5 (b) indicates that the most effective value (either the largest value above one or smallest value below one) of the partition coefficient of HMF occurs at the tie lines of longest length (i.e. at small fraction of HMF), which is equivalent to previous results shown in Figure 3, calculated using infinite dilution of HMF in an equilibrated binary alcohol and water system.

Methyl isobutyl ketone (MIBK) is another extractant frequently considered in experimental studies. Motivated by its frequent use we decided to examine how well COSMO-SAC can predict the literature data for this system despite the fact that MIBK is not an alcohol. The COSMO-SAC predictions for binary LLE of water and the MIBK system are compared with experimental data(Stephenson, 1992) in the supporting information (Figure S1). The two phase region at different temperatures is also predicted for MIBK but the error is larger than that in alcohol system. This is because accuracy of the current COSMO-SAC model among different classes of compounds is different (as pointed out in previous publications for COSMO-based methods (Hsieh et al., 2010;

Mullins et al., 2006; Xiong et al., 2014)). However, the predicted HMF partition ratio (~ 0.8) in water and MIBK system using COSMO-SAC is close to ~ 1.1 (the experimental determined value using reaction mixtures after fructose conversion to HMF at $\sim 50\%$ fructose conversion (Dumesic et al., 2009; Roman-Leshkov et al., 2006)). This finding is promising and encouraging, especially after taking into consideration the uncertainty in the composition of the liquid mixtures used in the experiments (concentrations of the other components is not known). Furthermore, COSMO-SAC predictions capture the experimental trends since the partitioning ratio is lower than that in 1-butanol highlighting its usefulness for extractant screening.

4. Conclusions

The partition coefficients of HMF between water and other solvent systems predicted using the computational method COSMO-SAC were in good agreement with experimentally determined values, indicating that this method can be useful in screening potential HMF extractants. In particular we measured the partition coefficients of HMF, levulinic acid, formic acid and fructose in a series of alcohol-water biphasic systems. This set of compounds is involved in a variety of biomass processes, e.g. the dehydration reaction of fructose to HMF. The COSMO-SAC method was used to predict the partition coefficients of these compounds and the liquid-liquid equilibrium phase diagrams. In agreement with experiment, C4 alcohols partitioning HMF into the alcohol phase with partition coefficients higher than the longer chain alcohols. Partition coefficients for HMF are higher than those for the other three solutes at every studied experimental condition. The partition coefficients in alcohol solvent are probably too low for efficient industrial extraction. Using a predictive activity coefficient model such as COSMO-SAC to predict HMF partition ratio and finding optimal solvents to extract HMF is interesting, and will be considered as part of our future work.

Acknowledgements

This material is based upon work supported by the Catalysis Center for Energy Innovation, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001004.

References

- Abrams, D.S., Prausnitz, J.M., 1975. Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems. *Aiche Journal* 21, 116-128.
- ADF2012, SCM Theoretical Chemistry, Amsterdam, The Netherlands.
- Christensen, C.H., Rass-Hansen, J., Marsden, C.C., Taarning, E., Egeblad, K., 2008. The renewable chemicals industry. *Chemsuschem* 1, 283-289.
- De Santis, R., Marrelli, L., Muscetta, P.N., 1976. Influence of temperature on the liquid-liquid equilibrium of the water-n-butyl alcohol-sodium chloride system. *Journal of Chemical & Engineering Data* 21, 324-327.
- Dumesic, J.A., Rométn-Leshkov, Y., Chheda, J.N., 2009. Catalytic Process for Producing Furan Derivatives in A Biphasic Reactor, U.S. Patent 7572925B2.
- Fredenslund, A., Jones, R.L., Prausnitz, J.M., 1975. Group-contribution estimation of activity coefficients in nonideal liquid mixtures. *Aiche Journal* 21, 1086-1099.
- Hsieh, C.M., Sandler, S.I., Lin, S.T., 2010. Improvements of COSMO-SAC for vapor-liquid and liquid-liquid equilibrium predictions. *Fluid Phase Equilibria* 297, 90-97.
- Kim, M.H., Kim, C.S., Lee, H.W., Kim, K., 1996. Temperature dependence of dissociation constants for formic acid and 2,6-dinitrophenol in aqueous solutions up to 175 degrees C. *Journal Of The Chemical Society-Faraday Transactions* 92, 4951-4956.
- Klamt, A., 1995. Conductor-like Screening Model for Real Solvents - A New Approach to the Quantitative Calculation of Solvation Phenomena. *Journal Of Physical Chemistry* 99, 2224-2235.
- Kruger, J.S., Nikolakis, V., Vlachos, D.G., 2012. Carbohydrate dehydration using porous catalysts. *Current Opinion in Chemical Engineering* 1, 312-320.
- Lin, S.T., Sandler, S.I., 2002. A priori phase equilibrium prediction from a segment contribution solvation model. *Industrial & Engineering Chemistry Research* 41, 899-913.
- Liu, F., Sivioththaman, S., Tan, Z., 2014. Solvent extraction of 5-HMF from simulated hydrothermal conversion product. *Sustain. Environ. Res.* 24, 149-157.
- Maczynski, A., Shaw, D.G., Goral, M., Wisniewska-Gocłowska, B., 2007a. IUPAC-NIST Solubility Data Series. 82. Alcohols with Water—Revised and Updated: Part 1. C4 Alcohols with Water. *Journal of Physical and Chemical Reference Data* 36, 59-132.
- Maczynski, A., Shaw, D.G., Goral, M., Wisniewska-Gocłowska, B., 2007b. IUPAC-NIST Solubility Data Series. 82. Alcohols with Water—Revised and Updated: Part 3. C6 Alcohols with Water. *Journal of Physical and Chemical Reference Data* 36, 399-443.
- Maczynski, A., Shaw, D.G., Goral, M., Wisniewska-Gocłowska, B., 2007c. IUPAC-NIST Solubility Data Series. 82. Alcohols with Water—Revised and Updated: Part 4. C7 Alcohols with Water. *Journal of Physical and Chemical Reference Data* 36, 445-484.
- Maczynski, A., Shaw, D.G., Goral, M., Wisniewska-Gocłowska, B., 2007d. IUPAC-NIST Solubility Data Series. 82: Alcohols with Water—Revised and Updated: Part 2. C5 Alcohols with Water. *Journal of Physical and Chemical Reference Data* 36, 133-190.
- Magnussen, T., Rasmussen, P., Fredenslund, A., 1981. UNIFAC Parameter Table For Prediction of Liquid-Liquid Equilibria. *Industrial & Engineering Chemistry Process Design and Development* 20, 331-339.
- Mullins, E., Oldland, R., Liu, Y.A., Wang, S., Sandler, S.I., Chen, C.C., Zwolak, M., Seavey, K.C., 2006. Sigma-profile database for using COSMO-based thermodynamic methods. *Industrial & Engineering Chemistry Research* 45, 4389-4415.
- Peniston, Q.P., 1956. Manufacture of 5-Hydroxymethyl 2-Furfural, U.S. Patent 2750394.

