

Prediction of aqueous solubilities of solid carboxylic acids with COSMO-RS

Bernd Schröder^{a,*}, Luís M.N.B.F. Santos^b, Isabel M. Marrucho^a, João A.P. Coutinho^a

^a CICECO, Departamento de Química, Universidade de Aveiro, 3810-193 Aveiro, Portugal

^b Centro de Investigação em Química, Faculdade de Ciências da Universidade do Porto, R. Campo Alegre 687, P-4169-007 Porto, Portugal

ARTICLE INFO

Article history:

Received 16 October 2009

Received in revised form

11 November 2009

Accepted 16 November 2009

Available online 23 December 2009

Keywords:

Organic acids

Aqueous solubilities

Carboxylic acids

COSMO-RS

COSMOtherm

Enthalpies of solution

ABSTRACT

The methodology for the prediction of aqueous solubilities of solid organic compounds, and their temperature dependence, based on the Conductor-like Screening Model for Real Solvents (COSMO-RS/COSMOtherm) procedure, is presented and evaluated. The predictive capability of the quantum chemistry based program and the applied methodology was tested on the most common solid carboxylic acids. From the temperature dependence of the solubilities, the mean apparent enthalpies of solution were derived. The results obtained for a set of 27 carboxylic acids, consisting of aromatic carboxylic acids, dicarboxylic acids, as well as hydroxycarboxylic acids, are in good agreement with the experimental solubility data and their dependence with the temperature. The mean apparent enthalpies of solution, although of the same order of magnitude of the experimentally values, seem to be systematically underestimated.

The good response of the applied methodology based on COSMOtherm/COSMO-RS on the prediction of non-ideal aqueous solubilities demonstrate the capability of this approach to predict and distinguish the solubility trends and magnitude, resulting from the substituent effects on the non-ideality of the aqueous solution.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The aqueous solubility of solid organic compounds is a fundamental thermodynamic property with importance reaching from the design and optimization of industrial processes in the chemical and pharmaceutical industry to the description of transport phenomena in environmental chemistry. Due to the considerable experimental efforts and the requirement of a substantial amount of high purity material, most temperature dependent solubility measurements are performed only for selected compounds. Therefore, the availability of reliable straightforward prediction methods is of foremost importance. Tools that can quickly estimate the solubility as a function of temperature and solvent composition are considered as being crucial in the engineering practice of today.

A variety of different models for aqueous solubilities has been described to date. Empirical correlations apart, the more sound thermodynamic approach is based on the solid–liquid equilibrium equation (Eq. (1))—describing the liquid phase non-ideality by activity-coefficient models such as the Wilson equation [1], NRTL [2] or UNIQUAC [3]. However, these are models which have no predictive capability to new systems. For that purpose, other approaches must be used such as solubility parameters [4] or group

contribution methods, requiring a strong data base on constituting groups, such as UNIFAC [5]. These methods have proved successful for many compounds providing acceptable values for temperature dependent solubilities, but they fail for molecules with several functional groups or systems where the solute–solvent interactions are complex. In the case of molecules where properties of a functional group depend on the remains of the molecule, properties become non-additive, rescinding the basic requirement for the validity of an additive model [6].

In what concerns the systems studied in this work, association equation-of-state models for the description of phase equilibria, like the cubic plus association equation of state (CPA EoS), have shown that they can accurately correlate the mutual solubilities of fatty acids and water, but require a high degree of parameterization [7].

Over the last years, the quantum chemically based Conductor-like Screening Model for Realistic Solvation (COSMO-RS), a novel approach for computing equilibrium thermodynamic properties, has proven to be a promising alternative when compared with group contribution methods, since it is based on a small number of adjustable parameters only, which are independent of molecular or structural information [8]. COSMO-RS is a surface interaction model, and contrary to group contribution methods, it considers molecular charge densities rather than the interaction of groups, which are obtained by molecular quantum chemical COSMO calculations. A key feature of COSMO-RS is the reduction of molecular

* Corresponding author. Tel.: +351 234 370200; fax: +351 234 370084.
E-mail address: bschrode@fc.up.pt (B. Schröder).

properties to a probability distribution of screening charges for the molecule solvated in a perfect conductor, the so-called sigma profile of the molecule. The sigma profiles are used to estimate the exchange energies between the substrate and the solvent. The calculation of the sigma profiles of the components is sensitive to the molecular configuration which is optimized in the ideal gas phase [9].

Carboxylic acids are widely used as intermediates in large-scale preparations of resins, plasticizers, dyes, inks and adhesives, pharmaceuticals as well as in the food industry [10–14]. They are also of interest in environmental chemistry, as they might occur as oxidation residues in the polyester industry, or they appear as metabolites. Carboxylic acids are also of biological importance and represent a particularly useful series of model compounds for studying hydrogen bonding and hydrophobic interactions [15–18]. As a consequence of their widespread use and importance, aqueous solubilities of carboxylic acids as a function of temperature might be considered being extensively measured, recommending them as test compounds for predictive methods, and yet, at times, discrepancies in published values, concerning accuracy and reproducibility of the gathered data, may occur.

The goal of the present study is to evaluate the COSMO-RS potential for the prediction of temperature dependence of solid–liquid equilibria (SLE), taking the aqueous solubility of 27 solid carboxylic acids as example. The chosen set of carboxylic acids consists of aromatic carboxylic acids, hydroxycarboxylic acids and dicarboxylic acids, with aqueous solubilities ranging from very low to highly soluble.

2. Methodology

2.1. Solid–liquid equilibrium thermodynamics of carboxylic acids

In a two-component solid–liquid equilibrium, the fugacity of the pure solid solute (defined as component 2) equals its fugacity in the solvent (defined as component 1), assuming complete immiscibility with the solvent in the solid state. Defining the standard-state fugacity as the fugacity of pure, subcooled liquid at a solution temperature T , an expression for the activity of the solute, valid at low pressures, can be derived if a single solid–solid-phase transition is assumed [4]:

$$\ln a_2 = \ln(\gamma_2 x_2) = \frac{\Delta_{\text{fus}}H(T_{\text{fus}})}{RT_{\text{fus}}} \left(\frac{T - T_{\text{fus}}}{T} \right) + \frac{\Delta_{\text{trs}}H(T_{\text{trs}})}{RT_{\text{trs}}} \left(\frac{T - T_{\text{trs}}}{T} \right) - \frac{\Delta_{\text{fus}}C_{p,m}(T_{\text{fus}})}{R} \left[\left(\frac{T - T_{\text{fus}}}{T} \right) + \ln \frac{T_{\text{fus}}}{T} \right] \quad (1)$$

with a_2 being the activity of the solute, γ_2 as the activity coefficient of the solute, x_2 its mole fraction, $\Delta_{\text{fus}}H(T_{\text{fus}})$ as the molar enthalpy of fusion of the pure solute, at the fusion temperature T_{fus} , $\Delta_{\text{trs}}H(T_{\text{trs}})$ the molar solid–solid-phase transition enthalpy, at the transition temperature T_{trs} , and $\Delta_{\text{fus}}C_{p,m}(T_{\text{fus}})$ as the molar heat capacity change upon fusion, at fusion temperature T_{fus} [4]. By adding up the respective terms, Eq. (1) can be adapted to situations where solid-phase transitions are occurring.

