Prediction of the temperature dependence of a polyether-water mixture using COSMO*therm*

By Jens Reinisch[1]*, Andreas Klamt[1,2], Frank Eckert[1] and Michael Diedenhofen[1]

[1] COSMO*logic* GmbH&CoKG, Burscheider Str. 515, 51381 Leverkusen, Germany

[2] Inst. of Physical and Theoretical Chemistry, University of Regensburg, Germany

* corresponding author, Tel.: +49-2171-363664, Fax.: +49-2171-731680, e-mail: reinisch@cosmologic.de

keywords: molecular simulation; COSMO-RS; liquid-liquid equilibria; solubility; lower critical solution

temperature

Abstract

As entry for the 6th industrial fluid property simulation challenge, the COSMO-RS method in its COSMO*therm*

implementation has been used to predict the mutual solubilities of dipropylene glycol dimethyl ether (DPGDME)

and water. The miscibility gaps and their unusual inverse temperature dependence have been qualitatively

correctly predicted with COSMO*therm* standard procedure. For quantitative agreement some adjustments based

on experimental data at 298 K turned out to be necessary. Fine tuning of the water and DPGDME areas led to a

good quantitative agreement with experimental data with a maximum deviation of 5.4 mass percent points,

which turned out to be the most accurate predictions within all challenge submissions.

1. Introduction

Latest since its success in the 1st Industrial Fluid Properties Simulation Challenge [1,2] the quantum chemically

based Conductor-like Screening Model for Real Solutes (COSMO-RS) [3-7] has become one of the industrial

standard tools for the simulation of fluid phase thermodynamics and especially for solvent and solute screening.

This demonstrates the demand for an alternative predictive method between the rather empirical, structure-

interpolating group contribution methods and the rather time-consuming force-filed based molecular dynamics

and Monte Carlo simulation techniques, which still suffer from the limitations of force filed transferability.

COSMO-RS combines quantum chemical calculations for solutes and solvents with statistical thermodynamics

of interacting surfaces and requires only a small set of empirically adjusted parameters. Since the

thermodynamic simulations only require seconds, COSMO-RS combines predictive power and calculation efficiency.

The 6th IFPSC requested the prediction of the temperature dependent liquid-liquid equilibrium (LLE) in the binary mixture of dipropylene glycol dimethyl ether (DPGDME) and water. Many glycol-di-ethers show an inverse temperature dependence of the water solubility, i.e. increased solubility at lower temperatures, resulting from a sophisticated interplay of hydrogen bonding and hydrophobic interactions. The proper simulation of this feature is a big challenge for any simulation method. Within this compound class, DPGDME is even more special, as it has no hydroxyl-group and can thus be used as solvent in proton sensitive systems. Hence the 6th IFPSC requested the simulation of a mixture of an unusual representative of an unusual class of conformationally flexible and relatively large solvent molecules with the small molecules of the very special, strongly hydrogen bonding solvent water. In the following we will outline how we addressed this challenge within the frame work of COSMO-RS.

2. The COSMO-RS method

A detailed description of COSMO-RS is given elsewhere [3-7] and beyond the scope of this article. Thus only the basic features required for the understanding of the results shall be presented here.

COSMO-RS is a two step procedure. In a first step, quantum chemical calculations have to be performed for all compounds of interest. In these calculations the continuum solvation model COSMO [7] is applied in order to simulate a virtual conductor environment for the molecule. In this environment, the solute molecule induces a polarization charge density σ on the interface of the molecule to the conductor, i.e. on the molecular surface, and these charges act back on the solute, generating a more polarized electron density than vacuum does. Solving the quantum mechanical self-consistency cycle the solute molecule is converged to its energetically optimal state in a conductor with respect to electron density and geometry. The standard quantum chemical method for COSMO-RS is density functional theory (DFT) and the DFT functional B88-P86 [8,9] with a triple zeta valence polarized basis set (TZVP) and RI approximation [10,11] has been used throughout this study. All DFT/COSMO calculations have been performed by the quantum chemical program TURBOMOLE [12,13]. The DPGDME molecules considered in this contest are of medium size so that a typical geometry optimization could be done in about 20 minutes on a single CPU. Due to the different isomers and the high degree of conformational flexibility the total amount of computer time for the conformational search and geometry optimizations sums up to more than a CPU-week. After the COSMO files for each species have been produced in this way, they can be reused

in COSMO*therm* for many different calculations. The quantum chemical calculations are thus only a one time investment for each compound. All further details and references regarding DFT/COSMO calculations with TURBOMOLE are given in reference [14].

