



# Henry's law constants of diacids and hydroxy polyacids: recommended values

S. Compernelle and J.-F. Müller

Belgian Institute for Space-Aeronomy, 1180 Brussels, Belgium

Correspondence to: S. Compernelle (steven.compernelle@aeronomie.be)

Received: 27 August 2013 – Published in Atmos. Chem. Phys. Discuss.: 26 September 2013

Revised: 3 February 2014 – Accepted: 5 February 2014 – Published: 14 March 2014

**Abstract.** In spite of the importance of diacids and functionalised diacids for organic aerosol formation through aqueous-phase processes in droplets and aerosol water, there seems to be no reliable set of experimental values for their Henry's law constants (HLCs). We show that their estimation through the use of infinite dilution activity coefficients is also prone to error. Here we present HLC values for diacids and hydroxy polyacids determined from solubilities, water activities and vapour pressures of solids or solutions, by employing thermodynamic relationships. The vapour pressures are found to be the largest source of error, but the analysis of the obtained HLC points to inconsistencies among specific vapour pressure data sets. Although there is considerable uncertainty, the HLC defined as aqueous concentration per unit gaseous partial pressure of linear  $\alpha$ - and  $\omega$ -diacids appear to be higher than estimated by the often cited review work of Saxena and Hildemann (1996).

## 1 Introduction

Henry's law constant (HLC) describes the partitioning of a compound between the gas phase and a liquid, highly dilute solution. In the atmosphere, such dilute solutions, with water as solvent, can be reached in cloud droplets. Aqueous aerosols represent another example of an aqueous phase in the atmosphere. Although in the latter case this phase is not a highly dilute aqueous phase due to the large concentration of other organic and inorganic molecules, Henry's law constants for the aqueous phase are still relevant as a reference point. Polyacids are important constituents in droplets and aerosols. An often-cited HLC compilation for atmospheric purposes is that of Sander (1999); these values are also avail-

able in the NIST Chemistry WebBook (Linstrom and Mallard). This compilation contains HLCs for several diacids and functionalised diacids, but most of them are not experimental but rather estimated through group contribution, taken from the review paper of Saxena and Hildemann (1996). We note that Mentel et al. (2004) did measure the HLC of glutaric acid.

In any case, HLC defined as aqueous concentration per unit gaseous partial pressure of dicarboxylic acids are likely very high, such that they will be almost completely dissolved in aqueous droplets. However, for aqueous aerosols, the total water content of the aqueous phase is much lower and the gas phase fraction could be significant. Furthermore, HLC estimation methods also need reliable HLC to fit their parameters.

Through thermodynamics, vapour pressures of liquids or solids, solubilities and activity coefficients are all related. The focus in this work is on HLC, but we will need the other quantities as well. Therefore, we first briefly review their thermodynamic relationships. For HLC, several definitions exist. We will follow here the convention taken by Sander (1999):

$$k_h \equiv \lim_{c_s, p_s \rightarrow 0} \frac{c_s}{p_s} = \lim_{x_s, p_s \rightarrow 0} \frac{x_s c_w}{p_s}, \quad (1)$$

with  $c_s$  the molar concentration of the solute,  $s$ , in the aqueous solution,  $p_s$  its partial pressure above it,  $x_s$  the mole fraction in the aqueous phase and  $c_w$  the molar concentration of pure water (55.6 M). As expressed by Eq. (1), they should in principle be measured at the infinite dilution limit (IDL). For an ideal solution where there is no difference in interaction between like or unlike molecules the partial pressure would equal  $p_s = x_s p_{L,s}^0$  (Raoult's law), with  $x_s$  the mole fraction of the solute and  $p_{L,s}^0$  its liquid saturation vapour

pressure. However, in general, solutions are not ideal, the solute molecules behave differently in water than in a liquid of pure *s*, and therefore an activity coefficient correction is needed.

$$p_s = \gamma_s x_s p_{L,s}^0 \quad (2)$$

$\gamma_s$  is the mole fraction based activity coefficient. It expresses the preference of the solute to the mixture, compared to the pure component reference state at the same temperature and pressure. Note that in Eq. (2) it is assumed that the gas phase of the solute behaves ideally, a reasonable assumption given that the partial pressure of the solute will be small. If the solvent is the same as the solute,  $\gamma_s = 1$ , while a value below (above) unity means that the solute prefers the mixture (solvent of pure *s*). Combination of Eqs. (1) and (2) leads to

$$k_h = \frac{c_w}{\gamma_s^\infty p_{L,s}^0}, \quad (3)$$

with  $\gamma_s^\infty$  the infinite dilution activity coefficient (IDAC) and  $p_{L,s}^0$  the liquid saturation vapour pressure. Suppose now that the solute is a crystalline solid in pure form. Above the solid, gas phase molecules with a solid state vapour pressure  $p_{Cr,s}^0$  will be present. Now if the solute is added to water beyond its solubility limit, a solid phase will form in the aqueous phase. If it can be assumed that this solid is the same as in the dry form (e.g. there is no incorporation of water in the crystal structure), one has

$$p_s^{\text{sat}} = p_{Cr,s}^0 = \gamma_s^{\text{sat}} x_s^{\text{sat}} p_{L,s}^0, \quad (4)$$

with  $x_s^{\text{sat}}$  the mole fraction solubility of the solid and  $\gamma_s^{\text{sat}}$  the activity coefficient at this point. In principle, a compound's pure liquid and solid state can only coexist at the melting curve, where  $p_{Cr}^0 = p_L^0$ . Hence if the substance is solid at the temperature of interest, the vapour pressure of the sub-cooled liquid is inaccessible from a thermodynamic point of view, as  $x_s$  cannot increase above  $x_s^{\text{sat}}$ . In practice, however, metastable, strongly supersaturated solutions might exist in small particles (Peng et al., 2001; Soonsin et al., 2010; Huisman et al., 2013), such that  $p_s \rightarrow p_{L,s}^0$  can be approached.  $p_{L,s}^0$  can also be related to  $p_{Cr,s}^0$ , as sublimation can be seen thermodynamically as first a melting of the solid, and then a vaporization of the resulting liquid. The fusion enthalpy and entropy have then to be taken into account:

$$p_{Cr,s}^0(T) = \exp\left(-\frac{\Delta H_{\text{fus}}(T)}{RT} + \frac{\Delta S_{\text{fus}}(T)}{R}\right) p_{L,s}^0(T). \quad (5)$$

Unfortunately,  $\Delta H_{\text{fus}}$  and  $\Delta S_{\text{fus}}$  can only be measured at the fusion point, where  $\Delta H_{\text{fus}} = T_{\text{fus}} \Delta S_{\text{fus}}$ , which is in our case often far above the temperature of interest (around room temperature). Thus, extrapolation schemes are necessary, e.g. by assuming a constant heat capacity difference,  $\Delta C_{p,ls}$ , be-

tween liquid and solid:

$$\frac{\Delta H_{\text{fus}}(T)}{RT} - \frac{\Delta S_{\text{fus}}(T)}{R} \approx -\frac{\Delta H_{\text{fus}}(T_{\text{fus}})}{R} \left(\frac{1}{T_{\text{fus}}} - \frac{1}{T}\right) + \frac{\Delta C_{p,ls}}{R} \left(1 - \frac{T_{\text{fus}}}{T} - \ln \frac{T}{T_{\text{fus}}}\right). \quad (6)$$

$\Delta C_{p,ls}$  is generally not available and has to be estimated, e.g. by neglecting it or imposing  $\Delta C_{p,ls} \approx \Delta S_{\text{fus}}(T_{\text{fus}})$ . This is the approach taken by e.g. Booth et al. (2010).