As a standard simplification throughout the chemical engineering literature the heat capacity change upon fusion is regarded as being independent of the temperature. The bracketed term multiplied with $\Delta_{\text{fus}}C_{p,m}$ is often considered as being small, as the opposite signs inside the bracket lead to near cancellation [19–21]. Nevertheless, as Prausnitz et al. [4] pointed out, the contribution of $\Delta_{\text{fus}}C_{p,m}$ may become large at temperatures far off the solute melting point. Neglecting this term might then introduce a considerable error into the solubility estimation [22].

Another approximation assumes the molar entropy of fusion $\Delta_{\text{fus}}S_m$ as being an estimate of $\Delta_{\text{fus}}C_{p,m}$ [23]. By implication, sol-

ubility data might be used together with fusion temperatures and enthalpies of fusion to determine $\Delta_{\text{fus}}C_{p,m}(T_{\text{fus}})$ [24].

Various phenomena might influence the reliability of the measured melting point parameter and the validity of assumptions on solubility data of carboxylic acids. Processes like decomposition, incipient sublimation prior to fusion or evaporation upon fusion can impede accurate melting property measurements. Melting points are occasionally inaccessible due to the occurrence of modifications causing phase conversions or degradation at high temperatures [25,26]. Different approaches to predict fusion temperatures, based on group contribution methods or quantitative activity–property relationships, are applied in the thermodynamic practice, but nevertheless, this is one of the thermophysical properties most difficult to estimate.

Concerning aqueous solubilities of carboxylic acids, a variety of compound-specific features have to be taken into account. For instance, carboxylic acids are known to form dimers in the gas phase and in non-polar solvents, whereas in aqueous solution, at low concentrations the intermolecular association of carboxylic acid molecules is replaced by hydrogen bonding to water [27]. Dimer formation in aqueous solution is subject of ongoing discussion [28]. At concentrations sufficiently low, even very weak electrolytes ($pK_a \sim 10$) dissociate to a certain extent. Extrapolation to activity coefficients of the solute at infinite dilution might be seriously flawed. However, in practical terms, the effect of dissociation of carboxylic acids on aqueous solubility may be considered negligible and, in general, they can be treated as non-ionized solutes, as for instance the dicarboxylic acids [29–31]. Therefore, the effect of dissociation is neglected in this work and the carboxylic acids considered here are treated as non-electrolytes.

2.2. Apparent molar enthalpies of solution

The temperature dependence of solubility is related to the molar enthalpy of solution:

$$\left\{ \frac{\partial \ln x}{\partial (1/T)} \right\} \left\{ 1 + \left(\frac{\partial \ln \gamma}{\partial \ln x} \right)_T \right\} = - \frac{\Delta_{\text{sol}}H_m}{R} \quad (2)$$

Since the values of the activity coefficients are unknown this term is neglected and only the apparent or *van't Hoff* molar enthalpies of solution are accessible [26]:

$$\left(\frac{\partial \ln x_2}{\partial (1/T)} \right) = - \frac{\Delta_{\text{sol}}^{\text{app}}H_m}{R} \quad (3)$$

High values of apparent molar enthalpies of solution indicate a strong temperature dependence of the respective solubilities. Deviations from a straight line of the plot ($\ln x_2$) vs. $(1/T)$ may occur due to the fact that the apparent molar enthalpy of solution itself is temperature dependent. Non-ideal behaviour of the solution and a change in composition of the solid phase in equilibrium with the saturated solution may contribute to this deviation as well. In some cases, distinct breaks in solubility curves may be attributed to transitions between different hydrated species [32]. Larger deviations from calorimetric enthalpies of solution can be caused by the impact of enthalpies of dilution, or from activity coefficient effects in Eq. (2), depending on the respective substance.

For an ideal solution, the apparent molar enthalpy of solution $\Delta_{\text{sol}}^{\text{app}}H_m$ (ideal) equals the enthalpy of fusion of the pure solid at the temperature T .

The regular solution assumption is another common approximation [4]:

$$\left(\frac{d \ln \gamma}{dT} \right)_x = - \frac{\ln \gamma_{\text{eq}}}{T} \quad (4)$$

Eq. (4) describes the change of the activity coefficient with temperature at constant concentration for regular solutions. But solutions in which the molecular volume differs between the solute and the solvent may cause deviations, as can be expected, for instance, in the case of weak acids. In general, solubility data in organic solvents show less deviation, when compared with aqueous solubility data, except at very low solubility concentrations [26].

2.3. COSMO-RS

COSMO-RS is a novel method for predicting thermodynamic properties of pure and mixed liquids on the basis of unimolecular quantum chemical calculations for the respective individual molecules. COSMO-RS combines the computational efficiency of the quantum chemical dielectric continuum solvation model COSMO with a statistical thermodynamics approach, based on the results of quantum chemical calculations [8,33,34]. The COSMO model is a variant of known dielectric continuum solvation methods. The solute molecules are assumed to be placed in a virtual conductor environment, where the solute induces a polarization charge density on the molecular surface, the interface between the molecule and the conductor. These charges act back on the solute and give rise to a more polarized electron density than in vacuum. During the quantum chemical self-consistency cycle, the solute molecule is converged to its energetically optimal state in a conductor, with respect to electron density.

The standard procedure of COSMO-RS calculations consists of two main steps: quantum chemical COSMO calculations for the molecular species involved and COSMO-RS statistical calculations performed within the COSMOtherm program [35,36]. The COSMO-RS calculation with COSMOtherm (Version C2.1 Release 01.05) describes all interactions between molecules as contact interactions of the molecular surfaces, and these interactions can be related to the screening charge densities of the interacting surface parts. The COSMO output provides the total energy of a molecule in its conductor environment and the 3D polarization density distribution on the surface of each molecule—the sigma profile. This information acts as an input for the statistical thermodynamic calculations and is independent of the solvent dielectric constant and temperature. Furthermore, COSMO-RS performs the statistical thermodynamic calculations and hence yields enthalpic and entropic information for each system. The COSMO-RS method depends only on a small number of general, element-specific adjustable parameters (predetermined from known properties of a small set of molecules), that are not specific for functional groups or types of molecules.

The calculation procedure and basis set parameterization of COSMO-RS/COSMOtherm calculations were performed using the respective molecular description file derived at the BP/TZVP level of theory (Turbomole [37,38], DFT/COSMO calculation with the BP density functional theory and the triple- ζ -valence polarized large basis set (TZVP) [39] basis set with fully optimized geometries at the same level of theory), and the parameter file BP TZVP C21 0025.

The feasibility of COSMO-RS/COSMOtherm's ability to predict phase equilibria was successfully demonstrated with the satisfactory quantitative prediction of mutual solubilities of hydrocarbons and water [40], the prediction of LLE and VLE of water and ionic liquids binary systems [41–44], as well as on SLE in different solvents [9,45].

Since COSMO-RS can process liquids only, it will treat solids as subcooled melts, and therefore, the Gibbs energies of fusion, $\Delta_{\text{fus}}G$, have to be known and provided in the input section [9]. Certain approximations have to be introduced, since the properties of subcooled melts far below the fusion temperature are difficult to quantify. The applied methodology is schematically depicted in Fig. 1.