The polarization charge density σ resulting from the DFT-COSMO calculations is a good local descriptor of the molecular surface polarity. In the second step of COSMO-RS, the statistical thermodynamics of the molecular interactions, this polarization charge density is used for the quantification of the interaction energy of pair-wise interacting surface segments with regard to the most important molecular interaction modes, i.e. electrostatics and hydrogen bonding. Thus the quantum chemical information about the polarization charge densities σ plays the key role for the evaluation of the molecular interactions in the liquid phase. As a prerequisite for an efficient statistical thermodynamics treatment, the 3D distribution of the polarization charges σ on the surface of each molecule X is converted into a surface composition function, a histogram function $p^X(\sigma)$ (called σ-profile), which describes the amount of surface with polarity σ of molecules. σ-profiles provide detailed information about the molecular polarity distribution [3-6], but they discard 3D geometric information. The σ-profiles of water and one representative of the DPGDME isomers are shown in figure 1.
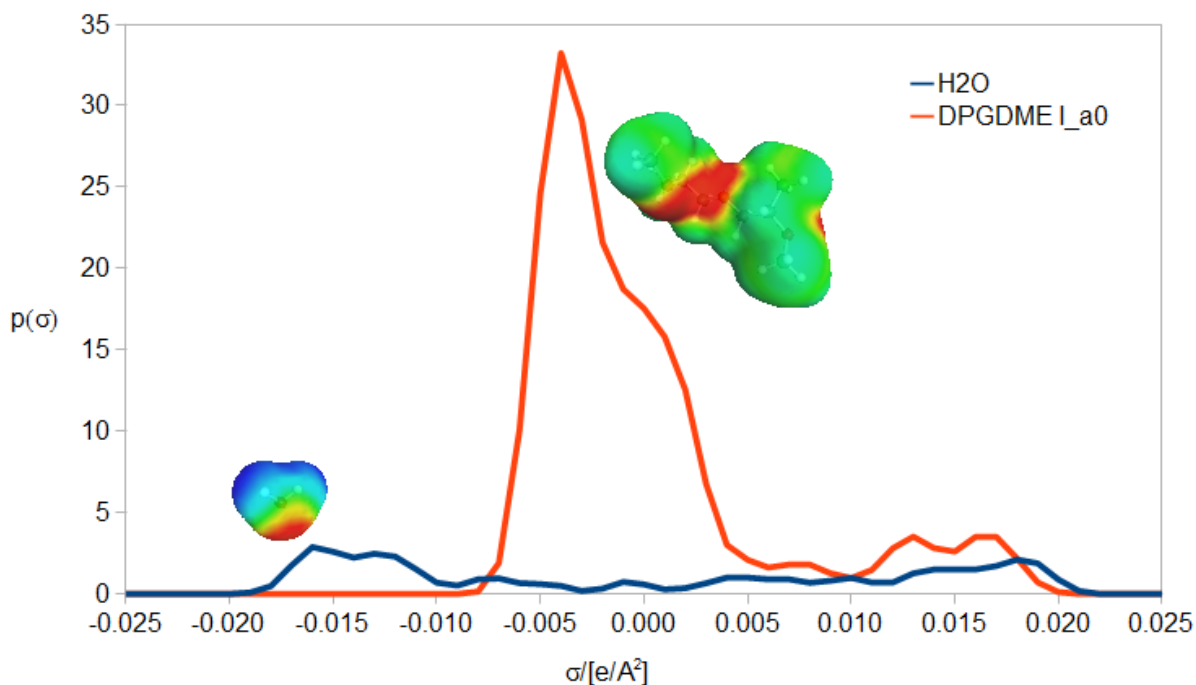

Figure 1: Illustration of the sigma profiles for water and the lowest energy conformation of the DPGDME isomer I.

Based on the surface pair interaction model introduced above, the statistical thermodynamics itself is done using a coupled set of non-linear equations for the activity coefficients of the surface segments. Finally, the chemical potentials of the compounds in a pure or mixed solvent are calculated by summation of the chemical potentials of the surface segments of a compound and slightly corrected by an empirical combinatorial term which takes into account surface areas and volumes of solutes and solvents as they result from the COSMO cavities.

