PROPERTY PREDICTION IN REACTIVE SOLUTIONS

Whitepaper

In reactive solutions, reaction educts and products are coexistent and their concentrations are unknown. Macroscopic property prediction however requires a well defined system composition. Thus, physical interactions in the liquid phase as well as the chemical equilibrium between the reacting species have to be taken into account in such systems. With standard COSMO-RS, interactions between molecules are described with statistical thermodynamics of interacting surface segments. Thereby, all spatial information is neglected. Concerted multiple contacts of molecules, like in compound dimers, or transformation to covalently bound species, like polymerization, are not captured with COSMO-RS. In this study COSMO-RS-DARE, an extension of COSMO-RS for reactive solutions, is evaluated for the binary liquid-liquid equilibrium system of formaldehyde + water + 1-butanol. After fitting the COSMO-RS-DARE parameters to experimental data for the phase compositions, the liquid-liquid equilibrium is calculated at several temperatures. The procedure is explained in detail.

INTRODUCTION

Property prediction of mixtures is complicated when reactions between solute and solvent or solute and solute occur and the chemical equilibrium is concentration dependent. Examples for such reactions are the concentration dependent dimerization of small organic acids (solute-solute reaction) or reversible polymerization. In experimental property measurements, the influence of such a reaction on the measured property is usually not separated from other effects. For computational prediction methods however, it is important to take into account that[1] in addition to the original solute(s) and solvent(s), reaction products are also present in the mixture and[2] as a result, the original phase composition is changed. Therefore, a reliable prediction method has to consider physical interactions in the liquid phase as well as the chemical equilibrium between the reacting species.

Formaldehyde reacts with water and alcohols to form oligomers. In aqueous solution, formaldehyde is mostly present as methylene glycol and poly(oxymethylene) glycols. In solution with alcohols, formaldehyde reacts similarly and forms hemiformal and poly(oxymethylene) hemiformals with the alcoholic species present. In a ternary mixture of formaldehyde, water, and alcohol, these reactions take place simultaneously, and the equilibrium concentration of monomeric formaldehyde is usually very small.

$$CH_2O + H_2O \rightleftharpoons HO(CH_2O)H \qquad (I)$$

$$HO(CH_2O)_{n-1}H + HO(CH_2O)H \rightleftharpoons HO(CH_2O)nH + H_2O \qquad (II)$$

$$CH_2O + ROH \rightleftharpoons HO(CH_2O)R \qquad (III)$$

$$HO(CH_2O)_{n-1}R + HO(CH_2O)R \rightleftharpoons HO(CH_2O)_nR + ROH \qquad (IV)$$

If the alcohol in the mixture is 1-butanol (or a monohydric alcohol with a longer aliphatic chain), the system will form two liquid phases I and II. The condition for the liquid-liquid phase equilibrium is

$$\mu_i^{I} + RT \ln x_i^{I} = \mu_i^{II} + RT \ln x_i^{II}$$

for all compounds *i*.

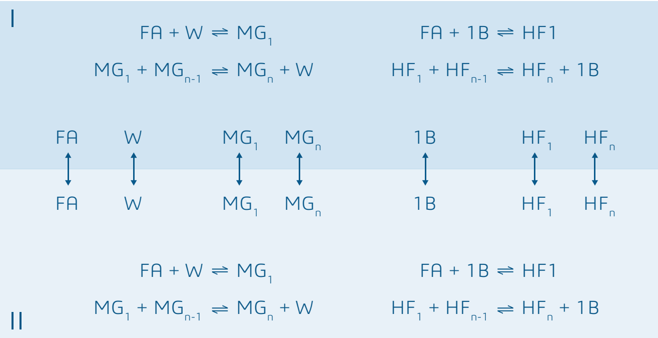


Figure 1: The two-phase liquid-liquid equilibrium system of formaldehyde + water + 1-butanol

Peschla et al.[1] reported experimental results for the chemical equilibrium in the binary formaldehyde – 1-butanol system and the ternary formaldehyde – water – 1-butanol system at temperatures from 278 to 348 K. Liquid-liquid phase equilibrium data are reported for temperatures of 298, 313, and333 K.

Computational Methods

COSMO-RS (COnductor like Screening MOdel for Real Solvents) is a universal theory to predict the thermodynamic equilibrium properties of liquids.[2] COSMO-RS thermodynamics is based on the statistical physics of interacting molecular surface segments. The polar and hydrogen bond interaction energies are quantified based on the surface screening charge densities, which result from a quantum chemical continuum solvation calculation. For details we refer the reader to the literature.[2,3] Due to its ability to treat mixtures at variable temperatures and to compute accurate solvation energies based on first-principles, it has become a much used method in chemical engineering and in different areas of chemistry.