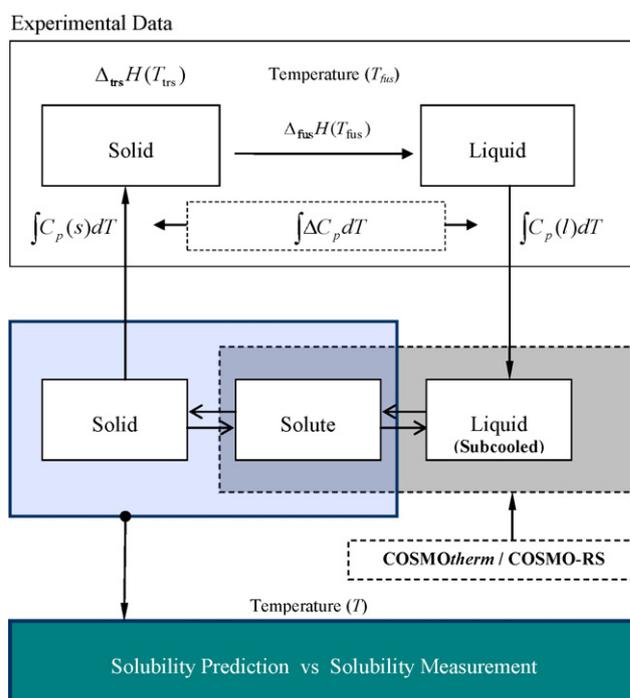


Fig. 1. Scheme of the hypothetical thermodynamic cycle used in the applied methodology.

The mole fractions of the respective solute were refined using the iterative automatic solubility calculation option of COSMO-RS/COSMOtherm:

$$\log_{10}(x_2^{\text{SOL}(i+1)}) = \frac{[\mu_2^{(P)} - \mu_2^{(1)}(x_2^{\text{SOL}(i)}) + \min(0, \Delta_{\text{fus}}G)]}{RT \ln 10} \quad (5)$$

where x_2^{SOL} is the mole fraction of solid 2 dissolved in the solvent phase 1 at saturation, $\mu_2^{(P)}$ is the chemical potential of pure compound 2, $\mu_2^{(1)}$ is the chemical potential of compound 2 at infinite dilution in the solvent compound 1. Existing conformers are weighted internally by COSMO-RS/COSMOtherm, using their respective COSMO-energy and their chemical potential, and the resulting solubility is averaged according to the Boltzmann distribution of the conformers. Low-energy conformers have been taken into account when relevant. This is of special importance for several carboxylic acids, which can form complex intramolecular hydrogen bond patterns in the neutral form. The sequence of relative stability of the different conformers found in quantum chemical calculations for the gaseous phase is not necessarily the same as in solution.

The solubility of selected carboxylic acids in aqueous solution, x_2 , has been calculated with COSMO-RS/COSMOtherm, in the temperature interval from 278.15 K to 368.15 K.

2.4. Compounds and literature data

Carboxylic acids have been extensively characterized in the past, both in terms of thermal melting parameters and aqueous solubilities, and the data retrieved from literature have been subject of critical analysis here. Table 1 lists the carboxylic acids chosen, as well as their respective melting properties literature data [$\Delta_{\text{fus}}H(T_{\text{fus}})$ —enthalpies of fusion at fusion temperature, T_{fus} —fusion temperature and $\Delta_{\text{fus}}C_{p,m}$ —heat capacity change upon fusion] and the temperature interval of the available aqueous solubility data. $\Delta_{\text{tpce}}^1 H_m(T_{\text{fus}})$ – the total phase change enthalpy at fusion temperature – includes phase transitions prior to melting for those compounds, whose enthalpies of phase transitions are known.

Table 1

Carboxylic acid thermophysical properties data ($\Delta_{\text{fus}}H(T_{\text{fus}})$)—enthalpies of fusion at fusion temperature, T_{fus} —fusion temperature and $\Delta_{\text{fus}}C_{p,m}$ —heat capacity change upon fusion), and the temperature range of aqueous solubility data used and its source.

	CAS registry number	$\Delta_{\text{fus}}H(T_{\text{fus}})$ (kJ mol ⁻¹)	T_{fus} (K)	$\Delta_{\text{fus}}C_{p,m}$ (JK ⁻¹ mol ⁻¹)	Temperature range (K)
<i>Aromatic carboxylic acids</i>					
Benzoic acid	65-85-0	18.07 [47]	395.52 [47]	60.18 [47]	281.15–337.15 [48]
p-toluic acid, 4-methylbenzoic acid	99-94-5	22.72 [47]	452.75 [47]	58.64 [46]	278.15–343.15 [49]
Salicylic acid, 2-hydroxybenzoic acid	69-72-7	27.09 ± 0.02 [50]	431.35 ± 0.2 [50]	~60	283.15–339.15 [51]
Acetylsalicylic acid, 2-acetoxybenzoic acid	50-78-2	29.80 ± 0.2 [52]	414.0 [52]	90.81 [52]	278.15–345.15 [49]
p-hydroxybenzoic acid	99-96-7	31.4 ± 0.8 [53]	487.15 [53]	63.1 [53]	278.15–345.15 [54]
m-hydroxybenzoic acid	99-06-9	35.92 [55]	474.8 [55]	~60	283.15–323.15 [55]
p-hydroxyphenylacetic acid,	156-38-7	28.0 ± 0.7 [53]	422.85 ± 0.7 [53]	59.7 [53]	283.15–298.15 [53]
2-Furancarboxylic acid	88-14-2	22.6 ± 0.3 [53]	402.5 ± 0.2 [56]	~60	279.15–341.15 [51]
Terephthalic acid,	100-21-0	63.428 [57]	702.2 [57]	100 [46]	278.15–331.65 [48]
benzene-1,4-dicarboxylic acid					
o-phthalic acid,	88-99-3	36.5 ± 1.2 [58]	463.45 ± 0.01 [58]	~100	283.15–338.15 [51]
benzene-1,2-dicarboxylic acid					
Isophthalic Acid,	121-91-5	43.2 ± 1.2 [58]	617.41 ± 0.04 [58]	~100	281.15–337.15 [48]
benzene-1,3-dicarboxylic acid					
Nicotinic acid, pyridine-3-carboxylic acid	59-67-6	27.9 [59]	510.45 [59]	~60	274.15–361.15 [60]
3-nitrobenzoic acid	121-92-6	21.4 [61]	414.15 [62]	~60	278.15–343.15 [63]
<i>Dicarboxylic acids</i>					
Oxalic acid, etanedioic acid	144-62-7	58.158 [64]	465.26 [64]	~50	278.15–338.15 [32]
Malonic acid, propanedioic acid	141-82-2	25.48 [47]	407.95 [47]	~60	278.15–338.15 [32]
Succinic acid, butanedioic acid	110-15-6	34.0 ± 0.3 [65]	455.2 ± 0.2 [65]	69.6 [65]	278.15–338.15 [32]
Glutaric acid, pentanedioic acid	110-94-1	21.1 ^a [65]	363.9 ± 6.4 [65]	83.6 [65]	279.15–342.15 [51]
Adipic acid, hexanedioic acid	124-04-9	33.7 ± 0.4 [65]	419.0 ± 1.4 [65]	88.6 [65]	278.15–338.15 [32]
Pimelic acid, heptanedioic acid	111-16-0	25.2 ^a [65]	368.2 ± 1.6 [65]	93.6 [65]	279.15–342.15 [51]
Suberic acid, octanedioic acid	505-48-6	41.8 ^a [65]	413.2 ± 0.2 [65]	98.6 [65]	280.15–349.65 [66]
Azelaic acid, nonanedioic acid	123-99-9	30.4 ^a [65]	372.4 ± 1.2 [65]	103.6 [65]	280.15–357.65 [66]
<i>Hydroxycarboxylic acids</i>					
Ascorbic acid,	50-81-7	29.2 [47]	465.15 [47]	~60	280.15–338.15 [51]
(R)-3,4-dihydroxy-5-((S)-1,2-dihydroxyethyl)furan-2(5H)-one					
Citric acid, 2-hydroxypropane-1,2,3-tricarboxylic acid	77-92-9	26.7 [47]	426.15 [47]	~70	278.15–338.15 [32]
Tartaric acid,	526-83-0	30.1 [47]	479.15 [47]	~70	278.15–338.15 [32]
2,3-dihydroxybutanedioic acid					
Levulinic acid, 4-oxopentanoic acid	123-76-2	9.22 [67]	306.15 [67]	~60	280.15–303.25 [66]
Malic acid, hydroxybutanedioic acid	6915-15-7	25.3 [47]	403.15 [47]	~60	278.15–338.15 [32]
Diglycolic acid,	110-99-6	26.4 [47]	421.15 [47]	~60	278.25–361.35 [66]
2-(carboxymethoxy)acetic acid					