All COSMO-RS calculations of this paper are performed using the COSMO*therm* program [15], which provides an efficient and flexible implementation of the COSMO-RS method. The parameterization BP_TZVP_C11_0101 is used, because it was already published in detail in previous papers [4,6]. In such parameterization the few parameters of the COSMO-RS method, i.e. two parameters for each chemical element and about 10 general parameters, were optimized based on a set of about 1000 thermodynamic data, mainly partition coefficients and vapor pressures (for details see reference [4]). Such parameterization depends on the underlying quantum chemistry level and may be considered as the equivalent of a specific force-field in MD/MC simulations, but it has much less adjusted parameters than typical force fields, since only global and element specific constants are required in COSMO-RS.

3. Results

The 6th IFPSC consists of the prediction of the mutual solubility of DPGDME and water at 283, 323, 333 and 353 K.

The mutual solubility for Proglyde DMM ™ (a commercial DPGDME product) at 298 K is given in the challenge and will be used as reference. In addition we present COSMO*therm* results with our standard procedure, i.e. without any fitting or adjustments.

3.1 Defining DPGDME: Isomers, diastereomers, conformers

DPGDME consists of three isomers, which are defined according to the challenge, see table 1.

Isomer	SMILES	No. of diastereomers	Mass-% in Proglyde TM	No. of conformers used
I	COC(C)COCC(C)OC	2	50%	20
II	COC(C)COC(C)COC	3	47%	28
III	COCC(C)OC(C)COC	2	3%	18

Table 1: DPGDME isomers, diastereomers and conformers.

As isomer III is only present at 3 mass-%, its influence on the total solubility is small. We will therefore neglect isomer III in our further considerations. Usually one would not expect isomer I+II and their corresponding diastereomers to show very different properties. However, because inverse temperature dependency and the corresponding lower critical solution temperature are crucially dependent on small differences in solution enthalpy and entropy, small differences may have major influence on the results. We thus consider all 5 diastereomers of isomer I+II to be relevant for our simulations. The relative distribution of the diastereomers is unknown and we will consider them weighted equally. DPGDME is a long flexible molecule and has many structural conformations within a relatively narrow energetic domain. We calculated different conformers with our tool COSMO*conf* [16] and limited the maximum number of used conformers to 10 per diastereomer. In our experience the error made by this limit is small compared to the overall COSMO-RS accuracy. For each diastereomer the conformers are automatically Boltzmann weighted during the COSMO*therm* calculation. In total we used 2 isomers, divided into 5 diastereomers with a total number of 48 conformers for each LLE calculation.

All molecules have been optimized using the BP functional, TZVP basis set applying the COSMO continuum solvation model.

