Extended Abstract of the PhD thesis

"Measuring and Modeling Thermodynamic Properties of Biological Solutions"

by Dr.-Ing. Christoph Held

Date work commenced: April 1st 2007
Date of the defence: December 1st 2011

1 Summary

Thermodynamic models provide the basis for the design of biochemical processes. However, for biological systems predictive or universally-valid models still do not exist in the literature. In this thesis thermodynamic properties of aqueous solutions with biological impact are considered, which may contain also alcohols (methanol, ethanol) as further solvents and solutes such as electrolytes (salts, acids, bases), amino acids/peptides, sugars, and osmolytes.

Prior to the modeling of these systems, a consistent and broad experimental data basis was established in a first step. Appropriate experimental data for model-parameter estimation especially are activity coefficients γ as they are a direct measure for the molecular interactions in mixtures. On top of these γ values, also solution densities and solubilities of the binary systems osmolyte/water, amino acid/water, salt/alcohol and of the (qua)ternary solutions osmolyte/salt/water, amino acid/salt/water, and amino acid/amino acid(s)/water were measured for broad both temperature and concentration ranges, respectively. This provides the basis for applying a thermodynamic model.

In this work the electrolyte Perturbed-Chain Statistical Association Theory (ePC-SAFT) developed by Cameretti *et al.* [1] in 2005 was used for modeling biological solutions. Components which form hydrogen bonds (e.g. water, alcohols, amino acids, or osmolytes) are described with five pure-component parameters whereas strong electrolytes (e.g. Na^+ in NaCl) possess two ion-specific parameters only. This allows for a quantitative description of thermodynamic properties of electrolyte/water [2], electrolyte/alcohol [3], and biomolecule (amino acid, osmolyte, sugar)/water [4, 5] solutions. Besides fully dissociated electrolytes or amino acids which are exclusively

present as neutral zwitterions, also weak electrolytes and charged amino acids were considered. Modeling such systems succeeded by implementing respective dissociation equilibria within ePC-SAFT which explicitly account for the prevailing pH-value. Based on these binary solute/solvent systems, densities, activity coefficients, and solubilities in (qua)ternary solutions (salt/water/alcohol(s) [6], amino acid/salt/water, amino acid/amino acid(s)/water [4, 7]) could be *predicted* with ePC-SAFT, i.e. additional fitting parameters were not required.

2 Problems Adressed

Thermodynamic tools for chemical substances and mixtures are widely available and accepted in industry and academia. In contrast, solutions with biological impact have been neglected by the modelers for a long time. This is due to the variety of system conditions in biological systems, which makes it almost impossible to provide a broad experimental data basis which is needed for a physically-sound model development. To give an example, not a single publication showing the salt influence on the vapor-liquid equilibrium of amino-acid solutions was available in the literature in the beginning of this thesis (April 2007). Even more, activity coefficients are mostly neglected in biological process calculations. Recently, also biologists have started to recognize that there are conditions where this assumption leads to wrong results or simply misinterpretation of results. Examples where the activity coefficients in biological solutions are crucial are (1) deciding whether biochemical reactions take place or not, (2) determining the amount of a component which is soluble in a solvent (its solubility), and (3) the magnitude of the osmotic pressure which arises by introducing solutes in a solvent.

The main focus of this work was to measure and model biological solutions. These solutions were restricted to water and alcohols as solvents, whereas many different solutes were considered: Weak and strong electrolytes (including acids and bases), amino acids and peptides, sugars and sugar alcohols, and osmolytes. The investigated properties were solution densities, solvent activity coefficients (vapor pressure, osmotic pressure), solute activity coefficients (solubility in solvents), and pH-value. The basis for a systematic modeling of such systems in this work is provided by the electrolyte Perturbed-Chain SAFT equation of state (ePC-SAFT), which was introduced by Cameretti *et al.* [1] in 2005.

3 State of the art

Since the early years of the last century phase-equilibria (e.g. vapor pressures or solubilities) of solvents have been correlated and calculated. In general, there are two different kinds of approaches to model the required substance properties: Gibbs excess-energy (G^E) models and equations of state (EOS). Independent of the kind of approach, the basis for modeling of electrolyte solutions is e.g. provided by the work of Debye and Hückel (DH) in 1923 [8] or by the Mean Spherical Approximation (MSA) introduced by Waisman and Lebowitz [9] in 1970 who solved the Ornstein-Zernike equation for a fluid of charged spheres of equal size.