^a $\Delta_{\text{tpce}}^{\text{I}}H_m(T_{\text{fus}})$: total phase change enthalpy at fusion temperature, includes phase transitions prior to melting.

Experimental data for the heat capacity change upon fusion are difficult to access and therefore, a certain scarcity of data exists in the literature. Also, its estimation with group contribution methods is not always reliable, as Domalski and Hearing stated for the case of dibasic carboxylic acids [46]. Thus, the heat capacity changes upon fusion considered here are assumed as temperature independent. For carboxylic acids with no information available at all, estimated values based on comparison with similar molecules were considered.

3. Results and discussion

3.1. Aqueous solubilities

The temperature dependence of the aqueous solubilities of 27 carboxylic acids was predicted with COSMO-RS/COSMOtherm and compared with available literature data. For all solutes, melting properties data are found in the literature. The apparent molar enthalpies of solution have been estimated for each substance, using the literature solubility data, and compared with those predicted with COSMO-RS/COSMOtherm, at $T = 298.15$ K, as reported in Table 2.

The influence of the quality of available thermodynamic melting properties (molar enthalpy of fusion at fusion temperature, fusion temperature, and molar heat capacity change upon fusion) is evaluated taking salicylic acid as example, and is depicted in Fig. 2. The further distant the melting temperature is from 298.15 K, the larger the deviations of the calculated solubilities from the experimental values become, while approaching lower temperatures, with emphasis on accuracy of the enthalpies of fusion, while at higher temperatures, errors in the temperature of fusion become more influential on the results.

3.1.1. Aromatic carboxylic acids

The aqueous solubilities of aromatic carboxylic acids are reasonably well predicted by COSMO-RS/COSMOtherm, as can be seen in the deviation plot shown in Fig. 3. The description of the experimental data is very satisfactory for a pure predictive model. Data for hardly soluble aromatic carboxylic acids like p-toluic acid are adequately reproduced. Only the solubility data of isophthalic acid deviate considerably, when compared with the results for the other aromatic carboxylic acids. These deviations may result from the quality of the reported thermodynamic properties at the melting point. The molar heat capacity difference between the solid and liquid of isophthalic acid might be overestimated: considering a value

Table 2
Apparent molar enthalpies of solution, $\Delta_{\text{sol}}^{\text{app}}H_m = -R(\partial \ln x_2 / \partial (1/T))$, at $T=298.15$ K, as estimated from experimental aqueous solubility data and COSMO-RS solubility predictions, as well as their deviations [$\Delta_{\text{sol}}^{\text{app}}H_m(\text{exp}) - \Delta_{\text{sol}}^{\text{app}}H_m(\text{COSMO})$].

Carboxylic acid	$\Delta_{\text{sol}}^{\text{app}}H_m$ (298.15 K; exp) (kJ mol ⁻¹)	$\Delta_{\text{sol}}^{\text{app}}H_m$ (298.15 K; COSMO) (kJ mol ⁻¹)	Deviation (exp – COSMO-RS) (kJ mol ⁻¹)
<i>Aromatic carboxylic acids</i>			
Benzoic acid	25.57	11.73	13.84
p-Toluic acid	24.12	12.30	11.82
Salicylic acid	23.81	15.76	8.05
Acetylsalicylic acid	22.93	14.27	8.66
p-Hydroxybenzoic acid ¹	41.88	10.50	31.38
m-Hydroxybenzoic acid	31.85	15.27	16.58
p-Hydroxyphenylacetic acid	39.46	16.78	22.68
2-Furancarboxylic acid	39.12	16.42	22.7
Nicotinic acid	14.02	12.89	1.13
3-Nitrobenzoic acid	28.30	11.05	17.25
o-Phthalic acid	34.95	10.08	24.87
Terephthalic acid	-7.17	8.85	-16.02
Isophthalic Acid	33.82	-2.78	36.6
<i>Dicarboxylic acids</i>			
Oxalic acid ²	26.37	13.79	12.58
Malonic acid	9.36	15.32	-5.96
Succinic acid	29.34	10.68	18.66
Glutaric acid	26.31	10.44	15.87
Adipic acid	36.18	12.12	24.06
Pimelic acid	41.86	28.86	13
Suberic acid	34.91	16.08	18.83
Azelaic acid	38.80	9.73	29.07
<i>Hydroxycarboxylic acids</i>			
Ascorbic acid	21.46	11.21	10.25
Citric acid ³	27.21	12.04	15.17
Tartaric acid	6.58	1.52	5.06
Levulinic acid	26.81	5.77	21.04
Malic acid	11.84	9.12	2.72
Diglycolic acid	22.46	12.12	10.34

According to reported equilibrium states of the solid acids at the interface solute–solvent as hydrated molecules, modified equations to calculate the apparent molar enthalpies of solution have been used: 1–p-hydroxybenzoic acid as monohydrate [72]; $\Delta_{\text{sol}}^{\text{app}}H_m = -R(1 - hmM_1) \frac{\partial \ln m}{\partial (1/T)}$; with $h=1$, $M_1=18.052$ g mol⁻¹ [49]; $m=0.0489$ mol kg⁻¹, $\partial \ln m / \partial (1/T) = -5037$ mol kg⁻¹ K (experimental) and $m=0.2155$ mol kg⁻¹, $\partial \ln m / \partial (1/T) = -1263$ mol kg⁻¹ K (derived from COSMO-RS data); 2–oxalic acid as dihydrate; $\Delta_{\text{sol}}^{\text{app}}H_m = -R(1 - hmM_1) \frac{\partial \ln m}{\partial (1/T)}$; with $h=1$, $M_1=18.052$ g mol⁻¹ [32]; $m=1.146$ mol kg⁻¹, $\partial \ln m / \partial (1/T) = -3308$ mol kg⁻¹ K (experimental) and $m=0.699$ mol kg⁻¹, $\partial \ln m / \partial (1/T) = -1702$ mol kg⁻¹ K (derived from COSMO-RS data); 3–citric acid as monohydrate; $\Delta_{\text{sol}}^{\text{app}}H_m = -R(1 - hmM_1) \left\{ \frac{\Phi}{m} + \left(\frac{\partial \Phi}{\partial m} \right)_T \right\} \frac{\partial m}{\partial (1/T)}$; with $\Phi=1.629$, $(\Delta \Phi / \Delta m) = 0.0598$ kg mol⁻¹, $h=1$, $M_1=18.052$ g mol⁻¹ [32] and references therein; $m=8.456$ mol kg⁻¹, $\Delta m / (\Delta (1/T)) = -15,300$ mol kg⁻¹ K (experimental) and $m=11.540$ mol kg⁻¹, $\Delta m / (\Delta (1/T)) = -9104$ mol kg⁻¹ K (derived from COSMO-RS data).

of 50 J K⁻¹ mol⁻¹ for this thermodynamic function leads to predicted aqueous solubilities considerably closer to the experimental ones, with a highest deviation of ca. one log-unit at $T=278.15$ K, approximating rapidly the experimental findings at high temperatures (0.16 log-units at $T=368.15$ K). Decomposition during melting is also a known feature of phthalic acid isomers [58].