If the solubility is low, such that  $x_s^{\text{sat}}$  is close to the IDL, one has

$$k_h \approx \frac{c_s^{\text{sat}}}{p_s^{\text{sat}}} = \frac{c_s^{\text{sat}}}{p_{Cr,s}^0} \approx c_w \frac{x_s^{\text{sat}}}{p_{Cr,s}^0}. \quad (7)$$

In this work, we will first investigate the reliability of HLC values of diacids present in the literature. Then we will evaluate the reliability of UNIFAC (UNIQUAC Functional-group Activity Coefficient)-type group contribution methods to obtain IDAC, as the IDAC is one ingredient, apart from the sub-cooled liquid vapour pressure  $p_L^0$ , to obtain HLC via Eq. (3). Next we will show how more reliable HLC values can be constructed using thermodynamic relationships and existing experimental data on solid state vapour pressures, solubilities and water activities. Vapour pressures of solids or solutions bear the largest uncertainty, with considerable disagreement between the results of different research groups (e.g. Booth et al., 2010; Soonsin et al., 2010; Cappa et al., 2007; Chattopadhyay and Ziemann, 2005). However, comparing the thus-derived HLC between different molecules will reveal clues as to the consistency of the vapour pressure data sets.

## 2 Case study on the reliability of compiled literature values for diacids

HLC estimation methods need of course experimental data to fit and/or test their models. Some studies (e.g. Raventos-Duran et al., 2010; Hilal et al., 2008; Modarresi et al., 2007) report also estimated vs. experimental HLC values for diacids. The origin and reliability of the experimental values is then of prime importance, as any estimation method can only be as good as the data on which it is based on. We choose here the data set compiled by Raventos-Duran et al. (2010) to fit the model GROMHE, but similar conclusions would be very likely drawn for other compilations. GROMHE is developed especially for atmospherically relevant compounds. It is based on a compilation which includes data for five diacids: oxalic, malonic, succinic, glutaric and adipic acids. Their HLC values as used by Raventos-Duran et al. (2010) are presented in Table 1.

The HLC of oxalic acid was taken from Gaffney et al. (1987), however this article refers to a conference abstract (Gaffney and Senum, 1984) we could not obtain, and it is therefore unclear how the value was originally obtained.

**Table 1.** HLC of diacids from the compilation of Raventos-Duran et al. (2010).

	HLC/(M atm <sup>-1</sup> )	Issues
oxalic	10 <sup>6.8 a</sup>	Origin not clear. Effective instead of intrinsic
malonic	10 <sup>6.6 b</sup>	Invalid application of Eq. (7). Too high $p_S^0$
succinic	10 <sup>9.4 b</sup>	Gross extrapolation of $p^0$ from high $T$
glutaric	10 <sup>8.3 c</sup>	HLC not experimental
adipic	10 <sup>8.3 b</sup>	Gross extrapolation of $p^0$ from high $T$

Obtained from:

<sup>a</sup> Gaffney et al. (1987)

<sup>b</sup> Meylan and Howard (2000)

<sup>c</sup> Hilal et al. (2008)

Furthermore, Gaffney et al. (1987) state that this HLC is an effective one at pH = 4, while the basis set of the GROMHE method should consist of intrinsic HLCs according to Raventos-Duran et al. (2010). For diacids, the relation between effective and intrinsic HLC is

$$k_h^{\text{eff}} = k_h^{\text{intr}} \left( 1 + \frac{K_{a,1}}{[\text{H}^+]} + \frac{K_{a,1}K_{a,2}}{[\text{H}^+]^2} \right). \quad (8)$$

Given that oxalic acid is a quite strong acid, with dissociation constants of  $K_{a,1} = 5.18 \times 10^{-2}$ ,  $K_{a,2} = 5.30 \times 10^{-5}$  (Apelblat, 2002), it will be mostly in ionised form at this pH, and  $k_h^{\text{eff}}/k_h^{\text{intr}} \approx 800$ .

The HLC of glutaric acid was taken from another compilation (Hilal et al., 2008). From there, one can follow the trace via the compilation of Modarresi et al. (2007) to the compilation of Sander (1999) and finally to the original source, Saxena and Hildemann (1996), which makes clear that this value is not experimental but rather estimated through a group-contribution method. This notion was explicitly mentioned in the compilation of Sander (1999), but was lost in the compilations of Modarresi et al. (2007); Hilal et al. (2008), and Raventos-Duran et al. (2010) directly or indirectly referring to it. We emphasise here that these last three compilations were made explicitly to develop and/or test HLC estimation methods; hence, it is required that the compiled HLC values are all experimental rather than estimated!

The HLCs of the other three diacids (malonic, succinic, and adipic acids) are all taken from the data compilation available in the EPI suite software (Meylan and Howard, 2000). Checking this compilation, it turns out that in all cases, the HLC does not refer to a directly measured value, but rather to an estimation obtained by combining solubility  $c_s^{\text{sat}}$  of the solid diacid and solid state vapour pressure  $p_{\text{Cr}}^0$ , using Eq. (7). This is not necessarily a problem; Eq. (7) is correct provided the solubility is low, such that one is close to the IDL. Of course both  $c_s^{\text{sat}}$  and  $p_{\text{Cr}}^0$  should be reliable to obtain a reliable  $k_h$ . The solubilities  $c_s^{\text{sat}}$  are all from the AQUASOL

database (Yalkowsky and Dannenfelser, 1992) and they reasonably agree with the values from primary references (Apelblat and Manzurola, 1987, 1989; Marcolli et al., 2004). However, the solubility mole fraction  $x^{\text{sat}}$  of malonic acid is about 0.22 (Apelblat and Manzurola, 1987), which is far from IDL, hence Eq. (7) does not hold. The  $x^{\text{sat}}$  of succinic and adipic acids are much lower (0.013 and 0.003 respectively), hence Eq. (7) should be reasonably valid.