To obtain electrolyte models, non-electrolyte approaches were combined with appropriate theories accounting for the Coulomb forces among ions. Electrolyte G^E models and electrolyte EOS have been developed so far. Examples are the electrolyte NRTL [10] or the Pitzer model [11]. Nasirzadeh *et al.* [12] used a MSA-NRTL model [13] as well as an extended Pitzer model of Archer [14]. However, a huge number of adjustable component parameters are needed for these models. Myers *et al.* [15] developed an electrolyte model based on the PR EOS. Although osmotic and activity coefficients could be described accurately, the PR EOS failed quantitatively in calculating liquid densities which is a commonly-known problem [1, 15]. Aiming at the description of even more complex biological systems, where the molecules exhibit a rod-like structure, segment-based models appear to be more appropriate compared to the models mentioned above, one among them being the SAFT model. Just to give some examples, Liu and co-workers [16] combined SAFT with the MSA primitive model to predict MIAC in two-salt solutions; however, they used salt-specific parameters. Radosz *et al.* recently published their SAFT1 [17, 18] and SAFT2 [19, 20] EOS yielding excellent results concerning the properties of aqueous single-salt and multi-salt solutions. Three individual-ion parameters as well as one additional salt parameter have to be adjusted for each electrolyte.

4 Key Innovations

In contrast to chemical mixtures, precise and reliable models are still hardly available for multi-component biological systems. This is due to the fact that systematic and reliable experimental data are still not available and that biological systems are dominated by complex and specific interactions. Thus, the industry often still uses

(semi-)empirical approaches, which possess a very high number of model parameters or even unphysical model parameters. Moreover, most of the existing models are trimmed to describe very specific systems, making model predictions impossible. To overcome this lack in data and in the modeling, the work contains the following key innovations:

- Experimental data: In this work, experimental osmotic coefficients and solubility data of biomolecule/water were determined. This data could explain why certain biomolecules act as compatible solutes whereas others do not. Further, the salt effect on this data was measured systematically with respect to a number of different ions and ionic concentration.
- Modeling binary systems: In this work biomolecule/water (amino acids, peptides, osmolytes, sugars) and electrolyte/water solutions were modeled with ePC-SAFT. The key innovation was the use of a minimal parameter set for the reasonable description of a huge number of systems (e.g. 48 parameters for more than 100 electrolyte solutions). Effects in weak-electrolyte solutions (ion pairing) could be described quantitatively.
- Predicting ternary systems: ePC-SAFT and the pure-component parameters were applied to predict thermodynamic properties of a huge number of multi-component biological systems that may contain different solvents, biomolecules (e.g. amino acids, sugars, osmolytes), and electrolytes. The key innovation is the ability to describe properties of systems or system conditions where component parameters had not been adjusted to (prediction). The small number of pure-component paramneters did not cause losses concerning accuracy or predictive capability.

To sum up, this work shows that ePC-SAFT is able to predict thermodynamic properties of multi-component biological systems. The combination of both, a small number of parameters and the predictive capability, is new and not published in the literature so far. The established data basis on osmotic coefficients and solubility in amino acid/electrolyte/water systems is also new.

5 Applications and Results

5.1 Experiments

Thermodynamic properties of binary and ternary mixtures of biological aqueous solutions were measured, among them solution densities, osmotic coefficients, and

solubilities. These solutions contained salts, amino acids/peptides, and osmolytes, respectively. The following substances were used for the experiments: LiCl, LiBr, LiI, NaCl, NaBr, NaI, $NaNO_3$, Na_2SO_4, KCl, KBr, KI, NH_4Cl, NH_4Br, NH_4I, L-valine, DL-valine L/DL-alanine, glycine, L-leucine, DL-norleucine, DL-norvaline, L-cysteine, L-methionine, diglycine, L-leucine, L-alanine, and L-valine, the osmolytes ectoine, hydroxyectoine, homoectoine, and DHMICA, and the solvents water, methanol, and ethanol, respectively.

5.2 Modeling Binary Solutions with ePC-SAFT

5.2.1 Strong-Electrolyte and Weak-Electrolyte Aqueous Solutions

Strong electrolytes are systems containing solute(s) (e.g. salts, acids) which are completely dissociated in the solvent(s). The strong electrolytes considered in this work are composed of the solvent(s) water, methanol, and ethanol and of the cations Cat^+ = [H^+, Li^+, Na^+, K^+, Rb^+, Cs^+, NH_4^+, $Choline^+$, Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}, Co^{2+}, Cu^{2+}, Fe^{2+}, Cr^{3+}] and the anions An^- = [F^-, Cl^-, Br^-, I^-, OH^-, SCN^-, ClO_3^-, ClO_4^-, $H_2PO_4^-$, BrO_3^-, NO_3^- , SO_4^{2-}, HPO_4^{2-}], respectively.