Experimental observations of solubilities, reporting a somewhat unexpected high aqueous solubility of p-hydroxyphenylacetic acid, when compared to p-hydroxybenzoic acid [49], were not found with COSMOtherm.

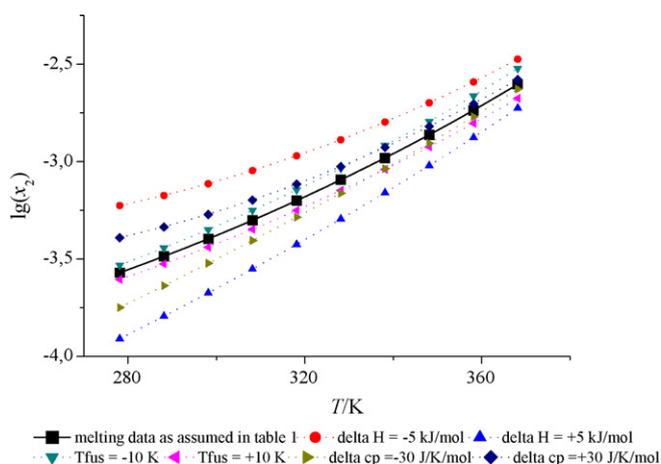


Fig. 2. Influence of melting properties on solubility predictions with COSMO-RS/COSMOtherm for salicylic acid.

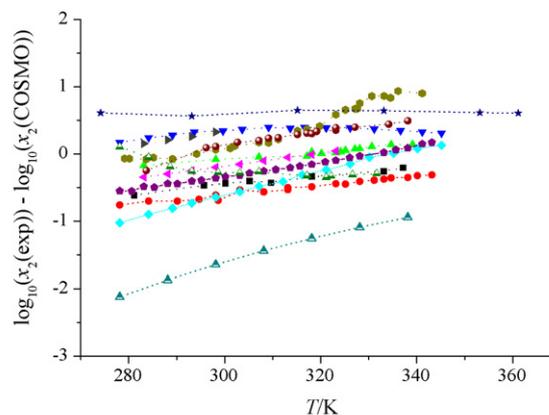


Fig. 3. Difference in solubilities between literature data and COSMO-RS/COSMOtherm, predictions for aromatic carboxylic acids.

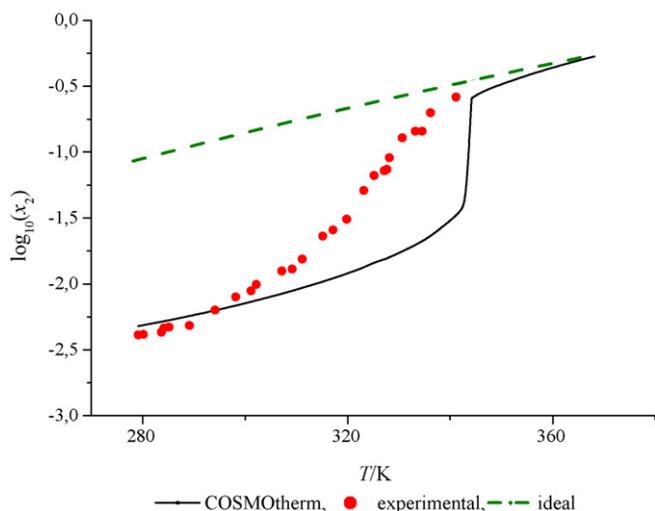


Fig. 4. 2-Furancarboxylic acid aqueous solubility as observed experimentally [51] and predicted by COSMO-RS/COSMOtherm.

In the experimental set of carboxylic acids examined, the usual COSMO-RS/COSMOtherm response displays a monotonous temperature dependence of aqueous solubility. Nevertheless, four compounds show a deviation from strict monotony in the high temperature region: 2-furancarboxylic acid, p-hydroxyphenylacetic acid, adipic acid, and azelaic acid. For the last three, there are no available experimental data for this high temperature region. The experimental and COSMO-RS/COSMOtherm values of aqueous solubilities of 2-furancarboxylic acid are shown in Fig. 4. Remarkably, COSMO-RS/COSMOtherm reproduces the experimentally observed deviation from monotony rather well. At higher temperatures, a leap in the amount of 2-furancarboxylic acid getting dissolved is noticed, with the system approaching ideal solution behaviour. The change from one solubility regimen into another might be related to changes in the intermolecular interactions between 2-furancarboxylic acid molecules in solution as the solution becomes diluted and the solvation of the acid molecules by water becomes dominant with fewer acid–acid interactions taking place.

3.1.2. Dicarboxylic acids

The aqueous solubilities of dicarboxylic acids are satisfactorily reproduced by COSMO-RS, as can be seen in the deviation plot presented in Fig. 5. For several alkylic dicarboxylic acids, phase transitions other than melting are known. Since COSMO-RS/COSMOtherm is not handling multiple phase transitions explicitly, the total phase change enthalpy at fusion temperature was assumed when available. The biggest difference between the molar enthalpy of fusion and the total phase change enthalpy is registered for suberic acid, with $(\Delta_{\text{TPCE}}^1 H_m(T_{\text{fus}}) - \Delta_{\text{fus}} H(T_{\text{fus}})) = 11.1 \text{ kJ mol}^{-1}$ [65]; nevertheless, the obtained solubility prediction is acceptable.

Depending on the molecule's nature, consideration of a variety of conformers is an important issue. For oxalic acid, for instance, five conformers have been considered [68–70]. Among them, the cTF conformer [68] is reported to form the most stable structure in clusters with two water molecules, altering the relative stability sequence of different conformers, in comparison with those found in quantum chemical studies for the gaseous phase [71].

Dicarboxylic acids with longer alkyl chains, like suberic and azelaic acid, exhibit significant lower solubilities than the other acids considered before; therefore, experimental determination of reliable aqueous solubility data is considered as being difficult [66].

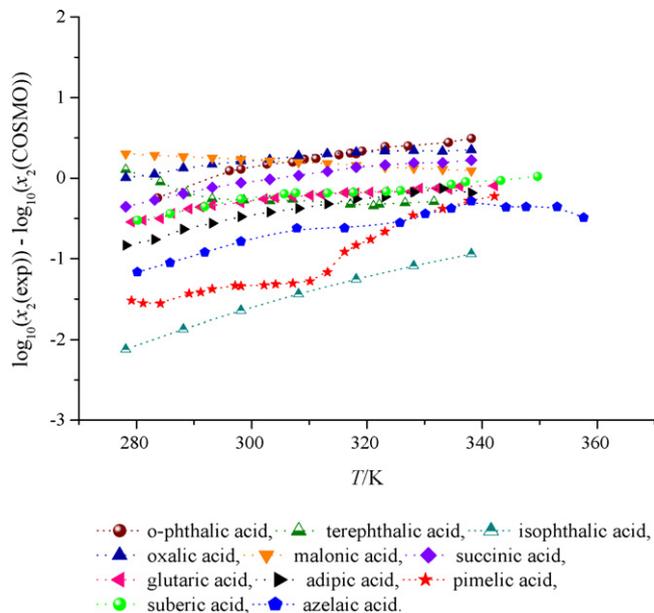


Fig. 5. Difference in solubilities between literature data and COSMO-RS/COSMOtherm predictions for dicarboxylic acids (including aromatic dicarboxylic acids).