3.2 COSMO-RS accuracy analysis and fine-tuning to experimental results

In order to estimate the typical error of water-ether systems, as a first step we calculated the infinite dilution activity coefficients for 11 water ether systems recently considered in a method comparison by Gerber and Soares [19]. For these 11 water-ether systems the root mean square error of COSMO-RS predicted infinite dilution activity coefficients is 1.1 ln-units corresponding to a factor of 3. This error is slightly larger than the typical COSMO-RS error (approx. 0.7 ln-units). Data for the miscibility gap of the Proglyde DMM ™ - water mixture at 298 K were given in the challenge as starting point for the simulations and for method calibration. Thus as a second step we analyzed the performance of COSMO-RS for these data and for some published VLE data of polyether –water mixtures. The COSMO-RS results reveal that the solubility of water in DPGDME at 298 K is predicted very well without any further adjustments. The DPGDME solubility in the water phase, however, is wrongly predicted by a factor 34 in terms of mole fraction, which seems to be a huge error. To investigate the origin of this misprediction, we analyzed the different enthalpy and entropy contributions in COSMO*therm* (see figure 2).
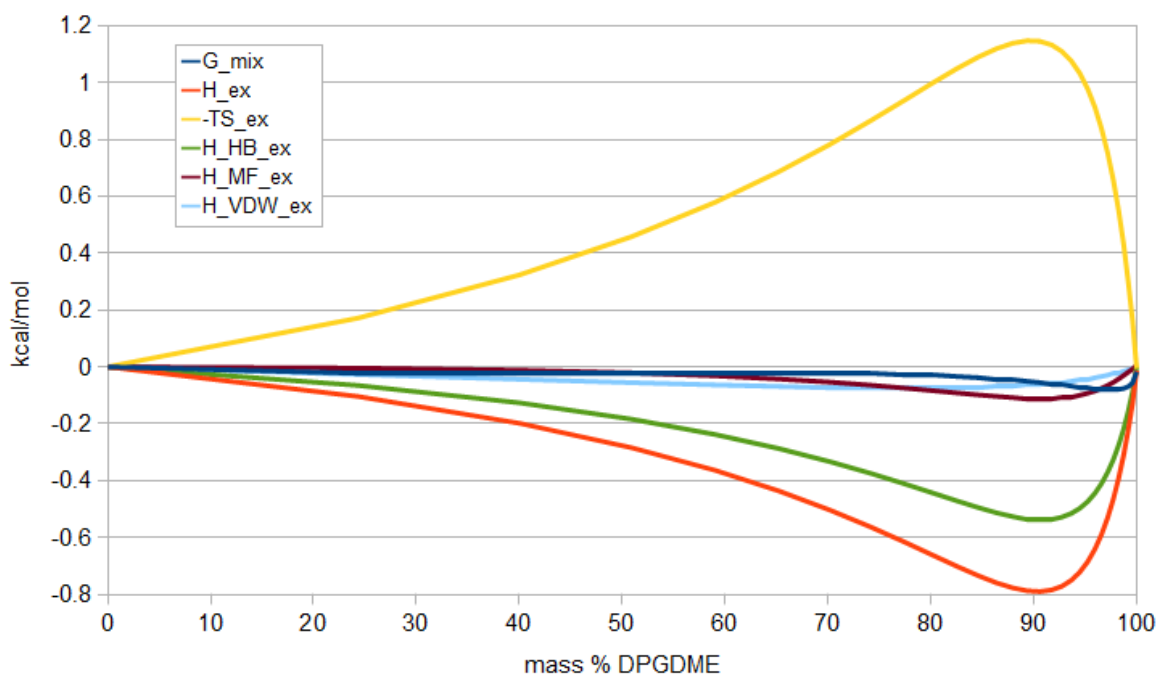
Figure 2: Energy and entropy contributions to the free energy of mixing (G_mix) at 298 K for the COSMO-RS (adj.) calculations. The trivial ideal mixing entropy is not shown to increase the clarity of the graph.

The excess enthalpy is clearly dominated by the hydrogen bonding term (H_HB_ex), while the electrostatic (H_MF_ex) and van der Waals (H_VDW_ex) contributions are of minor importance. The excess enthalpy is overcompensated by the excess entropy (-TS_ex) so that the free energy of mixing (G_mix) is very close to zero. Such almost complete compensation of entropic and enthalpic effect is very difficult to simulate, because small errors in either contribution can have large influence on the final results. The G_mix-curve is especially flat on the left hand side, leading to a very high sensitivity of the DPGDME solubility in water on small details of the parameterization and thus to the large misprediction observed when using the standard parameterization of COSMO-RS.

Compound specific fine-tuning can be achieved by rescaling of the molecular surface areas [20], which corresponds to multiplying their σ-profiles with a factor and leaving other features unchanged. Due to the non-linearity of the COSMO-RS equations, the effect of such scaling cannot be predicted in a simple fashion. The only way to see the effect is to evaluate the COSMO-RS equations in a certain system. To show that a simple adjustment to only 2 data points already yields good results, we did not use other external data sources and the general parameterizations is thus unchanged. The only free parameters were one scaling factor of the total surface area for each compound. The parameters were adjusted manually by trial and error starting from scaling

factors of unity and stopping, when the agreement of the water mole fraction was within 0.01, because such deviation is already below COSMO-RS accuracy and most likely also below the experimental measurement accuracy. In order to reproduce the experimental calibration data at 298 K, the water area was finally increased by 46% and the DPGDME area was reduced by 43%. All isomers, diastereomers and conformers have been scaled with the same factor. The corresponding results are labeled COSMO-RS (adj.).

Method	T/K	Results in mole fraction				Results in weight %			
		Phase1		Phase2		Phase 1		Phase 2	
		H2O	DPGDME	H2O	DPGDME	H2O	DPGDME	H2O	DPGDME
experimental	298	0.2981	0.7019	0.9436	0.0564	4.50	95.50	65.02	34.98
COSMO-RS	298	0.2937	0.7063	0.9983	0.0017	4.41	95.59	98.53	1.47
COSMO-RS (adj.)	298	0.2941	0.7059	0.9464	0.0536	4.42	95.58	66.21	33.79

Table 2: Experimental and COSMO-RS results for 298 K.