Implementing ion pairing is in particular reasonable for salts which are derivatives of weak acids, like acetates. Strong electrolytes are subjected to characteristic ion-specific effects, e.g. their mean ionic activity coefficients (MIAC) are ordered in the sequence $\gamma^*_{Li+} > \gamma^*_{Na+} > \gamma^*_{K+}$. In contrast, electrolytes containing the acetate anion show a reversed sequence of experimental MIACs $\gamma^*_{K+} > \gamma^*_{Na+} > \gamma^*_{Li+}$, shown in Figure 1 .

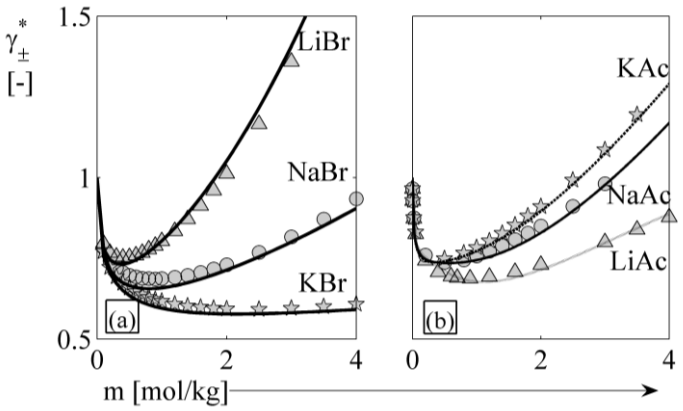

Figure 1: Experimental mean ionic activity coefficients of (a) alkali bromides and (b) alkali acetates in water at 25°C. The symbols represent experimental data [21]. Lines are modeling results wit ePC-SAFT.

5.2.2 Aqueous Biomolecule Solutions

Biomolecules considered in this work are amino acids (glycine, alanine, serine, proline, hydroxyproline, valine, leucine, arginine, lysine, threonine, asparagine, tyrosine, histidine, cysteine, methionine, aspartic acid, glutamic acid, α-ABA, α-isoABA, β-ABA, γ-ABA, α-AVA, γ-AVA), peptides (diglycine, triglycine, dialanine, Gly-Ala, Ala-Gly), osmolytes (ectoine, hydroxyectoine, homoectoine, DHMICA), sugars and sugar alcohols (sucrose, maltose, galactose, glucose, mannose, fucose, xylose, mannitol, sorbitol) and urea, respectively.

The overall absolute relative deviations ARD between experiment and PC-SAFT modeling for the more than 40 considered biomolecule/water systems are very small ($ARD_{density}$ < 0.1 %, $ARD_{vapor\ pressure}$ < 0.5 %, $ARD_{activity\ coefficient}$ < 1 %) with the highest error in solubility data ($ARD_{solubility}$ < 4 %). Presumably, this is caused by uncertainties in the experiments and the applied group-contribution method (for the melting properties) as well as by the neglect of heat-capacity values. Moreover, most of the solubility calculations were performed within a broad temperature region (0-100°C). Figure 2 shows a typical modeling result for some biomolecules.

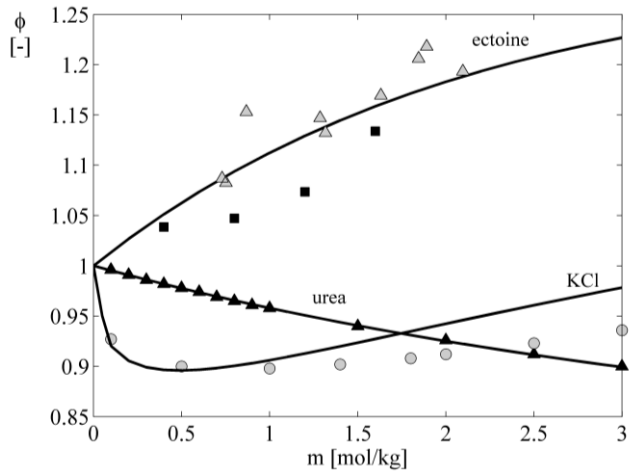

Figure 2: Osmotic coefficients in aqueous solutions at 30°C. Symbols are experimental data (KCl/water: circles [21], ectoine/water: full squares [this work] and open triangles unpublished data by the group of Kunz [22], and urea/water: full triangles [23]), lines are calculations with ePC-SAFT.