COSMO-RS DARE[4] (Dimerization, Aggregation, and Reaction Extension) is an extension of COSMO-RS which equilibrates molecular species with the corresponding entities in the reaction products (or dimer or aggregate, respectively). Molecular conformers and the corresponding parts of the reaction product are treated in an artificial conformer equilibrium. The contacts between reaction product entities are described using artificial surface segments. Only predefined contacts are allowed for such surface segments. For details, we refer the reader to the corresponding publication.[4]

COSMO-RS DARE is implemented in COSMOtherm using so-called Interaction Energy Indices (IEI numbers). The reacting compounds are treated as pseudo-conformers in the compound input of COSMOtherm. To describe the reaction product, an interaction energy is assigned to two interacting confomers.

To illustrate this procedure, we will describe it in more detail for an equilibrium reaction A + B ↔ A-B. First we have to provide the interacting species:

- The first conformer of compound A is the free compound A, and the second conformer is a COSMO metafile of the reaction product A-B. In the reaction product all atoms of the reaction partner B are set to zero by atomic weights (refer to the COSMOtherm Reference Manual for details). Thus the second conformer describes compound A after the reaction, or, more accurately, the "A" entity in the reaction product.
- Compound B is constructed similarly: one conformer is pure compound B and a second conformer is a meta-file of A-B with all atomic weights of A set to zero.
- The conformers described by the COSMO meta-file have to be identified by a unique interaction energy index (IEI number). This is done with the IEI=i command in the same line where the conformer is given.

Then we have to describe the interaction between the species in the reaction:

- The interaction energy has to be assigned in the mixture part of the COSMOtherm input file. This is done with the en_IEI={i1 i2 c_H c_S} option, where i1 and i2 are the IEI numbers of the reacting compound as given in the compound input section.
- c_H and c_S are parameters for the enthalpic and entropic contribution to the interaction energy of the two IEI numbered compounds. c_H and c_S are [kcal/mol] and [kcal/mol K], respectively.
- The interaction Gibbs free energy between the two IEI compounds i1 and i2 is calculated from

$$G\,(i_1, i_2) = -2\,\Delta^0 + c\_H - T\,c\_S$$

  where Δ^0 is the energy difference between the first conformer (e.g. A) and the meta-file conformer. The value of Δ^0 is calculated by COSMOtherm.
- It is possible to give several en_IEI entries in one mixture input line (for the treatment of several different reactions).

It is important to notice that the conformers described by the meta-files, i.e. the "A" and "B" entities in the reaction product, cannot be treated in the ususal way for conformer equilibrium. This is not possible because the quantum chemical COSMO energy of a COSMO metafile is not known.

In this study, we have to take into account the following species: Formaldehyde and 1-butanol form hemiformal HF1 and different poly(oxymethylene) hemiformals HFn. With water, formaldehyde forms methylene glycol MG1 and poly(oxymethylene) glycols MGn. Thus, we require the following pseudo-conformers in COSMOtherm:

For formaldehyde:

- monomeric formaldehyde
- the formaldehyde entity from the hemiformal HF1 ($CH_3(CH_2)_3OCH_2OH$)
- a central formaldehyde entity from a poly(oxymethylene) hemiformal. We use the poly(oxymethylene) hemiformal with three formaldehyde units, HF3 ($CH_3(CH_2)_3O(CH_2O)_3H$).
- the formaldehyde entity from the methylene glycol MG1 ($HOCH_2OH$)
- a central formaldehyde entity from a poly(oxymethylene) glycol. We use the poly(oxymethylene) glycol with three formaldehyde units, MG3 ($HO(CH_2O)_3H$).