Renon, H., Prausnitz, J.M., 1968. Local compositions in thermodynamic excess functions for liquid mixtures. *Aiche Journal* 14, 135-144.

Roman-Leshkov, Y., Chheda, J.N., Dumesic, J.A., 2006. Phase Modifiers Promote Efficient Production of Hydroxymethylfurfural from Fructose. *Science* 312, 1933-1937.

Román-Leshkov, Y., Dumesic, J.A., 2009. Solvent Effects on Fructose Dehydration to 5-Hydroxymethylfurfural in Biphasic Systems Saturated with Inorganic Salts. *Topics in Catalysis* 52, 297-303.

Saha, B., Abu-Omar, M.M., 2014. Advances in 5-hydroxymethylfurfural production from biomass in biphasic solvents. *Green Chemistry* 16, 24-38.

Stephenson, R.M., 1992. Mutual Solubilities - Water Ketones, Water Ethers, and Water Gasoline Alcohols. *Journal of Chemical and Engineering Data* 37, 80-95.

Stephenson, R.M., Stuart, J., 1986. Mutual Binary Solubilities - Water Alcohols and Water Esters. *Journal of Chemical and Engineering Data* 31, 56-70.

Stephenson, R.M., Stuart, J., Tabak, M., 1984. Mutual Solubility of Water and Aliphatic-alcohols. *Journal of Chemical and Engineering Data* 29, 287-290.

van Putten, R.J., van der Waal, J.C., de Jong, E., Rasrendra, C.B., Heeres, H.J., de Vries, J.G., 2013. Hydroxymethylfurfural, A Versatile Platform Chemical Made from Renewable Resources. *Chemical Reviews* 113, 1499-1597.

Wang, T., Nolte, M.W., Shanks, B.H., 2014. Catalytic dehydration of C6 carbohydrates for the production of hydroxymethylfurfural (HMF) as a versatile platform chemical. *Green Chemistry* 16, 548-572.

Xiong, R., Sandler, S.I., Burnett, R.I., 2014. An Improvement to COSMO-SAC for Predicting Thermodynamic Properties. *Industrial & Engineering Chemistry Research* 53, 8265-8278.

Table 1. HMF extraction selectivity ($R_{\text{HMF,c}}/R_{\text{i,c}}$; i= Fructose, LA or FA) at 30°C for each of the solvents tested.

Extractant	Selectivity to HMF over		
	Fructose	LA	FA
1-BuOH	53.2	1.42	1.55
1-PeOH	166.5	1.40	2.39
1-HexOH	137.8	1.23	2.48
1-HepOH	139.9	0.33	6.72
i-BuOH	66.1	1.53	1.78
i-PeOH	155.3	1.57	2.12

Table 2. Activity coefficients calculated from COSMO-SAC for HMF and fructose in each phase and the partition coefficients at 30°C.