3.1.3. Hydroxycarboxylic acids

As a subgroup of carboxylic acids, the examined hydroxycarboxylic acids show the least deviations in the solubilities predictions. COSMO-RS differentiates between high solubilities of alkylic hydroxycarboxylic acids and the somewhat lower solubilities of hydroxycarboxylic acids with aromatic core, as shown in Fig. 6.

A subdivision of the chosen solid carboxylic acids into aromatic carboxylic acids, hydroxycarboxylic acids and dicarboxylic acids, with the sometimes overlapping combination of different functional groups, is revealing the response of COSMO-RS/COSMOtherm, being subject to the nature of the acid structure, visible in a distinct stratification on the solubility scale.

The applied methodology differentiates well between different magnitudes of solubilities, according to various functionalities in the molecule and the resulting variations in their solution thermodynamics, as can be seen exemplified in Fig. 7, where the

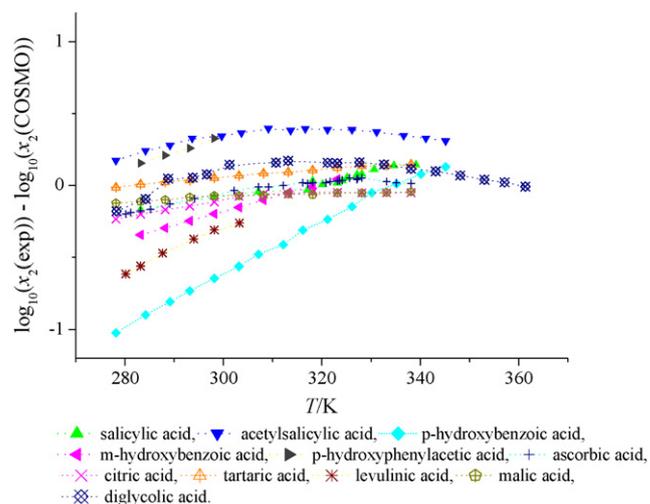


Fig. 6. Difference in solubilities between literature data and COSMO-RS/COSMOtherm predictions for hydroxycarboxylic acids (including aromatic carboxylic acids with hydroxyl groups).

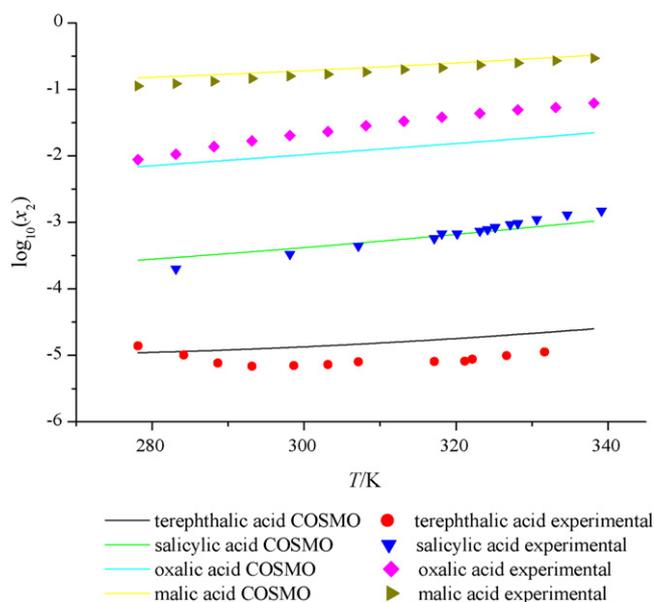


Fig. 7. Comparison between COSMO-RS/COSMOtherm predictions and experimental data for selected compounds of each of the families studied.

COSMO-RS/COSMOtherm predictions are compared against experimental data for a compound of each of the families studied.

3.2. Deviation from ideal solubility

Experimental and calculated activity coefficient values for the studied systems are compared in Fig. 8a and b. The values are spread over the axis of ordinates and they are ordered according to the type of ideality deviations resulting from the interactions between the solute and solvent molecules. Due to their lower affinity towards water, aromatic, substituted aromatic as well as dicarboxylic acids with long alkyl chains display significant positive deviations from Raoult's law, whereas dicarboxylic acids with shorter alkyl chains and hydroxycarboxylic acids show near-ideal solution behaviour. Oxalic acid exhibits the most distinctive negative deviation from Raoult's law, indicating the presence of preferential interactions with water.

COSMO-RS/COSMOtherm's modelling of activity coefficients provides an adequate description of the non-ideality of the studied systems, both in qualitative and quantitative terms, although the magnitude of the activity coefficients is slightly compressed, when compared with the experimental activity coefficients.

3.3. Apparent molar enthalpies of solution

The apparent molar enthalpies of solution, $\Delta_{\text{sol}}^{\text{app}}H_m$, at $T=298.15$ K, were estimated using experimental and COSMO-RS solubility data by fitting the solubility data to a quadratic equation:

$$\ln x_2 = a_0 + a_1 \left(\frac{1}{T}\right) + a_2 \left(\frac{1}{T}\right)^2 \quad (6)$$

leading to

$$\frac{\partial \ln x_2}{\partial (1/T)} = a_1 + 2a_2 \left(\frac{1}{T}\right) \quad (7)$$

According to Eq. (3), the apparent molar enthalpies of solution, $\Delta_{\text{sol}}^{\text{app}}H_m$, at T , can be derived as

$$\Delta_{\text{sol}}^{\text{app}}H_m(T) = -R \left(a_1 + 2a_2 \left(\frac{1}{T}\right) \right) \quad (8)$$

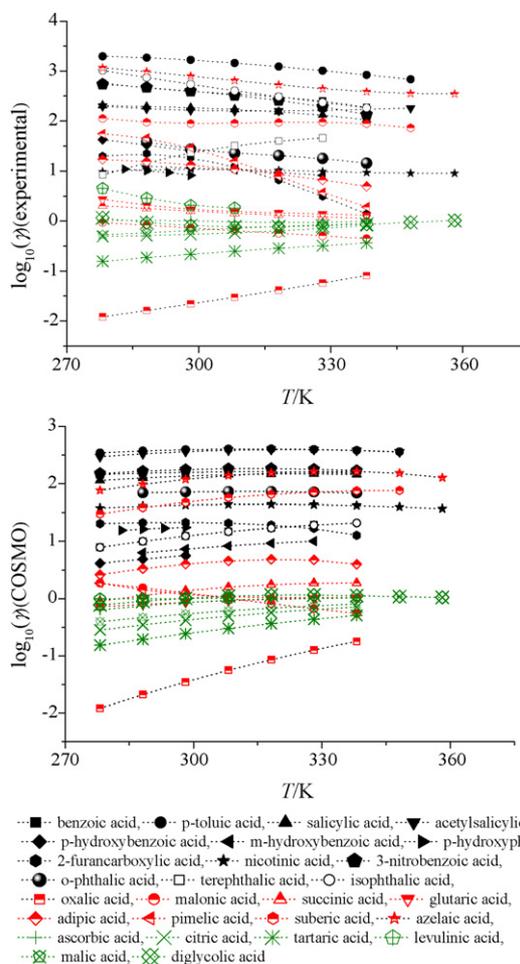


Fig. 8. Activity coefficients for the studied systems, as derived from literature data and predicted by COSMO-RS/COSMOtherm.