For 1-butanol:

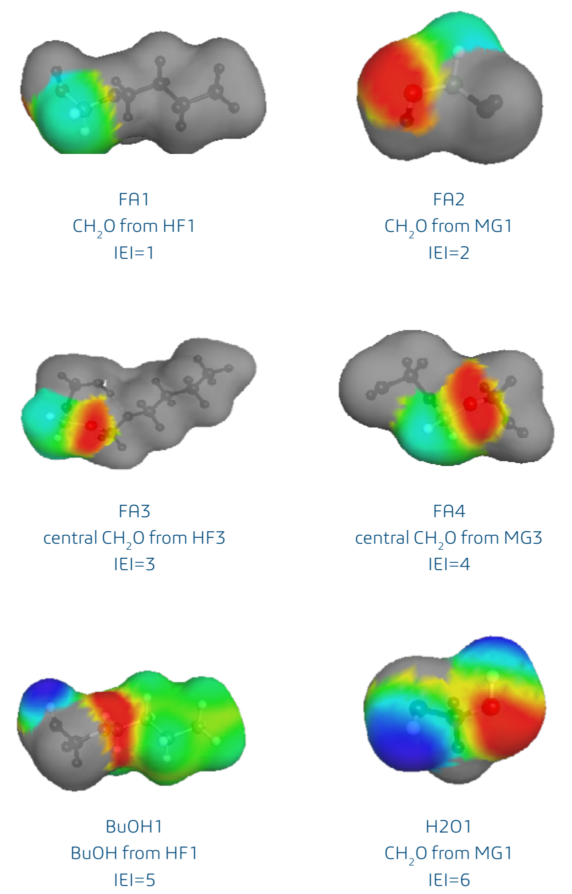
- pure 1-butanol
- the 1-butanol entity from the hemiformal HF1. This fragment is also used to model the 1-butanol entity from the poly(oxymethylene) hemiformal.

For water:

- pure water
- the water entity from the methylene glycol MG1. This fragment is also used to model the water entity from the poly(oxymethylene) glycol.

Since the 1-butanol and water entities are terminal fragments, we can reasonably assume that the corresponding fragments from hemiformal or methylene glycol and the higher oligomers poly(oxymethylene) hemiformal or poly(oxymethylene) glycol , respectively, are equivalent. We won't use the same approximation for the formaldehyde fragments, however, because the two formaldehyde fragments have different next neighbors. Figure 2 shows the DARE conformers used in this study.

The COSMO surfaces were obtained from quantum chemical geometry optimizations on the BP/TZVP level, using TURBOMOLE 7.0[5]. The corresponding BP_TZVP_C30_1501 parameterization was used in the COSMOtherm calculations.[6]



FA1
CH_2O from HF1
IEI=1

FA2
CH_2O from MG1
IEI=2

FA3
central CH_2O from HF3
IEI=3

FA4
central CH_2O from MG3
IEI=4

BuOH1
BuOH from HF1
IEI=5

H2O1
CH_2O from MG1
IEI=6

Figure 2: DARE conformers used to model the chemical equilibrium and the liquid-liquid equilibrium.

PARAMETER FITTING

The interaction energy parameters c_H and c_S for the oligomerization reactions of formaldehyde with 1-butanol and water are not known. Therefore, we will fit them to the experimental data for the chemical equilibrium of the different formaldehyde species in the binary formaldehyde – 1-butanol and ternary formaldehyde – 1-butanol – water systems, starting with the binary system. Additional interaction energy parameters for the ternary system are fitted in a second step of the procedure. Breaking down the optimization procedure into two steps reduces the complexity of the process.

The interaction energy parameters for the contacts between the reaction product entities were fitted on a grid such that the deviation between the calculated and the experimental mole fractions are minimized. It is of course possible to use other optimization methods. Table 1 shows the 10 parameter sets with the lowest sum of MAD (mean absolute deviation) for the mole fraction which were selected for the next step. An exemplary input file for the COSMOtherm calculation is shown in Figure 3.

set	H_13	S_13	H_15	S_15	MAD sum	G_13	G_15			
						at 313.15 K				
1	-3.6	0.007	-3.1	0.003	27.015	-5.79	-4.04			
2	-4.2	0.005	-3.1	0.003	27.023	-5.77	-4.04			
3	-3.8	0.006	-3.1	0.003	27.038	-5.68	-4.04			
4	-4.6	0.003	-3.3	0.002	27.066	-5.54	-3.93			
5	-3.2	0.008	-3.1	0.003	27.080	-5.71	-4.04			
6	-4	0.005	-3.3	0.002	27.088	-5.57	-3.93			
7	-3.4	0.007	-3.3	0.002	27.110	-5.59	-3.93			
8	-3.4	0.008	-2.9	0.004	27.119	-5.91	-4.15	-5.75	-4.03	average
9	-3.2	0.009	-2.9	0.004	27.120	-6.02	-4.15	0.13	0.06	mean abs dev
10	-3.4	0.008	-3.1	0.003	27.136	-5.91	-4.04	0.026	0.0070	s2