3.3 COSMOtherm LLE calculations

A single COSMO*therm* calculation (parameterization BP_TZVP_C11_0101) was performed for the prediction of the water-DPGDME LLE at each temperature. Each diastereomer has been entered as a single compound with fixed relative abundance to the other diastereomers at the starting point. Thus a 6 component (five diastereomers + water) liquid-liquid-phase-equilibrium was solved for each temperature. The results for combined DPGDME are presented in table 3 in comparison to the experimental data points used for the challenge evaluation. A graphical comparison of the predicted and experimental LLE can be found in figure 3.

Method	T/K	Experimental LLE results in mass % (Donate)				Predicted LLE results in mass %			
		Phase1		Phase2		Phase 1		Phase 2	
		H2O	DPGDME	H2O	DPGDME	H2O	DPGDME	H2O	DPGDME
COSMO-RS (adj.)	283	16.4	83.6	56.8	43.2	11.1	88.9	56.7	43.3
COSMO-RS (adj.)	323	4.8	95.2	80.6	19.4	1.4	98.6	77.6	22.4
COSMO-RS (adj.)	333	5.0	95.0	84.0	16.0	1.1	98.9	80.5	19.5
COSMO-RS (adj.)	353	5.0	95.0	88.0	12.0	0.8	99.2	84.5	15.5

Method	T/K					H2O	DPGDME	H2O	DPGDME
COSMO-RS	283					8.9	91.1	97.7	2.3
COSMO-RS	323					2.2	97.8	99.3	0.7
COSMO-RS	333					1.9	98.1	99.4	0.6
COSMO-RS	353					1.5	98.5	99.6	0.4

Table 3: Temperature dependent LLE points for the water DPGDME system at the temperatures requested for the challenge prediction. The experimental data have been taken from [17].
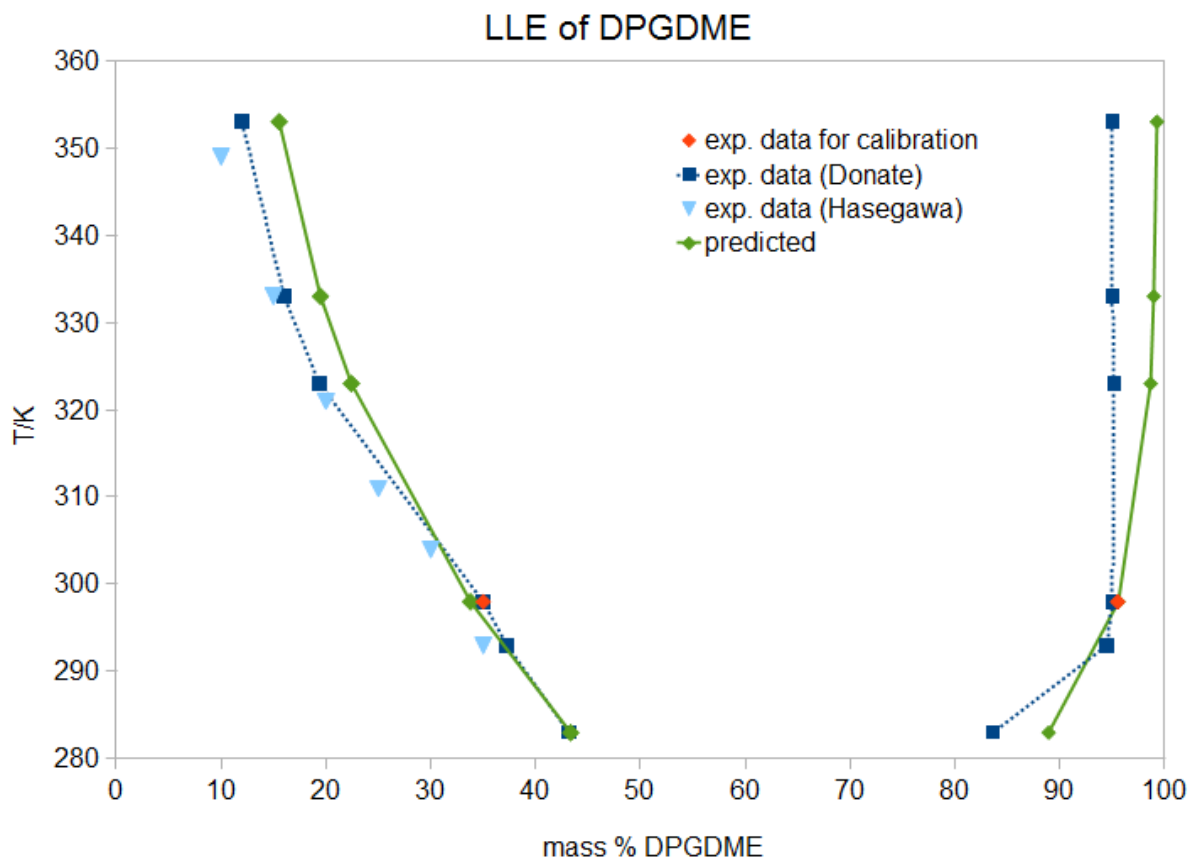
Figure 3: Comparison of the predicted (adj.) and experimental temperature dependent LLE points for the water DPGDME system. Experimental data are taken from internal measurement of Dow chemical by Donate [17]and Hasgeawa [18]