Table 2 lists the apparent molar enthalpies of solution, $\Delta_{\text{sol}}^{\text{app}}H_m$, at $T=298.15$ K as well as their deviations [$\Delta_{\text{sol}}^{\text{app}}H_m(\text{exp}) - \Delta_{\text{sol}}^{\text{app}}H_m(\text{COSMO})$].

The experimental molar enthalpies of solution of carboxylic acids are typically endothermic and small. The values derived from COSMO-RS are also endothermic but in general ca. (10–20 kJ mol⁻¹) lower than the ones obtained from experimental solubility data. In other words, in the case of the studied carboxylic acids, COSMO-RS systematically underestimates the temperature dependence of aqueous solubilities. The reason stays unclear; the deviation found exceeds the error normally expected for the temperature dependence of the chemical potential of the liquid compartment (which corresponds to ca. 4 kJ mol⁻¹) [73]. But taking into account that the standard molar enthalpies of vaporization of carboxylic acids, obtained with COSMO-RS/COSMOtherm, are also underestimated by the same order of magnitude (10–15 kJ mol⁻¹), it seems reasonable to suggest that the systematic underestimation of the temperature dependence of aqueous solubilities arises from deficiencies on the COSMO-RS treatment of the carboxylic acid liquid compartment, as depicted in Fig. 1.

4. Conclusions

Carboxylic acids, as a class of compounds, have several known issues related with their solution in water, ranging from possible dimer and hydrated species formation to a variety of conformers existing on the molecular level. Although a variety of thermophysical property data exist, several difficulties in their experimental

determination have been reported, including interference of melting with the onset of decomposition processes. Having this present, COSMO-RS/COSMOtherm is able to predict the temperature dependence of the aqueous solubilities of the 27 carboxylic acids tested here with an acceptable accuracy. Prerequisite for the application is the availability of accurate thermophysical properties. COSMO-RS/COSMOtherm's accuracy of the predicted solubilities seems to be sufficient in practical terms, while the derived apparent molar enthalpies of solution have a tendency to be underestimated. The bulk of the obtained COSMO-RS/COSMOtherm aqueous solubilities are slightly overestimated—a fact more pronounced for temperatures far from the triple point temperature.

List of symbols

a	activity
$\Delta_{\text{fus}}C_{p,m}$	molar heat capacity change upon fusion
$\Delta_{\text{fus}}G$	Gibbs energies of fusion
$\Delta_{\text{fus}}H$	molar enthalpy of fusion
$\Delta_{\text{trs}}H$	molar solid-phase transition enthalpy
$\Delta_{\text{tpce}}^1 H_m$	total molar phase change enthalpy
$\Delta_{\text{sol}}H_m$	molar enthalpy of solution
$\Delta_{\text{sol}}^{\text{app}}H_m$	apparent or Van't Hoff molar enthalpy of solution
m	molality
R	universal gas constant = $8.314472(15) \text{ J K}^{-1} \text{ mol}^{-1}$
$\Delta_{\text{fus}}S_m$	molar entropy of fusion
T	temperature
T_{fus}	fusion temperature
T_{trs}	transition temperature
x	mole fraction
x_2	mole fraction of solid 2 dissolved in the solvent phase at saturation

Greek letters

γ	activity coefficient
$\mu_2^{(P)}$	chemical potential of all pure compounds 2
$\mu_2^{(1)}$	chemical potential of compound 2 at infinite dilution in compound 1

Acknowledgments

Bernd Schröder is grateful to the Fundação Para a Ciência e Tecnologia, Lisboa, Portugal, and the European Social Fund (ESF) under the 3rd Community Support Framework (CSF), for the award of a Post-Doc scholarship SFRH/BPD/38637/2007.