Table 1: Parameter sets and MAD sum for the binary formaldehyde – 1-butanol system.

```
ctd=BP_TZVP_C30_1501.ctd
cdir=/software/COSMOtherm-C30-1501/CTDATA-FILES
fdir=../Results_of_job_BP-TZVP-COSMO wconf wcmn
!!
[  f =formaldehyde_c0.cosmo comp=FA
   f =FA-hemiformal.mcos IEI=1
   f =FA-polyoxymethylenehemiformal-3.mcos ] IEI=3
[  f =1-butanol_c0.cosmo comp=1B
   f =1-butanol_c1.cosmo
   f =1B-hemiformal.mcos ] IEI=5
tk=278.15 x={0.5446 0.4554} EN_IEI={1 3 <H_13> <S_13>}
EN_IEI={1 5 <H_15> <S_15>}
```

Figure 3: Compound section and exemplary mixture input for a COSMOtherm calculation using DARE pseudo-conformers and interaction energy indices for the binary system formaldehyde – 1-butanol. For fitting the parameters H_13, S_13, H_15, and S_15 vary in this input type.

In Figure 4 experimental and calculated mole fractions of formaldehyde units are plotted at various temperatures, using the final set of parameters (set 7). Visual differences to mole fraction calculated with parameter set 1 (not shown) are negligible. When comparing to UNIFAC results published by Peschla et al.[1] it is important to keep in mind that COSMO-RS-DARE requires 4 parameters for the binary formaldehyde – 1-butanol system in addition to the standard COSMO-RS parameterization, while UNIFAC requires much more parameter fitting: 10 binary interaction parameters and 10 group parameters.