Figure 2 further allows some discussion on the physical effect of osmolytes. Whereas the ectoines are known to be used as osmolytes, other amino acids or urea are not. Moreover, although almost every microorganism possesses ion pumps, they often rather produce compatible solutes against stresses instead of the energetically cheaper alternative to accumulate ions from their surrounding. Ectoine causes high osmotic

coefficients whereas the opposite is valid for solutions containing salt or urea, i.e. the osmotic coefficient *decreases* by adding salts or urea. Decreasing osmotic coefficients cause lower osmotic pressures π ($\pi \sim \phi$), i.e. ineffective anti-stress agents.

5.3 Predicting (Qua)Ternary Solutions with ePC-SAFT

Up to this point, binary solvent/solute mixtures have been investigated. Model parameters have been fitted for ions, amino acids, small peptides, osmolytes, and sugars. However, biological solutions never contain one solute or one solvent only. This chapter discusses the applicability of ePC-SAFT for predicting thermodynamic properties in (qua)ternary solutions containing multiple solutes in one solvent (water/ biomolecule/biomolecule(s) and water/biomolecule/salt) on the one hand and solutions containing one salt in multiple solvents (water/alcohol(s)/salt) on the other hand.

5.3.1 Osmotic coefficients

In the previous section it was shown that compatible solutes (e.g. ectoines) cause increased ϕ values whereas the opposite was found for incompatible solutes (e.g. urea). For solute (using the example urea in Figure 3) concentrations higher than 0.5 m this effect seems to be still valid in systems where a third component (0.5 m glycine in Figure 3) is successively added, i.e. osmotic coefficients decrease for adding urea independent of the presence of glycine. Also this behavior is predicted quantitatively by PC-SAFT. However, at urea concentrations lower than 0.5 m this behavior is reversed, which causes a maximum of the osmotic coefficients in the case of urea. This behavior can be explained by the fact that at very low urea concentrations compared to 0.5 m glycine, glyine will dominate the phase behavior in the system (osmotic coefficients in the ternary system similar to the binary 0.5 m glycine/water solution) whereas at higher urea concentrations urea will be the phase-dominant compound (osmotic coefficients in the ternary system similar to the binary urea/water solution).

The model can also be applied to predict osmotic coefficients of systems containing salts. Figure 4 shows osmotic coefficients of aqueous valine/NaCl and valine/$NaNO_3$ solutions for varying salt molalities (0.5 - 3 m) at 25°C. Also here, ePC-SAFT allows for precisely predicting the experimental data. The ARD between ePC-SAFT prediction and experimental data of the 22 glycine/salt and valine/salt systems measured in this work for salt molalities of 0.5 and 1 m only is 1.19 %. That is, the predictive capability makes osmotic experiments dispensable in amino acid/salt solutions.

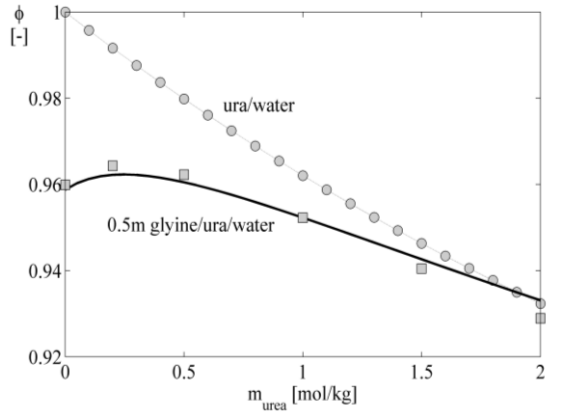

Figure 3: Osmotic coefficients of aqueous urea solutions containing no glycine or 0.5 m glycine as function of urea molality at 25°C. Thin lines and circles are modeled and experimental osmotic coefficients of the binary glycine-free urea/water system at 25°C, respectively. Squares and thick lines represent data from [24] and predictions with PC-SAFT of the ternary 0.5m glycine/urea/water system .