However, for all these three diacids, the values for  $p_{\text{Cr}}^0$  are questionable. The values for succinic and adipic acids are from the handbook of Yaws (1994). In this compilation, the vapour pressures of these two acids refer to the liquid phase; the vapour pressures at 25 °C present in the compilation of Meylan and Howard (2000) are obtained by first extrapolating liquid vapour pressures over 157 and 128 K below their melting points respectively, and then converting to solid state vapour pressures. The exact procedure of this conversion is not clear to us, but the reported values  $p_{\text{Cr}}^0$  seem consistent with a simple approximate procedure like that of Yalkowsky (1979) where only the fusion temperature is from experiments, rather than the more precise procedure where both fusion temperature and enthalpy are taken from experiments. Finally, for malonic acid, the vapour pressure originates from the handbook of Jordan (1954). The cited vapour pressure (0.2 Pa) is however orders of magnitude higher than the ones from recent experiments ( $10^{-4}$ – $10^{-3}$  Pa, see e.g. the overview table of Soonsin et al. (2010)).

Given that strong reservations can be made for each of the HLC values in Table 1, the need for a reliable HLC set for diacids is clear.

### 3 Infinite dilution activity coefficients: usefulness in obtaining HLC

Following Eq. (3), HLC can be obtained by knowledge of the liquid vapour pressure,  $p_L^0$ , and the IDAC  $\gamma_s^\infty$ . Note that diacids and hydroxy polyacids are solid at room temperature, whereas  $p_L^0$  and  $\gamma_s^\infty$  are required for the subcooled liquid. Let us disregard the problem of obtaining a subcooled vapour pressure, and focus here on obtaining a reliable value of  $\gamma_s^\infty$ .

For the binary acid–water mixtures considered in this work, activity data is mostly restricted to the water component. In principle, this is not a problem, as it follows from the Gibbs–Duhem relation that for a binary mixture, knowledge of the activity coefficient of one component results in knowledge of the activity coefficient of the other component. The Gibbs–Duhem relation in its derivative and integral form is expressed as (see e.g. Mansoori, 1980)

$$x_w \left( \frac{\partial \ln \gamma_w}{\partial x_w} \right)_{T,P} = x_s \left( \frac{\partial \ln \gamma_s}{\partial x_s} \right)_{T,P}, \quad (9)$$

$$\ln \gamma_s = -\frac{x_w}{x_s} \ln \gamma_w + \int_0^{x_w} \frac{\ln \gamma_w(t)}{(1-t)^2} dt, \quad (10)$$

where  $x_w$  and  $x_s$  denote the mole fraction of water and solute respectively, and  $\gamma_w$  and  $\gamma_s$  are the corresponding activity coefficients. If the  $\gamma_s^\infty$  is desired, Eq. (10) reduces to

$$\ln \gamma_s^\infty = \int_0^1 \frac{\ln \gamma_w(t)}{(1-t)^2} dt; \quad (11)$$

hence, to obtain the IDAC of the solute, any functional form of  $\gamma_w(t)$  should match the experimental water activity coefficients over the entire concentration range.

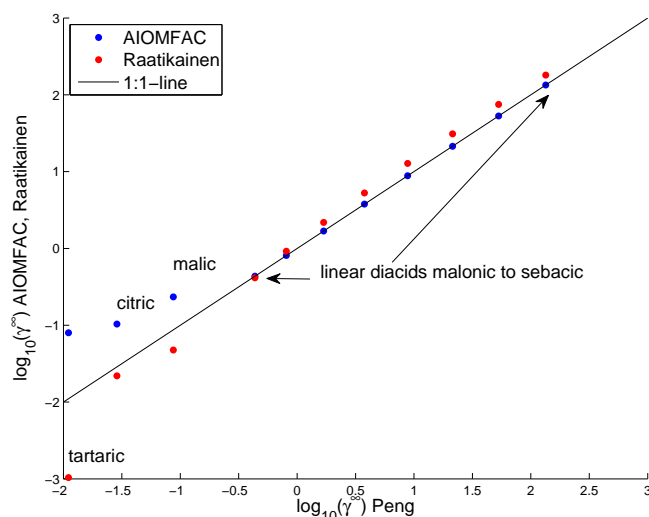
Activity coefficients can be estimated by fitting an activity coefficient expression (e.g. Margules, Van Laar, Wilson, UNIQUAC; see Prausnitz et al., 1999; Carlson and Colburn, 1942) to activity coefficient data of a particular binary system.  $\gamma_s^\infty$  can be derived from the parameters of this mixture-specific model. Another way to obtain the activity coefficient is through the use of group-contribution methods. UNIFAC (Fredenslund et al., 1975; Hansen et al., 1991) is arguably the most popular activity coefficient estimation method based on this group-contribution concept. While a mixture-specific model will generally perform better than a group-contribution model for that specific mixture, its use is limited to that binary mixture. A group-contribution method like UNIFAC, on the other hand, can be used to predict the activity coefficients of more complex mixtures, including molecules for which no experimental data is available.

Peng et al. (2001) and Raatikainen and Laaksonen (2005) provided new UNIFAC parameterisations (called UNIFAC-Peng and UNIFAC-Raatikainen hereafter) using activity and/or solubility data of mixtures with water and diacids or functionalised diacids. A close relative of UNIFAC-Peng is AIOMFAC (Aerosol Inorganic-Organic Mixtures Functional-group Activity Coefficient) (Zuend et al., 2011), as it inherited some of its parameters, while other parameters were inherited from the UNIFAC parameterisation of Marcolli and Peter (2005) (UNIFAC-MP). UNIFAC-MP was adapted to better describe monoalcohols and polyols. It should be mentioned that AIOMFAC has the widest scope of the aforementioned models; it can also describe organic-inorganic and water-inorganic interactions, and is therefore extremely useful for aqueous systems containing both salts and organics.

One would expect that these three models (UNIFAC-Peng, UNIFAC-Raatikainen, AIOMFAC) would give similar IDACs for diacids and hydroxy polyacids, as they were based on experimental data for these molecules. The estimated IDACs are compared in Fig. 1.

### 3.1 Hydroxy polyacids

Unexpectedly, large discrepancies, of up to one order of magnitude, show up for the IDAC of malic, tartaric and citric acids as calculated by the various methods. AIOMFAC predicts a higher IDAC than UNIFAC-Peng for all three molecules. UNIFAC-Raatikainen gives IDAC values quite