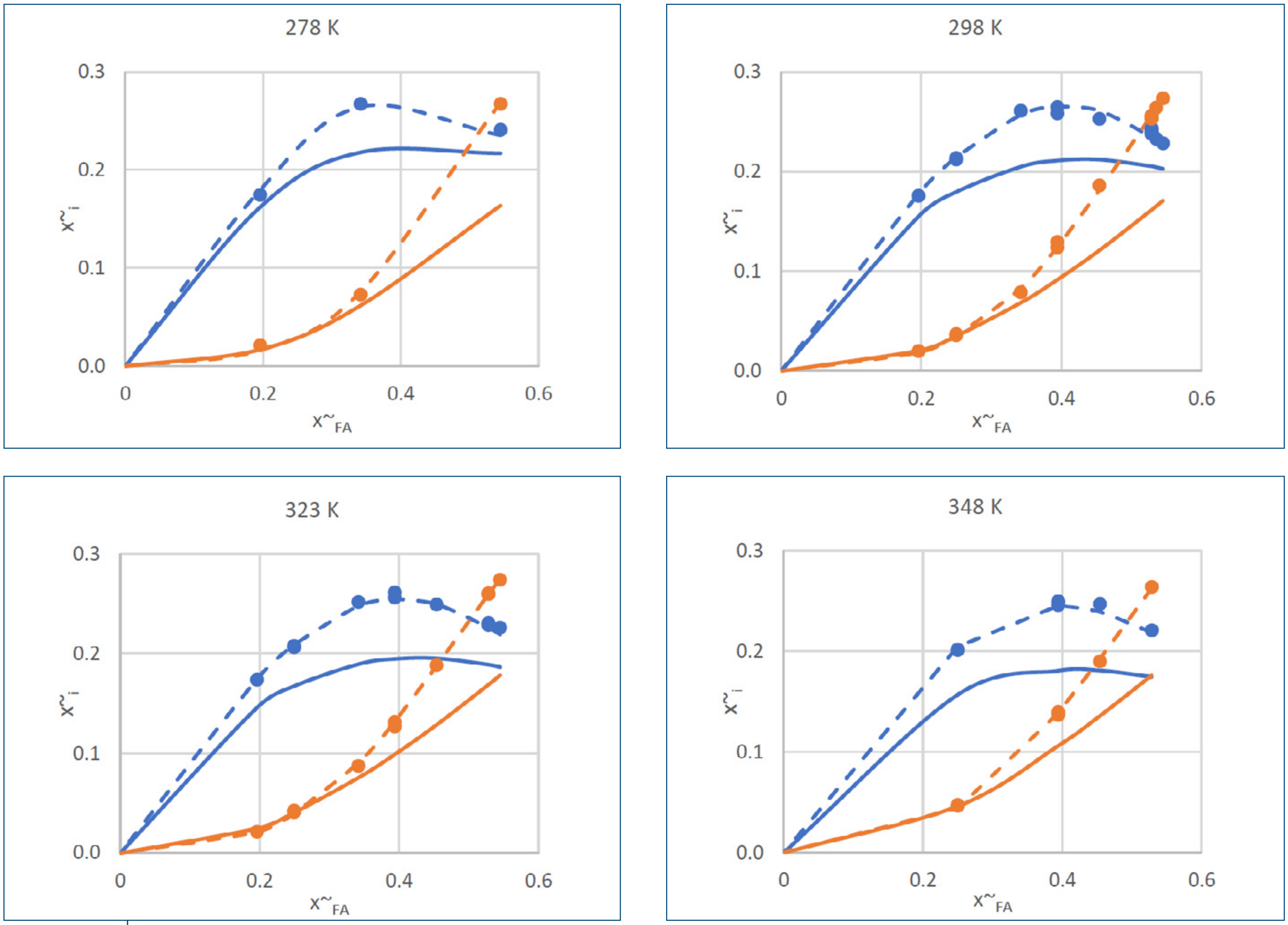


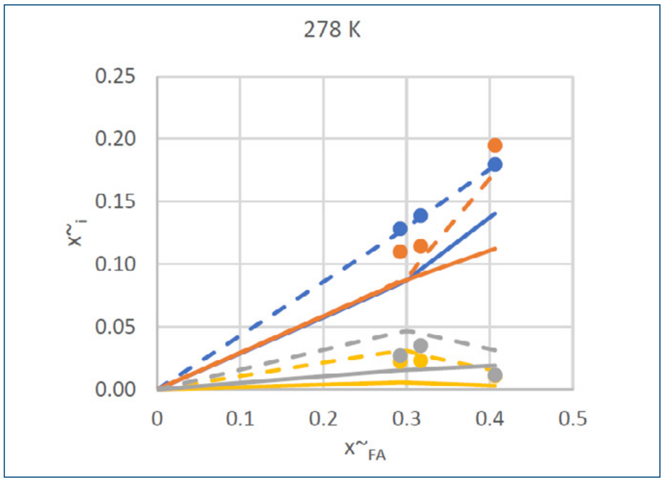
Figure 4: Experimental and calculated mole fractions of formaldehyde units in hemiformal HF1 and poly(oxymethylene) hemiformal HF3 at 278 K, 298 K, 323 K, and 348 K. Blue dots: exp mole fraction of HF1, blue lines: calc mole fraction of HF1 (dashed: UNIFAC, solid: COSMOtherm), orange dots: exp mole fraction of HF3 units, orange lines: calc mole fraction of HF3 units (dashed: UNIFAC, solid: COSMOtherm). UNIFAC results taken from ref. 1.

In the second step the IEI parameters for the reactions of formaldehyde and water to methylene glyclol and poly(oxymethylene) glycols were optimized. The interaction energy parameters for the contacts between the reaction product entities were fitted on a grid, using the fitted IEI parameter sets for formaldehyde and 1-butanol fragments and keeping them fixed. An example input file is shown in Figure 5. As in the first step, the deviation between calculated and experimental mole fractions are minimized on a grid. This procedure results in ~ 500 parameters sets with a MAD sum of less than 40. Figure 6 shows the experimental and calculated mole fractions of formaldehyde units at various temperatures, using the final set of parameters.

```
ctd=BP_TZVP_C30_1501.ctd cdir=/software/
COSMOtherm-C30-1501/CTDATA-FILES
fdir=../Results_of_job_BP-TZVP-COSMO wconf wcmn
!!
[  f =formaldehyde_c0.cosmo comp=FA
   f =FA-hemiformal.mcos IEI=1
   f =FA-metyleneglycol.mcos IEI=2
   f =FA-polyoxymethylenehemiformal-3.mcos IEI=3
   f =FA-polyoxymetyleneglycol-3.mcos ] IEI=4
[  f =1-butanol_c0.cosmo comp=1B
   f =1-butanol_c1.cosmo
   f =1B-hemiformal.mcos ] IEI=5
[  f =h2o_c0.cosmo comp=W
   f =W-metyleneglycol.mcos ] IEI=6
tk=278.15 x={0.29 0.26 0.45} EN_IEI={1 3 -3.6 0.007} EN_
IEI={1 5 -3.1 0.003} EN_IEI={2 4 <H_24> <S_24>} EN_IEI=
{2 6 <H_26> <S_26>}
```

Figure 5: Compound section and exemplary mixture input for a COSMOtherm calculation using DARE pseudo-conformers and interaction energy indices for the ternary system formaldehyde – 1-butanol – water. The parameters H_24, S_24, H_26, and S_26 vary in this input type.