With adjusted areas the deviation of the predicted and experimental solubilities are within 5.4 mass percent points for the DPGME rich phase and below 3.6 mass percent points for the water rich phase. The inverse temperature dependence is predicted correctly.

4. Conclusions

The COSMO-RS method in its COSMO*therm* implementation has been used to predict the LLE of the water-DPGDME system at 4 different temperatures. Direct predictive application of COSMO-RS leads to a qualitatively correct, but quantitatively insufficient description of the LLE behavior in this system. Adjusting two scaling parameters to the experimental data at 298 K leads to very satisfying predictions for the LLE at higher and lower temperatures. An analysis of the different enthalpic and entropic contributions of this polyether-water mixture reveals an almost perfect compensation of large competing contributions. This delicate

balance is the reason why some parameter fitting was required for a quantitative description of the LLEs, while usually COSMO-RS can be applied to LLE prediction without special fine-tuning. Nevertheless, performing best in this IFPSC, COSMO-RS is proven to be a very efficient and reasonably accurate tool for physical property predictions even in such demanding systems.

References

[1] F. Case, A. Chaka, D.G. Friend, D. Frurip, J. Golab, R. Johnson, J. Moore, R.D. Mountain, J. Olson, M. Schiller, J. Storer, Fluid Phase Equilibria **217** (2004) 1-10

[2] A. Klamt, F. Eckert, Fluid Phase Equilibria **217** (2004) 53–57

[3] A. Klamt, *J. Phys. Chem.* **99** (1995) 2224

[4] A. Klamt, V. Jonas, T. Buerger, J.C.W. Lohrenz, *J. Phys. Chem. A* **102** (1998) 5074

[5] A. Klamt, F. Eckert, W. Arlt, Annual Review Chem. Biomol. Eng. **1** (2010) 101-122

[6] A. Klamt, *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier, Amsterdam, (2005)

[7] A. Klamt, G. Schüürmann, *J. Chem. Soc. Perkin Trans. 2* (1993) 799

[8] A. D. Becke, *Phys. Rev. A* **38** (1988) 3098

[9] J. P. Perdew, *Phys. Rev. B* **33** (1986) 8822

[10] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, *Chem. Phys. Letters* **242** (1995) 652

[11] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, *Theor. Chem. Acc.* **97 (**1997) 119

[12] R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chem. Phys. Letters* **162** (1989) 165

[13] Turbomole, Version 6.20, TURBOMOLE GmbH, Germany, 2010.

[14] A. Schäfer, A. Klamt, D. Sattel, J.C.W. Lohrenz, F. Eckert, *Phys. Chem. Chem. Phys.* **2** (2000) 2187

[15] COSMO*therm* program, Release C2.1, Revision 01.07, COSMO*logic* GmbH&CoKG, Leverkusen, Germany (2007)

[16] A. Klamt, F. Eckert, M.J.Diedenhofen (2009) J. Phys. Chem. B **113** (2009) 4508–4510

[17] Internal report of Dow chemical, measurements by Donate et. al.

[18] Internal report of Dow chemical, measurements by Hasegawa et. al.

[19] R.P. Gerber and R. de P. Soares, *Ind. Eng. Chem. Res.*, Article ASAP, **DOI:** 10.1021/ie901947m

[20] A. Klamt, Fluid Phase Equilibria **206** (2003) 223-235