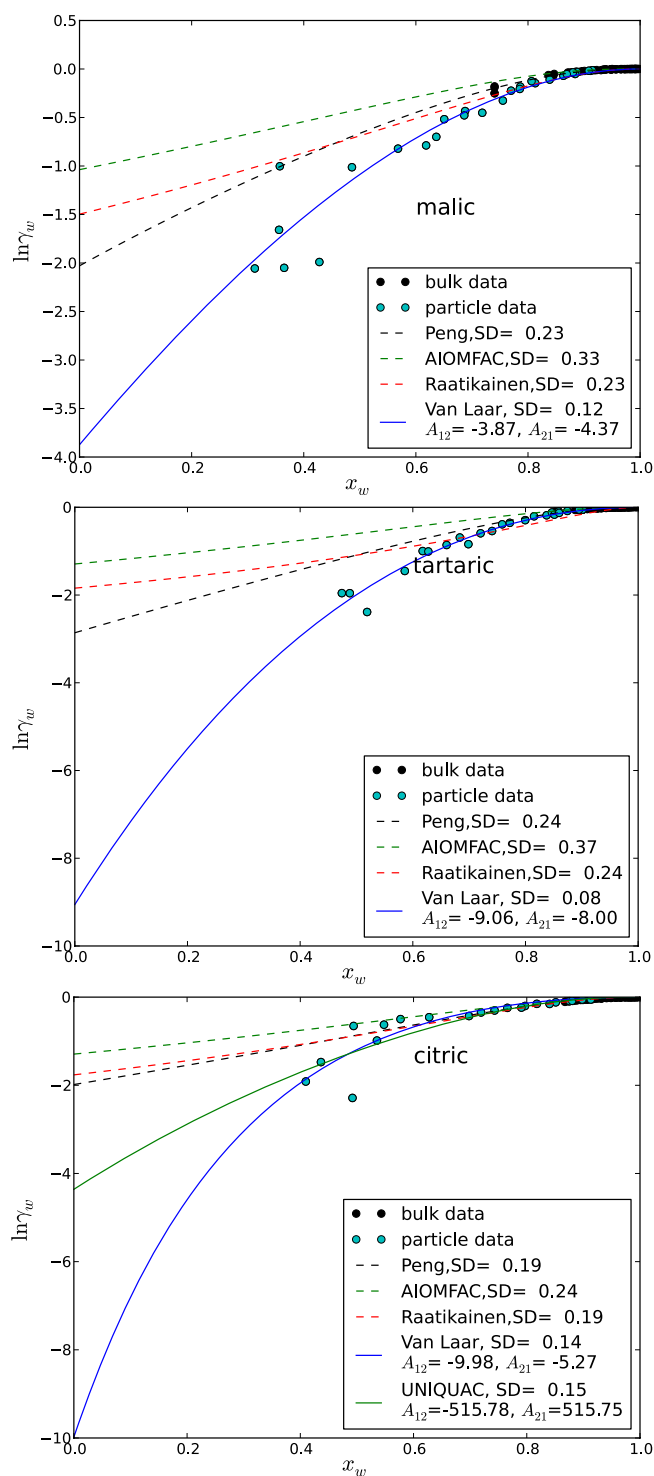
**Fig. 1.** Comparing  $\log_{10} \gamma^\infty$  at 298.15 K as estimated by different methods: AIOMFAC and UNIFAC-Raatikainen vs. UNIFAC-Peng. For the linear diacids, AIOMFAC and UNIFAC-Peng are identical.

close to UNIFAC-Peng for malic and citric acids, but for tartaric acid it predicts an IDAC an order of magnitude lower.

Figure 2 shows experimental  $\ln \gamma_w$  data for the three hydroxy polyacids, mixture-specific fittings and predictions through group-contribution methods. The data in the supersaturation range is from Peng et al. (2001), obtained by electrodynamic balance measurements on particles. Measurements on subsaturated bulk solutions is from several data sources (see Appendix A for an overview). It is clear that the particle data is more scattered and coarse than the bulk data. Some observations that can be made from these plots are as follows:

- for malic and tartaric acids, water activity in the supersaturation range is in the order AIOMFAC > UNIFAC-Peng > experimental;
- for tartaric acid, UNIFAC-Peng and UNIFAC-Raatikainen predict a quite different IDAC  $\ln \gamma_s^\infty$  (Fig. 1). Still, their standard deviation (SD) vs. the experimental data is similar in  $\ln \gamma_w$ . Close inspection reveals that UNIFAC-Peng matches best the bulk  $\ln \gamma_w$  data in the subsaturation range (see Appendix A).

AIOMFAC returns a higher  $\ln \gamma_w$  than UNIFAC-Peng over the entire concentration range for all three hydroxy acids (see Fig. 2). From Eq. (11), this explains the systematically higher IDAC predicted by AIOMFAC compared to UNIFAC-Peng. UNIFAC-Peng shows a lower SD vs. the experimental data compared to AIOMFAC. The reason for the discrepancy between UNIFAC-Peng and AIOMFAC can be attributed to the fact that AIOMFAC's hydroxy-water interaction parameters are from UNIFAC-MP, which was developed for monoalcohols and polyols but not for hydroxy acids. Note however



**Fig. 2.** Logarithm of water activity coefficient at 25 °C in function of mole fraction water, for malic, tartaric and citric acids. Particle experimental data is from Peng et al. (2001), bulk experimental data is from several sources (see Appendix A). Fitted activity coefficient expressions (Van Laar, UNIQUAC) are also included, as well as estimations by UNIFAC-Peng, AIOMFAC and UNIFAC-Raatikainen. The standard deviation vs. the experimental data is also given.

that in the subsaturation range (see Appendix A) AIOMFAC matches better the experimental data than UNIFAC-Peng for tartaric acid, while UNIFAC-Peng matches better for malic and citric acids.

Although UNIFAC-Peng has the lowest SD in  $\ln \gamma_w$  for the three acids, it overestimates the  $\ln \gamma_w$  data in the supersaturation region for malic and tartaric acids. Therefore, we fitted the data with the commonly used Margules, Van Laar, Wilson and UNIQUAC (UNIversal QUAsiChemical) activity coefficient expressions (e.g. Prausnitz et al., 1999; Carlson and Colburn, 1942, see also Appendix A). In Fig. 2 we present the most successful and unique fittings (e.g. for these particular cases, the Margules function gave results very close to the Van Laar function). The resulting parameter set is then used to obtain the solute  $\gamma_s^\infty$ . All fittings extrapolate to a lower  $\gamma_w^\infty$  than the group-contribution methods, and different solute  $\gamma_s^\infty$  are obtained: 0.01 for malic acid,  $3 \times 10^{-4}$  for tartaric acid, and  $5 \times 10^{-3}$  or  $1 \times 10^{-2}$  for citric acid if the Van Laar fitting and, respectively, the UNIQUAC fitting is used.

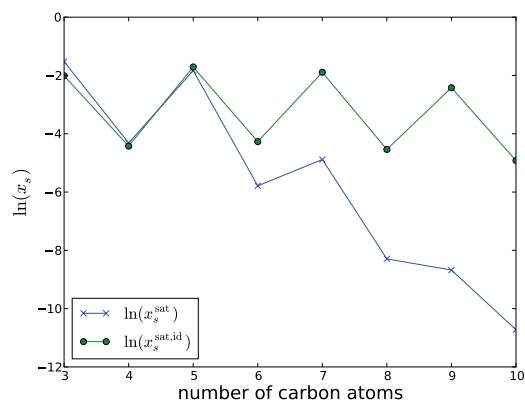
However, as the data in the supersaturation region is scattered and coarse, these fittings are not well constrained. This is shown clearly for citric acid, where the Van Laar and the UNIQUAC fittings have a comparable SD, but are quite different in the supersaturation region. We conclude therefore that these IDAC estimations are not an optimal basis to derived HLCs, even if reliable, subcooled liquid vapour pressures were available.