All ~500 sets of optimized IEI parameters were used to calculate the Liquid-Liquid Equilibrium (LLE) of the ternary system formaldehyde – 1-butanol – water at different temperatures. Best agreement with experimental data was obtained with the following parameters: H_13=-3.4, S_13=0.007, H_15=-3.3, S_15=0.002, H_24=-2.8, S_24=0.005, H_26=-0.1, S_26=0.001. This is our final parameter set for the modeling of the ternary system formaldehyde– 1-butanol – water.



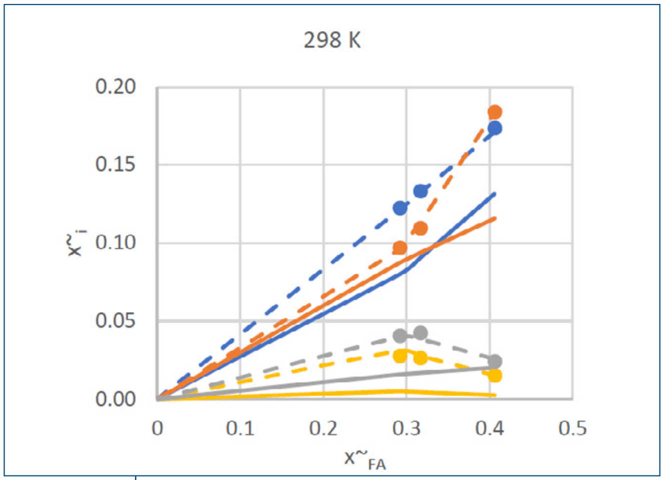
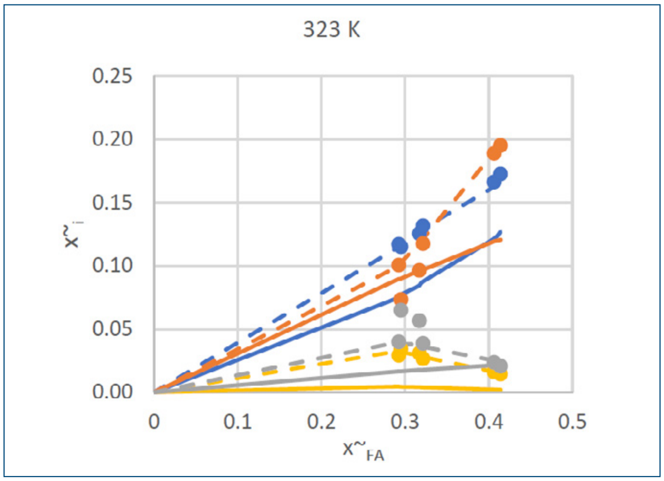
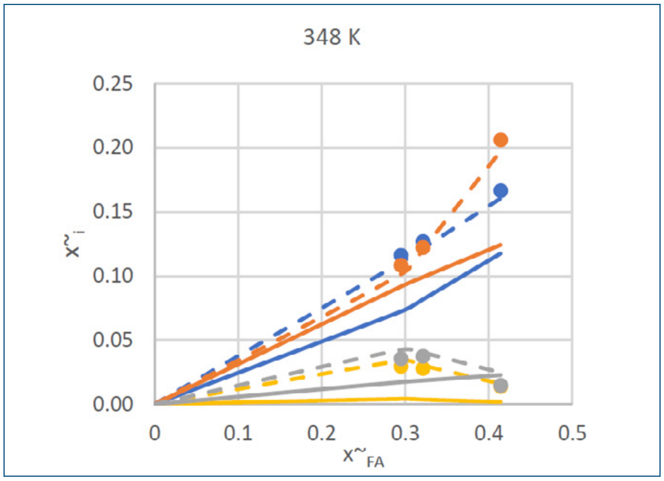


Figure 6: Experimental and calculated mole fractions of formaldehyde units in hemiformal HF1, poly(oxymethylene) hemiformal HF3, methylene glycol MG1, and poly(oxymethylene) glycol MG3 at 278 K, 298 K, 323 K, and 348 K. Blue dots: exp. mole fraction of HF1, blue lines: calc. mole fraction of HF1 (dashed: UNIFAC, solid: COSMOtherm), orange dots: exp. mole fraction of HF3 units, orange lines: calc. mole fraction of HF3 units (dashed: UNIFAC, solid: COSMOtherm), grey dots: exp. mole fraction of MG1, grey lines: calc. mole fraction of MG1 (dashed: UNIFAC, solid: COSMOtherm), yellow dots: exp. mole fraction of MG3 units, yellow lines: calc. mole fraction of HF3 units (dashed: UNIFAC, solid: COSMOtherm). UNIFAC results taken from ref. 1