Solute	Solvent	Activity Coefficient		Partition Coefficient	
		Phase I (water-rich phase)	Phase II (alcohol-rich phase)	$R_{i,\infty} = C_i^{\text{organic}} / C_i^{\text{aqueous}} = C^I / C^I \times (\gamma_i^I / \gamma_i^{\text{II}})$	
				COSMO-SAC (2013)-ADF	Experiment
HMF	1-butanol	4.73	1.17	1.39	1.59
	1-pentanol	5.57	1.55	0.89	1.14
	1-hexanol	5.82	1.85	0.65	0.63
	1-heptanol	5.94	2.42	0.42	0.33
Fructose	1-butanol	0.25	1.72	0.051	0.030
	1-pentanol	0.27	5.33	0.013	0.006
	1-hexanol	0.28	8.76	0.007	0.004
	1-heptanol	0.28	22.23	0.002	0.002

Table 3. The individual contributions to the activity coefficients in Phase II (alcohol-rich phase): γ_i^{res} is from the restoring free energy (electrostatic contribution); γ_i^{disp} is from the dispersive contribution; and γ_i^{comb} is from the combinatorial free energy (cavity contribution). The temperature is 30°C.

Solute	Solvent	γ_i^{res}	γ_i^{disp}	γ_i^{comb}
HMF	1-butanol	1.53	1.00	0.76
	1-pentanol	1.70	1.00	0.91
	1-hexanol	1.92	1.00	0.97
	1-heptanol	2.42	1.00	1.00
Fructose	1-butanol	2.46	1.12	0.63
	1-pentanol	5.57	1.16	0.83
	1-hexanol	8.30	1.16	0.91
	1-heptanol	18.88	1.19	0.99

Scheme 1. Schematic of fructose dehydration, HMF rehydration and humin formation reactions. Removing HMF from the reaction mixture as it is produced is a frequently used method for minimizing its degradation to side products.

Figure 1. Comparison of the liquid–liquid equilibrium predictions with experimental data for binary systems of water and (a) 1-butanol, (b) 1-pentanol, (c) 1-hexanol and (d) 1-heptanol, respectively. The open circles are the experimental data taken from Stephenson *et al.* (Stephenson and Stuart, 1986; Stephenson *et al.*, 1984). The solid and dashed lines are the predictions from COSMO-SAC (2013)-ADF (Xiong *et al.*, 2014) and COSMO-SAC (2010) (Hsieh *et al.*, 2010), respectively, w_1 is weight percent of the alcohol solvent in the two equilibrated liquid phases, and T is the system temperature.

Figure 2. Measured concentrations of the alcohol solvents in the water-rich phase (a) and water concentrations in the alcohol solvent-rich phase (b) after the partitioning of a single solute in alcohol solvent and water solution at 30°C. Values labeled “No solute” are the equilibrium concentrations for the binary alcohol solvent and water system, which are similar to the literature values(De Santis et al., 1976; Maczynski et al., 2007a, b, c, d; Stephenson and Stuart, 1986; Stephenson et al., 1984). The numbers on the x-axis represent the mixtures with different alcohols: 1) 1-butanol, 2) 1-pentanol, 3) 1-hexanol, 4) 1-heptanol, 5) isobutanol, and 6) isopentanol, respectively.

Figure 3. Alcohol-water partition coefficients of HMF, LA, FA, and fructose at 30°C. The measured data (R_c , this work) are shown in (a) and the comparisons of COSMO-SAC (2013)-ADF(Xiong et al., 2014) predictions ($R_{c,\infty}$) with the measured data are shown in (b) for HMF, (c) for LA, (d) for FA, and (e) for fructose. The numbers on the x-axis represent the different alcohols: 1) 1-butanol, 2) 1-pentanol, 3) 1-hexanol, 4) 1-heptanol, 5) isobutanol, and 6) isopentanol, respectively. In b,c,d and e, the open symbols are the measured data and the solid symbols are the predictions. The lines are to guide the eye.

Figure 4. σ -profiles (multiplied by the surface area) from COSMO-SAC (2013)-ADF(Xiong et al., 2014) for (a) the non-HB parts, and (b) the HB parts of the sigma profiles (b) of the four linear alcohols: 1-butanol, 1-pentanol, 1-hexanol and 1-heptanol. (HB: hydrogen-bonding; σ : charge density).

Figure 5. (a) the predicted liquid-liquid equilibrium (LLE) phase diagram at 30°C for the HMF-alcohol-water system from COSMO-SAC (2013)-ADF(Xiong et al., 2014). The dashed lines are the tie lines connecting the equilibrium concentrations in the two phases. (b) The predicted alcohol-water partition coefficients (R_c) using eqn. 1 as a function of the length of the tie lines. The partition coefficients at the longest length of each tie line are about the same as the values calculated using eqn. 2 and are comparable to the experimental data.