### 3.2 Linear diacids

For linear diacids, UNIFAC-Peng and AIOMFAC become identical (Fig. 1). UNIFAC-Raatikainen gives an only slightly higher IDAC for the longer chain diacids. However, this does not guarantee that they agree with experiment. Before proceeding further, let us first consider the peculiarities of linear diacid solubilities in more detail.

It is well known that several properties of linear diacids, such as melting point, fusion enthalpy, solubility and solid state vapour pressure, follow an even–odd alternation pattern with the number of carbon atoms in the chain. This is caused by the more stable crystal structure of linear diacids with an even number of carbon atoms (Thalladi et al., 2000). In the case of solubility, this leads to a lower solubility of the diacids with an even number of carbon atoms (Fig. 3). On top of the even–odd alternation pattern, the solubility decreases with the number of carbon atoms. One can view the dissolution of a solute in a solvent as first a melting process and second a mixing process.

$$\ln x_s^{\text{sat}} = \frac{\Delta S_{\text{fus}}(T)}{R} - \frac{\Delta H_{\text{fus}}(T)}{RT} + \frac{\Delta S_{\text{mix}}(T)}{R} - \frac{\Delta H_{\text{mix}}(T)}{RT} \quad (12)$$



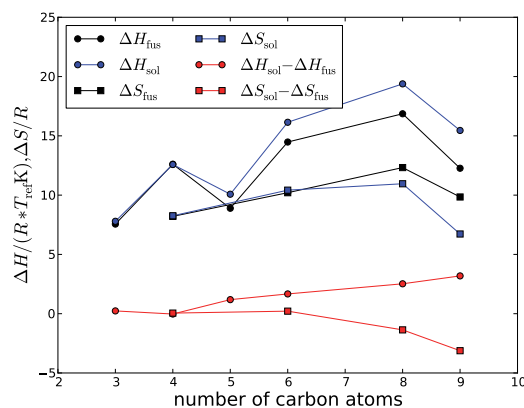
**Fig. 3.** Logarithm of the solubility mole fraction at 298.15 K of linear diacids, and of the ideal solubility mole fraction, as calculated by Eq. (13). Solubility data was taken from Apelblat and Manzurola (1987, 1989, 1990), and Bretti et al. (2006), fusion data from Booth et al. (2010); Roux et al. (2005).

The ideal solubility model (Yalkowsky and Wu, 2010) assumes ideal mixing, such that the finite solubility is only caused by the melting process:

$$\ln x^{\text{sat,id}} = \frac{\Delta S_{\text{fus}}(T)}{R} - \frac{\Delta H_{\text{fus}}(T)}{RT} \approx \frac{\Delta S_{\text{fus}}(T_{\text{fus}})}{R} - \frac{\Delta H_{\text{fus}}(T_{\text{fus}})}{RT} \quad (13)$$

where we used approximation (Eq. 7) and assumed a zero  $\Delta C_p$ . The fusion data was obtained from Booth et al. (2010), and Roux et al. (2005), and (if applicable) the sum over different solid–solid transition points was taken. From Fig. 3, it is clear that apart from the even–odd alternation pattern, there is only a small dependence of the ideal solubility with chain length. Therefore, the lowering of solubility with chain length must be due to a more difficult mixing, or equivalently an increase in activity coefficient. We showed previously (Compennolle et al., 2011) that for the longer chain diacids (starting from C7) UNIFAC-Peng and UNIFAC-Raatikainen underestimate  $\gamma_s^{\text{sat}}$ , which should be close to  $\gamma_s^\infty$  for these low-soluble acids. These longer chain molecules were not in the data set used to develop UNIFAC-Peng or UNIFAC-Raatikainen, and this can explain the lower performance of both methods for these compounds.

A deeper insight in the driving factors behind the solubility of linear diacids can be achieved by comparing fusion enthalpies and entropies with solution enthalpies and entropies close to IDL, as done in Fig. 4. The solution enthalpy and entropy for the low-solubility diacids succinic, adipic, suberic and azelaic acids was derived by fitting Van't Hoff equations to the temperature-dependent solubility data of Apelblat and Manzurola (1987); Yu et al. (2012), and Apelblat and Manzurola (1990). Note that for suberic acid, we chose the data of Yu et al. (2012) over that of Apelblat and Manzurola (1990) as the solubility varied more continuously with tem-



**Fig. 4.** Enthalpies and entropies of fusion and solution for linear diacids, and their differences. Note that for the enthalpies  $\Delta H/(R T_{\text{ref}})$ ,  $T_{\text{ref}} = 298.15 \text{ K}$  is plotted, and for the entropies  $\Delta S/R$ .

perature. Solubility data for pimelic acid is also available (Apelblat and Manzurola, 1989) but the solubility vs. temperature curve is quite irregular; this could mean that several solid–solid transitions take place (e.g. due to uptake of water in the crystal). We therefore omitted the pimelic acid data. Malonic and glutaric are highly soluble in water, hence their solution enthalpy and entropy derived from solubility would be far from IDL. Instead, we took the solution enthalpies derived from calorimetric measurements at low concentrations (Taniewska-Osinska et al., 1990).

Given that the  $\Delta H_{\text{fus}}$  and  $\Delta S_{\text{fus}}$  are obtained at a higher temperature than  $\Delta H_{\text{sol}}$  and  $\Delta S_{\text{sol}}$ , it is not fully justified to simply take their differences to obtain  $\Delta H_{\text{mix}}$  and  $\Delta S_{\text{mix}}$ , but for qualitative purposes it will probably suffice. One notices that from glutaric acid on,  $\Delta H_{\text{sol}} - \Delta H_{\text{fus}} \approx \Delta H_{\text{mix}}$  gradually increases; meaning that the energetic interactions acid–water become less strong compared to the acid–acid interactions in the pure melt. This is unfavourable for the solution process and is one reason why  $x_s^{\text{sat}}$  decreases with chain length. Furthermore, one notices that the entropy of mixing  $\Delta S_{\text{sol}} - \Delta S_{\text{fus}} \approx \Delta S_{\text{mix}}$  decreases, especially for the longer chain molecules, again causing  $x_s^{\text{sat}}$  to decrease. For azelaic acid, this entropic effect has become the dominant contribution to the low solubility.

Such a strongly negative  $\Delta S_{\text{mix}}$  is typical for dissolution of hydrophobic molecules in water. The presence of the hydrophobic chain causes the water molecules to reorder themselves, resulting in an entropy decrease. There are many examples of this effect in the HLC compilation of Abraham et al. (1990), for example for the series of linear 1-alkanols. For linear  $\alpha$ - and  $\omega$ -diacids, and presumably also for linear  $\alpha$ - and  $\omega$ -diols, with both tails hydrogen bonding, it takes a longer chain before this hydrophobic effect becomes important.

**Table 2.**  $x_s^{\text{sat}}$  and  $\gamma_s^\infty/\gamma_s^{\text{sat}}$  data at 25 °C. See Appendix A for the derivation and the used experimental  $a_w$  data. Note that for the linear diacids, UNIFAC-Peng and AIOMFAC return identical results.