As shown in Figure 7, the LLE predicted with COSMO-RS-DARE is in much better agreement with the experimental results than the LLE predicted with pure COSMO-RS. Differences are particularly apparent for the water-rich phase. The largest deviations from experimental results occur for the system without formaldehyde. It should be noted that in this case, which corresponds to the binary 1-butanol – water system, the COSMO-RS-DARE approach has no effect on the system, and the deviations correspond to the COSMO-RS error. For higher formaldehyde concentrations the difference between COSMO-RS and COSMO-RS-DARE becomes larger and the COSMO-RS-DARE predictions are much closer to experimental results than the COSMO-RS predictions.

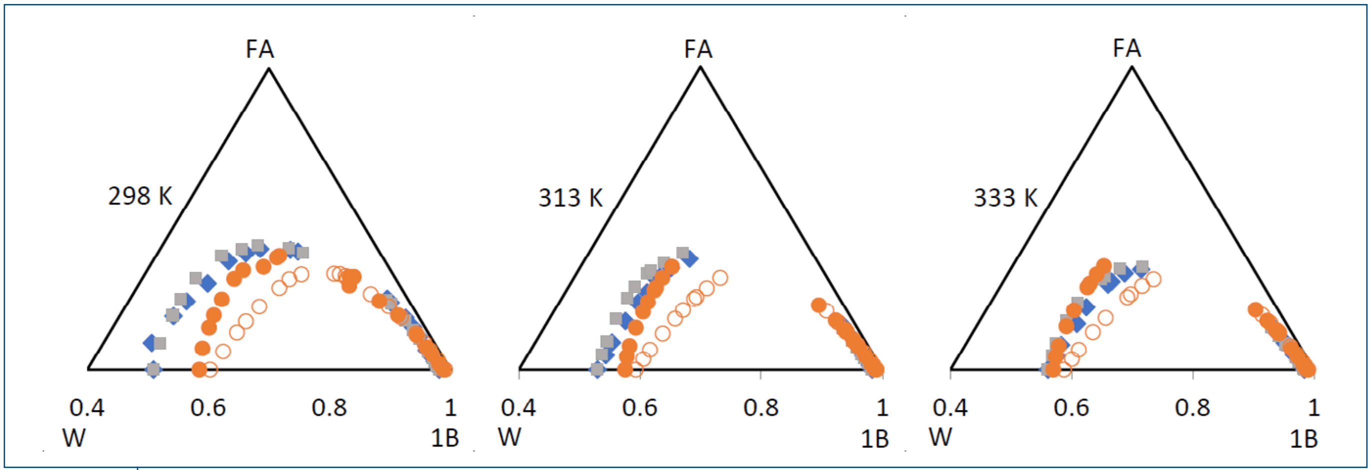


Figure 7: Liquid-liquid phase equilibrium of the ternary system formaldehyde (FA) – 1-butanol (1B) – water (W) at 298 K, 313 K, and 333 K. Blue diamonds: experimental results, grey squares: UNIFAC predictions [1], full orange circles: COSMO-RS-DARE predictions, open orange circles: COSMO-RS predictions. UNIFAC results taken from ref. 1.

When comparing with UNIFAC results it should be taken into account that, while the UNIFAC results are overall closer to experimental results, much more parameter fitting is involved. The simplified UNIFAC model used by Peschla et al. requires 32 fitted group interaction parameters and 14 group size and surface parameters. Furthermore, it relies on experimental data for chemical reaction equilibrium constants and vapor pressure, which implies additional lab work has to be done.