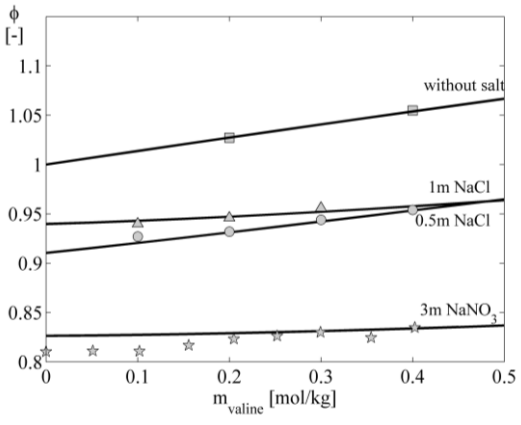

Figure 4: Osmotic coefficients of valine/salt/water solutions at 25°C. Symbols are experimental data from this work [25] (squares: without salt, triangles: 1 m NaCl, circles: 0.5 m NaCl, stars: 3 m $NaNO_3$). Lines are predictions with ePC-SAFT, i.e. the modeling is based on the pure-component salt and amino-acid parameters only.

5.3.2 Mean Ionic Activity Coefficients (MIAC)

In this thesis a modeling strategy was developed for the description of ions in mixed solvents. Mixing rules for permittivity and the mean solvated ion dimaeters were propsed [3]. With this, ePC-SAFT is able to describe the MIACs up to the respective NaCl solubility limits with an average deviation (ARD) of 6.7 % between 31 and 91 mol% ethanol in the salt-free solvent system (see Figure 5). No additional adjustable parameters were introduced so that these results are pure predictions. Even more, due to the low solubility of NaCl in ethanol, the Na^+ and Cl^- ePC-SAFT parameters were not adjusted to osmotic or activity coefficients of NaCl in ethanol but only to other salts containing Na^+ and Cl^- (i.e. to binary solution containing ethanol and a salt (NaI, LiCl, and NH_4Cl)). This again reveals the predictive capability of ePC-SAFT and the advantage of applying ion-specific model parameters. In sum, applying ePC-SAFT allows for modeling the MIAC and the densities of the systems water/MeOH + salt (KCl, NaCl, NaBr, LiCl) and water/EtOH + salt (KCl, NaCl, NaBr, NaI) with comparatively small overall ARD values (0.12 % for the densities and 3.57 % for the MIACs)[6].

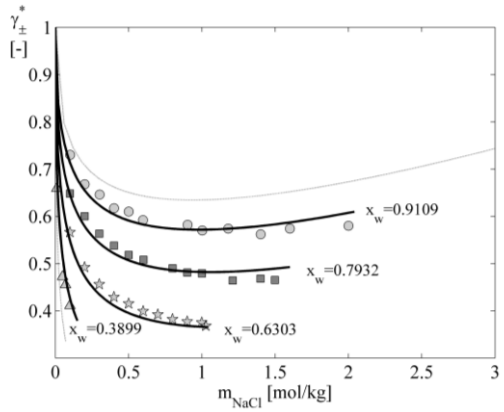

Figure 5: Experimental and modeled mean ionic activity coefficients of NaCl in water/ethanol mixtures at 25°C and salt-free water fractions between 39 and 91 mol%. The symbols represent experimental data [26]. The full lines are predictions with ePC-SAFT. The grey lines represent the MIACs of NaCl in the binary water/NaCl and ethanol/NaCl solutions.

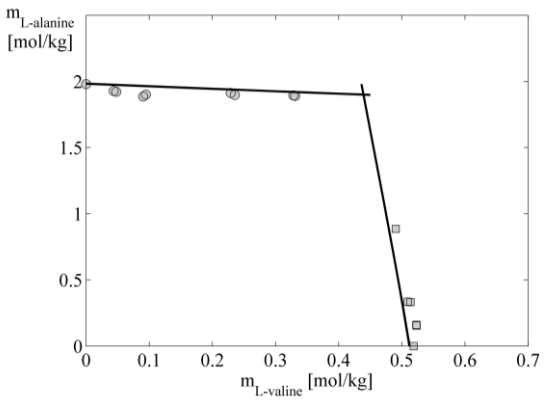

Figure 6: Solubility of L-alanine (circles) in the presence of L-valine and vice versa (squares) at 25°C. Symbols are experimental data [7], lines are predictions with PC-SAFT. Both amino acids decrease the solubility of the other amino acid.