Molecule	$x_s^{\text{sat}}$	This work	Clegg <sup>e</sup>	$\gamma_s^\infty/\gamma_s^{\text{sat}}$		
				Peng <sup>f</sup>	AIOMFAC	Raatikainen <sup>f</sup>
oxalic	0.02301 <sup>b</sup>	N/A	1.0 <sup>a</sup>		0.91	1.04
malonic	0.2176 <sup>b</sup>	0.59	0.69		0.74	1.12
succinic	0.01337 <sup>b</sup>	1.20	1.20		1.08	1.10
glutaric	0.16 <sup>c</sup>	3.10	2.87		2.37	2.40
adipic	0.00307 <sup>b</sup>	1.00	N/A		1.08	1.06
pimelic	0.007565 <sup>c</sup>	N/A	N/A		1.33	1.24
suberic	0.0002504 <sup>b</sup>	N/A	N/A		1.01	1.01
azelaic	0.0001706 <sup>b</sup>	N/A	N/A		1.01	1.01
sebacic	$2.2 \cdot 10^{-5}$ <sup>d</sup>	N/A	N/A		1.00	1.00
malic	0.1578 <sup>b</sup>	0.52	0.50	0.51	0.67	0.24
tartaric	0.1435 <sup>b</sup>	0.29	N/A	0.20	0.36	0.01
citric	0.1321 <sup>b</sup>	0.25	N/A	0.29	0.47	0.22
SD( $\log_{10}(\gamma_s^\infty/\gamma_s^{\text{sat}})$ ) <sup>g</sup>				0.10	0.14	0.63

<sup>a</sup> The effect of acid dissociation is significant for oxalic acid. Clegg and Seinfeld (2006a) calculated the activity of the undissociated acid and concluded that, within the data uncertainty, Raoult's law could be assumed. For the other acids the dissociation is a minor effect.

<sup>b</sup> Apelblat and Manzurola (1987).

<sup>c</sup> Apelblat and Manzurola (1989).

<sup>d</sup> Bretti et al. (2006).

<sup>e</sup> The activity calculator for dicarboxylic acid solutions available at the E-AIM website (Clegg and Seinfeld, 2006a, b) was used, taking explicit dissociation into account and considering the activity coefficient of the undissociated acid.

<sup>f</sup> Peng: UNIFAC-Peng. Raatikainen: UNIFAC-Raatikainen.

<sup>g</sup> Standard deviation in  $\log_{10}(\gamma_s^\infty/\gamma_s^{\text{sat}})$  of the group-contribution method vs. the values derived in this work.

#### 4 HLC data: results

From Eq. (7) it follows that HLC can be derived from solubility and solid state vapour pressure data, provided the solubility is low enough such that the IDL is a good approximation. But even if the compounds are quite water soluble, as is the case for the short-chained linear diacids and the hydroxy polyacids, one can still derive the HLC. Indeed, the combination of Eqs. (3) and (4) leads to

$$k_h = \frac{\gamma_s^{\text{sat}}}{\gamma_s^\infty} \frac{c_w x_s^{\text{sat}}}{P_{\text{Cr},s}^0} \quad (14)$$

The ratio  $\gamma_s^{\text{sat}}/\gamma_s^\infty$  can be retrieved if sufficient water activity data is available in the concentration range  $x_s = [0, x_s^{\text{sat}}]$ . From Eq. (10) one derives

$$\ln \frac{\gamma_s^\infty}{\gamma_s^{\text{sat}}} = \frac{1 - x_s^{\text{sat}}}{x_s^{\text{sat}}} \ln \gamma_w(\tilde{x}_w) + \int_{\tilde{x}_w}^1 \frac{\ln \gamma_w(t)}{(1-t)^2} dt, \quad (15)$$

$$\tilde{x}_w = 1 - x_s^{\text{sat}}.$$

The important point is that the integral no longer involves the supersaturation region, where data is typically less precise and/or scarce. Solubility data  $x_s^{\text{sat}}$  (see Table 2) was taken from Apelblat and Manzurola (1987, 1989, 1990), and Bretti et al. (2006), and was consistent with other solubility data

(Marcolli et al., 2004). Functional expressions for  $\ln \gamma_w(t)$  were fitted using only subsaturation water activity data; the details are given in Appendix A and a summary of the results in Table 2.  $\gamma_s^\infty/\gamma_s^{\text{sat}}$  estimations using UNIFAC-Peng, AIOMFAC and UNIFAC-Raatikainen are also presented, as well as estimations from the activity calculator for aqueous dicarboxylic acid solutions available at the E-AIM website (<http://www.aim.env.uea.ac.uk/aim/accent2/>) and described by Clegg and Seinfeld (2006a, b). Similarly to this work (see Appendix A) this activity calculator consists of models individually fitted to specific diacid–water systems. The resulting  $\gamma_s^\infty/\gamma_s^{\text{sat}}$  estimations are therefore very close to our work. Among the group-contribution methods, UNIFAC-Peng gives results closest to our work. UNIFAC-Raatikainen deviates the most, and gives an exceptionally low value in the case of tartaric acid.

For the longer linear chain diacids (C6 and higher), we did not find water activity data in the subsaturation range in the literature. In most cases, their solubility is low enough, such that  $\gamma_s^\infty/\gamma_s^{\text{sat}} \approx 1$  can be assumed. This was confirmed by  $\gamma_s^\infty/\gamma_s^{\text{sat}}$  using UNIFAC or AIOMFAC. A somewhat higher value is predicted only for pimelic acid.

To derive the HLC values, the  $\gamma_s^\infty/\gamma_s^{\text{sat}}$  derived in this work are used for the linear diacids C3–C5 and the hydroxy polyacids. For oxalic acid  $\gamma_s^\infty/\gamma_s^{\text{sat}} = 1$  is assumed following Clegg and Seinfeld (2006a). For pimelic acid, the value

**Table 3.** HLC data at 25 °C and dissolution enthalpies. The last column provides the source of the solid state vapour pressure, unless otherwise indicated. Recommended values are in bold type.