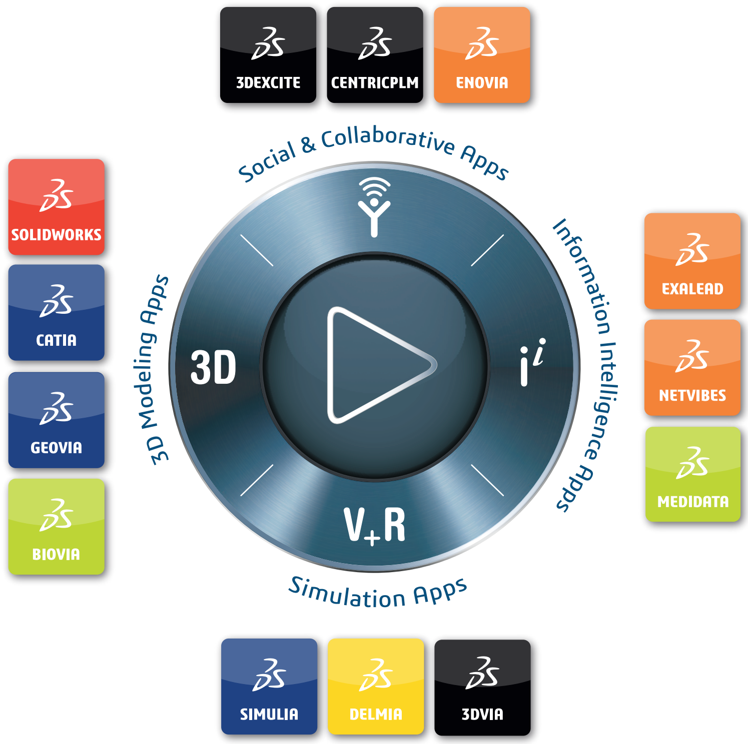
CONCLUSION

In this study, COSMO-RS-DARE parameters were fitted to the reactions occurring in the ternary system formaldehyde– 1-butanol– water at various temperatures. We were able to show that for such a complex reactive system, the COSMO-RS-DARE approach works well and leads to better predictions for the ternary LLE of the system, especially for higher formaldehyde concentrations.

COSMO-RS-DARE requires fitting of 2 parameters for each reaction occurring in a system; for systems under isothermal conditions the parameters can be combined to one parameter, thus simplifying the fitting procedure. For the temperature dependent reactions considered in this study, 8 energy interaction parameters were fitted in total. The corresponding UNIFAC study[1] required 46 interaction and group parameters, some of which were taken from a previous study of the binary formaldehyde – water system.[7]

REFERENCES

1. "Chemical Equilibrium and Liquid-Liquid Equilibrium in Aqueous Solutions of Formadelhyde and 1-Butanol", Peschla, R.; Coto Garcia, B.; Albert, M.; Kreiter C.; Maurer G.; Ind. Eng. Chem. Res. 2003, 42, 1508-1516.
2. Klamt, A.; The COSMO and COSMO-RS solvation models Wiley Interdisciplinary Reviews: Computational Molecular Science, John Wiley & Sons, Inc., 2011.
3. Klamt A.; COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design, Elsevier Science Ltd., Amsterdam, The Netherlands, 2005.
4. "Consideration of dimerization for property prediction with COSMO-RS-DARE" Sachsenhauser, T.; Rehfeld, S.; Klamt, A.; Eckert, F., Fluid Phase Equilibria 2014, 382, 89-99. doi: http://dx.doi.org/10.1016/j.fluid.2014.08.030
5. TURBOMOLE V7.0 2015, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
6. COSMOtherm, Version (e.g. C3.0, release 1501), COSMOlogic GmbH & Co KG, www.cosmologic.de.
7. "Vapor-Liquid Equilibria for Aqueous Solutions of Formaldehyde and Methanol", Albert, M.; Coto Garcia, B.; Kuhnert, Ch.; Peschla, R.; Maurer, G., AIChE J. 2000, 46, 1676-1687.



Our **3D**EXPERIENCE® platform powers our brand applications, serving 11 industries, and provides a rich portfolio of industry solution experiences.

Dassault Systèmes, the **3D**EXPERIENCE Company, is a catalyst for human progress. We provide business and people with collaborative virtual environments to imagine sustainable innovations. By creating 'virtual experience twins' of the real world with our **3D**EXPERIENCE platform and applications, our customers push the boundaries of innovation, learning and production.

Dassault Systèmes' 20,000 employees are bringing value to more than 270,000 customers of all sizes, in all industries, in more than 140 countries. For more information, visit **www.3ds.com**.

Europe/Middle East/Africa
Dassault Systèmes
10, rue Marcel Dassault
CS 40501
78946 Vélizy-Villacoublay Cedex
France

Asia-Pacific
Dassault Systèmes K.K.
ThinkPark Tower
2-1-1 Osaki, Shinagawa-ku,
Tokyo 141-6020
Japan

Americas
Dassault Systèmes
175 Wyman Street
Waltham, Massachusetts
02451-1223
USA

DASSAULT SYSTEMES | The **3D**EXPERIENCE® Company