5.3.3 Solubilities

This section deals with the amino-acid solubility in (qua)ternary aqueous solutions with one pure amino acid in the solid phase. Depending on the kind of amino acid which is added to an already saturated amino-acid solution, the solubility behavior differs remarkably. In case of two hydrophobic amino acids the added solutes decrease their mutual solubility. That is, alanine, leucine, or valine are less soluble if another hydrophobic amino acid is present (leucine or valine) which is illustrated exemplarily in Figure 6. A totally different picture is valid for two hydrophilic amino acids which cause a mutual increase of their water solubilities. This behavior can also be predicted with PC-SAFT (k_{ij} =0, results not shown).

To sum up, in this thesis solution densities, activity coefficients, and solubilities in aqueous solutions containing electrolytes and biomolecules were measured and modeled. Ions and biomolecules were modeled with two and five pure-component parameters. This allowed for quantitative model predictions (i.e. parametrization of model parameters only to binary solute/water systems) in multi-solute and multi-solvent aqueous systems over a broad range of system conditions.

6 References

1. *Cameretti, L.F., G. Sadowski, and J.M. Mollerup, Industrial & Engineering Chemistry Research, 2005* ***44*** *3355-3362; ibid., 8944.*
2. *Held, C., L.F. Cameretti, and G. Sadowski, Fluid Phase Equilibria, 2008* ***270*** *87-96.*
3. *Held, C., et al., Chemical Engineering Science, 2012* ***68*** *328-339.*
4. *Held, C., L.F. Cameretti, and G. Sadowski, Industrial & Engineering Chemistry Research, 2011* ***50*** *131-141.*
5. *Held, C., T. Neuhaus, and G. Sadowski, Biophysical Chemistry, 2010* ***152*** *28-39.*
6. *Held, C., et al., Chemical Engineering Science, 2011submitted.*
7. *Grosse Daldrup, J.B., et al., Industrial & Engineering Chemistry Research, 2009* ***49*** *1395-1401.*
8. *Debye, P. and E. Hückel, Physikalische Zeitschrift, 1923* ***9*** *185-206.*
9. *Waisman, E. and J.L. Lebowitz, Journal of Chemical Physics, 1970* ***52*** *4307-4311.*
10. *Chen, C.C., et al., Aiche Journal, 1982* ***28*** *588-596.*
11. *Pitzer, K.S., Journal of Physical Chemistry, 1973* ***77*** *268-277.*
12. *Nasirzadeh, K., R. Neueder, and W. Kunz, Industrial & Engineering Chemistry Research, 2005* ***44*** *3807-3814.*
13. *Papaiconomou, N., et al., Physical Chemistry Chemical Physics, 2002* ***4*** *4435-4443.*
14. *Archer, D.G., Journal of Physical and Chemical Reference Data, 1991* ***20*** *509-555.*
15. *Myers, J.A., S.I. Sandler, and R.H. Wood, Industrial & Engineering Chemistry Research, 2002* ***41*** *3282-3297.*
16. *Liu, W.B., Y.G. Li, and J.F. Lu, Fluid Phase Equilibria, 1999* ***160*** *595-606.*
17. *Tan, S.P., H. Adidharma, and M. Radosz, Industrial & Engineering Chemistry Research, 2005* ***44*** *4442-4452.*
18. *Ji, X.Y., et al., Industrial & Engineering Chemistry Research, 2005* ***44*** *7584-7590.*
19. *Tan, S.P., et al., Journal of Physical Chemistry B, 2006* ***110*** *16694-16699.*
20. *Ji, X.Y. and H. Adidharma, Industrial & Engineering Chemistry Research, 2006* ***45*** *7719-7728.*
21. *Lobo, V.M.M. and J.L. Quaresma, Handbook of Electrolyte Solutions, Parts A and B1989, Amsterdam: Elsevier.*
22. *Ramsauer, B. and W. Kunz, Measurement of Osmotic Coefficients in Ectoine/Water Solutions between 15 and 35°C (personal communication), 2010: University of Regensburg.*
23. *Ellerton, H.D. and P.J. Dunlop, The Journal of Physical Chemistry, 1966* ***70*** *1831-1837.*
24. *Rafflenbeul, L., et al., Zeitschrift Fur Naturforschung C, 1973* ***C 28*** *533-554.*
25. *Held, C., et al., Journal of Chemical Thermodynamics, 2013, to be submitted.*
26. *Esteso, M.A., et al., Journal of Solution Chemistry, 1989* ***18*** *277-288.*