Molecule	$k_h$ Matm <sup>-1</sup>	$-\Delta H_{g \rightarrow aq}$ kJ mol <sup>-1</sup>	Source for $p_{Cr}^0$
Linear diacids			
oxalic	$6.0 \times 10^6$		Booth et al. (2010)
	$5.2 \times 10^8$		Soonsin et al. (2010)
	<b><math>6.2 \times 10^8</math></b>		Soonsin et al. (2010) <sup>a</sup>
malonic	$3.6 \times 10^9$	73	Booth et al. (2010)
	$2.6 \times 10^{10}$	88	Soonsin et al. (2010)
	<b><math>3.9 \times 10^{10}</math></b>	<b>92</b>	Soonsin et al. (2010) <sup>a</sup>
	<b><math>9.4 \times 10^9</math></b>	<b>113</b>	Cappa et al. (2008)
	$4.0 \times 10^9$	73	Bilde et al. (2003)
succinic	$3.1 \times 10^9$	90	Ribeiro da Silva et al. (1999) <sup>c</sup>
	$5.6 \times 10^8$	62	Booth et al. (2010)
	$1.0 \times 10^{10}$		Soonsin et al. (2010)
	<b><math>4.2 \times 10^9</math></b>	<b>94</b>	Soonsin et al. (2010) <sup>a</sup>
	<b><math>2.0 \times 10^9</math></b>	<b>97</b>	Cappa et al. (2007) <sup>b</sup>
	$4.6 \times 10^8$	88	Chattopadhyay and Ziemann (2005)
	$1.4 \times 10^9$	107	Bilde et al. (2003)
glutaric	$9.8 \times 10^8$	81	Salo et al. (2010)
	$1.7 \times 10^9$	90	Ribeiro da Silva et al. (2001) <sup>c</sup>
	$6.9 \times 10^8$	98	Booth et al. (2010)
	$6.1 \times 10^9$		Soonsin et al. (2010)
	<b><math>5.2 \times 10^9</math></b>	<b>97</b>	Soonsin et al. (2010) <sup>a</sup>
	<b><math>2.4 \times 10^9</math></b>	<b>109</b>	Cappa et al. (2007) <sup>b</sup>
	$7.2 \times 10^8$	107	Chattopadhyay and Ziemann (2005)
	$3.2 \times 10^8$	66	Bilde et al. (2003)
adipic	$3.4 \times 10^8$	76	Salo et al. (2010)
	$1.1 \times 10^9$	92	Ribeiro da Silva et al. (1999) <sup>c</sup>
	$1.9 \times 10^9$		Mentel et al. (2004) <sup>e</sup>
	$2.9 \times 10^9$	79	Booth et al. (2010)
	<b><math>6.7 \times 10^9</math></b>	<b>105</b>	Cappa et al. (2007) <sup>b</sup>
	$5.7 \times 10^8$	106	Chattopadhyay and Ziemann (2005)
	$1.3 \times 10^9$	114	Bilde et al. (2003)
pimelic	$3.0 \times 10^8$	57	Salo et al. (2010)
	<b><math>8.2 \times 10^9</math></b>	<b>121</b>	Cappa et al. (2007) <sup>b</sup>
	$1.3 \times 10^9$	92	Chattopadhyay and Ziemann (2005)
	$3.3 \times 10^8$	115	Bilde et al. (2003)
suberic	$1.8 \times 10^9$	129	Salo et al. (2010)
	$2.6 \times 10^9$	105	Ribeiro da Silva et al. (1999) <sup>c</sup>
	<b><math>7.8 \times 10^9</math></b>	<b>120</b>	Cappa et al. (2007) <sup>b</sup>
	$4.3 \times 10^8$	100	Chattopadhyay and Ziemann (2005)
	$9.2 \times 10^8$	136	Bilde et al. (2003)
azelaic	$1.0 \times 10^8$	53	Salo et al. (2010)
	<b><math>9.0 \times 10^9</math></b>	<b>140</b>	Cappa et al. (2007) <sup>b,d</sup>
	$1.3 \times 10^8$	100	Chattopadhyay and Ziemann (2005)
	$1.1 \times 10^8$	115	Bilde et al. (2003)
sebacic	$2.0 \times 10^7$	58	Salo et al. (2010)
	$1.5 \times 10^9$	118	Ribeiro da Silva et al. (1999) <sup>c</sup>
	<b><math>7.7 \times 10^9</math></b>		Cappa et al. (2007) <sup>b</sup>
	$8.4 \times 10^7$		Chattopadhyay and Ziemann (2005)
	$1.4 \times 10^7$		Salo et al. (2010)
hydroxy polyacids			
malic	$2.7 \times 10^{10}$		Booth et al. (2010)
tartaric	$1.6 \times 10^{10}$		Booth et al. (2010)
	<b><math>7 - 93 \times 10^{16}</math></b>		Huisman et al. (2013) <sup>f</sup>
	$1.7 - 22 \times 10^{17}$		Huisman et al. (2013) <sup>g</sup>
citric	$2.1 - 28 \times 10^{15}$		Huisman et al. (2013) <sup>h</sup>
	$7.9 \times 10^{10}$		Booth et al. (2010)
	<b><math>&gt; 2 - 60 \times 10^{16}</math></b>		Huisman et al. (2013) <sup>f</sup>
	$> 1.5 - 42 \times 10^{16}$		Huisman et al. (2013) <sup>g</sup>
	$> 0.6 - 18 \times 10^{16}$		Huisman et al. (2013) <sup>h</sup>

<sup>a</sup> Results using  $p_{Cr}^0$  from a saturated solution.

<sup>b</sup>  $p_{Cr}^0$  data measured at relatively high  $T$ : from 318–358 K for succinic acid to 353–385 K for sebacic acid.

<sup>c</sup>  $p_{Cr}^0$  data measured at relatively high  $T$ : from 339–357 K for malonic acid to 367–377 K for azelaic acid.

<sup>d</sup> The  $p_{Cr}^0$  value of  $1.0 \times 10^8$  Pa in Table 1 of Cappa et al. (2007) is likely a typo. Comparing with  $\Delta H_{sub}$  and  $\Delta S_{sub}$  in the same table reveals that the value  $1.0 \times 10^7$  Pa should be taken.

<sup>e</sup> This value is not derived from  $p_{Cr}^0$  data but is directly measured.

<sup>f</sup> This value is not derived from  $p_{Cr}^0$  data, but from  $p_L^0$  data, using Eq. 3 and  $\gamma_s^\infty$  estimated in Section 3.1. The spread originates from the uncertainty in Huisman et al. (2013)'s  $p_L^0$  data. The true uncertainty will be higher due to uncertainty in  $\gamma_s^\infty$ .

<sup>g</sup> Estimated using Eqs. (5), (7) and (14), with fusion properties taken from Booth et al. (2010) and assuming  $\Delta C_{p,ls} \approx 0$ .

<sup>h</sup> As <sup>g</sup>, but assuming  $\Delta C_{p,ls} \approx \Delta S_{fus}(T_{fus})$ .

estimated by UNIFAC-Peng is used. For the other linear diacids (C6,C8–C10)  $\gamma_s^\infty/\gamma_s^{sat} = 1$  is adopted given their low solubility.

In Table 3, the HLC data at 25 °C derived from Eq. (15) or Eq. (7) are presented. As especially the solid state vapour pressures disagree between different sources, we grouped the data in Table 3 per solid state vapour pressure reference. Where possible, also the enthalpy of gas-phase dissolution is given, calculated as  $\Delta H_{g \rightarrow aq} = \Delta H_{sol} - \Delta H_{sub}$ .