References

- [1] Perry's Chemical Engineers' Handbook, eighth ed., (Chemical Engineers Handbook) McGraw-Hill, New York, 1997.
- [2] H. Renon, J.M. Prausnitz, *AIChE J.* 14 (1968) 135–144.
- [3] D.S. Abrams, J.M. Prausnitz, *AIChE J.* 21 (1975) 116–128.
- [4] J.M. Prausnitz, R.N. Lichtenthaler, E. Gomes de Azevedo, *Molecular Thermodynamics of Fluid-Phase Equilibria*, third ed., Prentice-Hall Inc., Upper Saddle River, New Jersey, 1999.
- [5] A. Fredenslund, R.L. Jones, J.M. Prausnitz, *AIChE J.* 21 (1975) 1086–1099.
- [6] S. Gracin, T. Brinck, Å.C. Rasmuson, *Ind. Eng. Chem. Res.* 41 (2002) 5114–5124.
- [7] M.B. Oliveira, M.J. Pratas, I.M. Marrucho, A.J. Queimada, J.A.P. Coutinho, *AIChE J.* 55 (2009) 1604–1613.
- [8] A. Klamt, *COSMO-RS from Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier, Amsterdam, 2005.
- [9] F. Eckert, A. Klamt, *AIChE J.* 48 (2002) 369–385.
- [10] S. Budavari (Ed.), *The Merck Index: An Encyclopedia of Chemicals, Drugs and Biologicals*, 12th ed., Merck Co. Inc., Whitehouse Station, NJ, 1996.
- [11] A.E. Williams, in: H.F. Mark, D.F. Othmer, C.G. Overberger, G.T. Seaborg (Eds.), *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 3, 3rd ed., Wiley, New York, 1978, pp. 778–792.
- [12] J.L. Opgrande, C.J. Dobratz, E.E. Brown, J.C. Liang, G.S. Conn, J. Wirth, J. Shelton, in: J.I. Kroschwitz, M. Howe-Grant (Eds.), *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 4, 4th ed., Wiley, New York, 1992, pp. 103–115.
- [13] C.M. Park, R.J. Sheehan, in: B. Elvers, S. Hawkins, G. Schultz (Eds.), *Ullmann's Encyclopedia of Industrial Chemistry*, vol. 18, fifth rev. ed. (Basel), 1991, pp. 991–1043.
- [14] D.E. Read, C.B. Purves, *J. Am. Chem. Soc.* 74 (1952) 116–119.
- [15] E.E. Schrier, M. Pottle, H.A. Scheraga, *J. Am. Chem. Soc.* 86 (1964) 3444–3449.
- [16] E. Suzuki, Y. Taniguchi, T. Watanabe, *J. Phys. Chem.* 77 (1973) 1918–1922.
- [17] K. Yamamoto, N. Nishi, *J. Am. Chem. Soc.* 112 (1990) 549–558.
- [18] M. Petukhov, G. Rychkov, L. Firsov, L. Serrano, *Protein Sci.* 13 (2004) 2120–2129.
- [19] M.B. Oliveira, V.L. Oliveira, J.A.P. Coutinho, A.J. Queimada, *Ind. Eng. Chem. Res.* 48 (2009) 5530–5536.
- [20] J.A.P. Coutinho, S.I. Andersen, E.H. Stenby, *Fluid Phase Equilib.* 103 (1995) 23–29.
- [21] J.A.P. Coutinho, J.-L. Daridon, *Energy Fuels* 15 (2001) 1454–1460.
- [22] C. Tsouopoulos, J.M. Prausnitz, *Ind. Eng. Chem. Fundam.* 10 (1971) 593–600.
- [23] S. Neau, S. Bhandarkar, E. Hellmuth, *Pharm. Res.* 14 (1997) 601–605.
- [24] F.L. Nordström, Å.C. Rasmuson, *J. Chem. Thermodyn.* 18 (40) (2008) 1684–1692.
- [25] J.S. Chickos, D.G. Hesse, *J. Org. Chem.* 55 (1990) 3833–3840.
- [26] F.L. Nordström, Å.C. Rasmuson, *Eur. J. Pharm. Sci.* 18 (36) (2009) 330–344.
- [27] J. Chocholoušová, J. Vacek, P. Hobza, *J. Phys. Chem. A* 107 (2003) 3086–3092.
- [28] J. Chen, C.L. Brooks, H.A. Scheraga, *J. Phys. Chem. B* 112 (2008) 242–249.
- [29] I.M. Kolthoff, MacMillan, New York, 1937.
- [30] M. Davies, D.M.L. Griffiths, *Trans. Faraday Soc.* 49 (1953) 1405–1410.
- [31] D.E. Giammar, D.A. Dzombak, *J. Solution Chem.* 27 (1998) 89–105.
- [32] A. Apelblat, E. Manzurola, *J. Chem. Thermodyn.* 19 (1987) 317–320.
- [33] A. Klamt, *J. Phys. Chem.* 99 (1995) 2224–2235.
- [34] A. Klamt, F. Eckert, *Fluid Phase Equilib.* 172 (2000) 43–72.
- [35] F. Eckert, A. Klamt, *COSMOtherm. Version C2.1 Release 01. 05*, COSMOlogic, GmbH & Co. Kg, Leverkusen, Germany, 2005.
- [36] F. Eckert, *COSMOtherm User's Manual Version C2.1, Release 01. 05*, COSMOlogic GmbH & Co. Kg, Leverkusen, Germany, 2005.
- [37] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* 162 (1989) 165–169.
- [38] A. Schäfer, A. Klamt, D. Sattel, J.C.W. Lohrenz, F. Eckert, *Phys. Chem. Chem. Phys.* 2 (2000) 2187–2193.
- [39] A. Schäfer, C. Huber, R. Ahlrichs, *J. Chem. Phys.* 100 (1994) 5829–5835.
- [40] A. Klamt, *Fluid Phase Equilib.* 206 (2003) 223–235.
- [41] M.G. Freire, S.P.M. Ventura, L.M.N.B.F. Santos, I.M. Marrucho, *Fluid Phase Equilib.* 268 (2008) 74–84.
- [42] M.G. Freire, P.J. Carvalho, R.L. Gardas, L.M.N.B.F. Santos, I. Marrucho, J.A.P. Coutinho, *J. Chem. Eng. Data* 53 (2008) 2378–2382.
- [43] M.G. Freire, P.J. Carvalho, R.L. Gardas, I.M. Marrucho, L.M.N.B.F. Santos, J.A.P. Coutinho, *J. Phys. Chem. B* 112 (2008) 1604–1610.
- [44] M.G. Freire, L.M.N.B.F. Santos, I.M. Marrucho, J.A.P. Coutinho, *Fluid Phase Equilib.* 255 (2) (2007) 167–178.
- [45] A.I. Caço, F.R. Varanda, M.J. Pratas de Melo, A. Dias, R. Dohrn, I.M. Marrucho, *Ind. Eng. Chem. Res.* 47 (2008) 8083–8089.
- [46] E.S. Domalski, E.D. Hearing, *J. Phys. Chem. Ref. Data* 22 (1993) 805–1159.
- [47] DIPPR DIADDEM Public v. 1.2.
- [48] A. Apelblat, E. Manzurola, N.A. Balal, *J. Chem. Thermodyn.* 38 (2006) 565–571.
- [49] A. Apelblat, E. Manzurola, *J. Chem. Thermodyn.* 31 (1999) 85–91.
- [50] F.L. Nordström, Å.C. Rasmuson, *J. Chem. Eng. Data* 51 (2006) 1668–1671.
- [51] A. Apelblat, E. Manzurola, *J. Chem. Thermodyn.* 21 (1989) 1005–1008.
- [52] D.R. Kirklin, *J. Chem. Thermodyn.* 32 (2002) 701–709.
- [53] S. Gracin, Å.C. Rasmuson, *J. Chem. Eng. Data* 47 (2002) 1379–1383.
- [54] A. Apelblat, E. Manzurola, *J. Chem. Thermodyn.* 22 (1990) 1527–1533.
- [55] F.L. Nordström, Å.C. Rasmuson, *Eur. J. Pharm. Sci.* 28 (2006) 377–384.
- [56] M.V. Roux, M. Temprado, P. Jiménez, C. Foces-Foces, M.V. Garcia, M.I. Redondo, *Thermochim. Acta* 420 (2004) 59–66.
- [57] L. Dian-Qing, L. Jiang-Chu, L. Da-Zhuang, W. Fu-An, *Fluid Phase Equilib.* 200 (2002) 69–74.
- [58] R. Sabbah, L. Perez, *Can. J. Chem.* 77 (1999) 1508–1513.
- [59] J. Sangster, *J. Phys. Chem. Ref. Data* 28 (1999) 889–930.
- [60] S.H. Yalkovskiy, Y. He, *Handbook of Aqueous Solubility Data*, CRC Press, Boca Raton, FL, 2003.
- [61] U.S. Rai, K.D. Mandal, *Bull. Chem. Soc. Jpn.* 63 (1990) 1496–1502.
- [62] A. Chacko, R. Devi, S. Abraham, B. Mathew, *J. Appl. Pol. Sci.* 96(2005) 1897–1905.
- [63] E. Manzurola, A. Apelblat, *J. Chem. Thermodyn.* 34 (2002) 1127–1136.
- [64] W. Omar, *J. Ulrich, Cryst. Growth Des.* 6 (2006) 1927–1930.
- [65] M.V. Roux, M. Temprado, J.S. Chickos, *J. Chem. Thermodyn.* 37 (2005) 941–953.
- [66] A. Apelblat, E. Manzurola, *J. Chem. Thermodyn.* 22 (1990) 289–292.
- [67] *CRC Handbook of Chemistry and Physics* 87th edition, Taylor and Francis, 2007.
- [68] E.M.S. Maçõas, R. Fausto, M. Pettersson, L. Khriachtchev, M. Räsänen, *J. Phys. Chem. A* 104 (2000) 6956–6961.
- [69] P.D. Godfrey, M.J. Mirabella, R.D. Brown, *J. Phys. Chem. A* 104 (2000) 258–264.
- [70] C. Chen, S.-H. Shyu, *Int. J. Quantum Chem.* 76 (2000) 541–551.
- [71] J.M. Hermida-Ramón, E.M. Cabaleiro-Lago, J. Rodríguez-Otero, *Chem. Phys.* 302 (2004) 53–60.
- [72] F.L. Nordström, Å.C. Rasmuson, *J. Pharm. Sci.* 95 (2006) 748–760.
- [73] F. Eckert, Private communication, 2009.