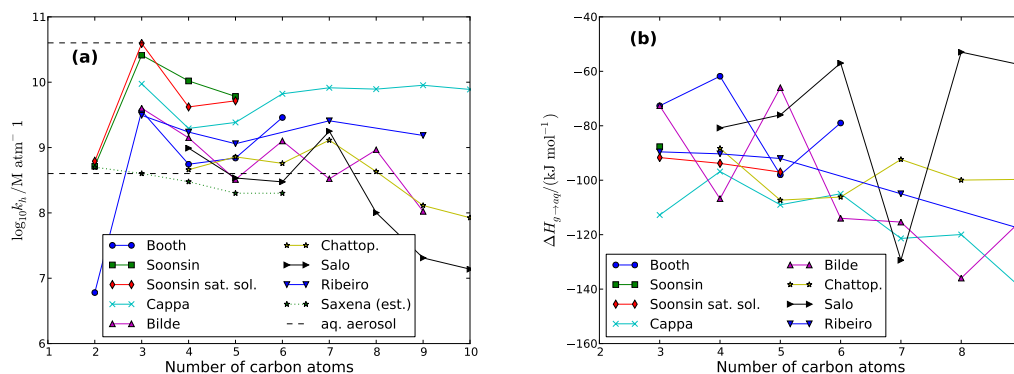
Huisman et al. (2013) provide liquid phase vapour pressures rather than solid state vapour pressures for tartaric and citric acids. In a first approach, we applied Eq. (3), using the IDACs from the fittings discussed in Sect.3. In a second approach,  $k_h$  was estimated combining Eqs. (5), (7) and (14), using fusion properties taken from Booth et al. (2010) and assuming  $\Delta C_{p,ls} \approx 0$  or  $\Delta C_{p,ls} \approx \Delta S_{fus}(T_{fus})$ . Both approaches return reasonably consistent results, although for tartaric acid, the result of the second approach depends strongly on the assumption for  $\Delta C_{p,ls}$ .

## 5 Discussion and conclusions

### 5.1 Consistency of solid state vapour pressure data

If other homologous series (linear alkanes, acids, 1-alkanols, 2-ketones, etc.) are any guide (Sander, 1999), one would expect a rather slow variation of the HLC of linear diacids with chain length compared to e.g. liquid vapour pressure  $p_L^0$ . For example, when going from acetic to hexanoic acid, HLC at 25 °C is lowered by a factor 4, while  $p_L^0$  is lowered roughly by a factor 400. Also, no even–odd alternation of  $k_h$  or gas dissolution enthalpy  $\Delta H_{g \rightarrow aq}$  with chain length is expected, as this is a peculiarity for properties involving the crystalline phase. Figure 5 presents the HLC and  $\Delta H_{g \rightarrow aq}$  of the linear diacids vs. carbon number, grouped per reference of solid state vapour pressure. The large variation in  $k_h$  reflects the variation in  $p_{Cr}^0$  from different data sources. Some of the lowest  $k_h$  and, in absolute value,  $\Delta H_{g \rightarrow aq}$  are found for Salo et al. (2010), especially for the longer chains C8–C10 where  $k_h$  lowers rapidly with chain length. This is likely due to samples that are not purely crystalline, a possibility acknowledged by these authors. For pimelic acid Salo et al. (2010) could distinguish two modes and they attributed the one with the lowest  $p^0$  to the crystalline phase. This is probably correct, as for this acid the derived HLC and  $\Delta H_{g \rightarrow aq}$  are more comparable to these derived from  $p_{Cr}^0$  data of other authors. To a smaller extent, also the HLC derived from the Chattopadhyay and Ziemann (2005) data lowers rapidly from C7 on. The dissolution enthalpies derived from the data of Booth et al. (2010) and of Bilde et al. (2003) exhibit a strong even–odd alternation – although in reverse directions – and contrary to expectation. This could be an indication of experimental artefacts in the measurement of  $\Delta H_{sub}$  in these works. Also the HLC data derived from Bilde et al. (2003)





**Fig. 5.**  $\log_{10} k_h$  (a) and  $\Delta H_{g \rightarrow aq}$  (b) of linear diacids, grouped per data source of solid state vapour pressure. The range where partitioning to aqueous aerosol, based on water content only, is neither complete nor negligible, is also indicated.

exhibit an even–odd alternation. The HLC data derived from Ribeiro da Silva et al. (1999, 2001) and Cappa et al. (2007, 2008) exhibit the smallest variation with chain length, more in line with the expectation. We recommend the HLC derived from the data of Cappa et al. (2007, 2008), as their measurement was closer to room temperature compared to that from Ribeiro da Silva et al. (1999, 2001). The HLC derived from the  $p_{Cr}^0$  measurements on saturated solutions of Soonsin et al. (2010) are also recommended; these authors make a convincing case that these are preferable over  $p_{Cr}^0$  measurements on the solid state.

Apart from diacids, Booth et al. (2010) also presented  $p_{Cr}^0$  data on hydroxy polyacids (malic, tartaric and citric acids). Using fusion enthalpy data, these data were then converted to subcooled liquid  $p_L^0$ . From these data, it followed that  $p_L^0(\text{tartaric}) > p_L^0(\text{succinic})$ , and  $p_L^0(\text{citric}) > p_L^0(\text{adipic})$ , which is counterintuitive, as one expects generally a lower  $p_L^0$  with increasing number of polar groups. However, one could argue that for molecules with many functional groups, it is difficult for the molecules to get optimal intermolecular bonding for all functional groups at once. Comparing  $k_h$  instead of  $p_L^0$  can provide a more stringent test; the small water molecules should more easily interact with all functional groups at once. From Table 3, one finds that  $k_h(\text{tartaric}) > k_h(\text{succinic})$  and  $k_h(\text{citric}) > k_h(\text{adipic})$ , which seems to be at least qualitatively correct. We recommend however the HLC derived from the data of Huisman et al. (2013). In this work, the same technique is used as Soonsin et al. (2010) used for linear diacids, and the expected order  $p_L^0(\text{tartaric}) < p_L^0(\text{succinic})$ , and  $p_L^0(\text{citric}) < p_L^0(\text{adipic})$  is preserved. The derived HLCs are about 6–7 orders of magnitude higher than those derived from the Booth et al. (2010) data.

## 5.2 Atmospheric implications

Notwithstanding the high variations in the derived HLC of the linear diacids, they are most often higher than the estima-

tions provided by the review work of Saxena and Hildemann (1996). For clouds, the liquid water content (LWC) varies between 0.1 and 1 g m<sup>-3</sup>, and for aqueous aerosols between 10<sup>-6</sup> and 10<sup>-4</sup> g m<sup>-3</sup> (Ervens et al., 2011). If partitioning between gas and aqueous phase is governed solely by Henry's law, the aqueous phase fraction,  $f_{aq}$ , of a species can be calculated from

$$f_{aq} = \frac{1}{k^*/k_h + 1}, \text{ with } k^* = \frac{\rho_w}{LWC} \frac{1}{RT}, \quad (16)$$

with  $\rho_w$  as the water density. For clouds,  $k^*$  is between  $4 \times 10^4$  and  $4 \times 10^5$  Matm<sup>-1</sup>. For oxalic acid, the lowest  $k_h$  value from Table 3 is  $6.0 \times 10^6$  Matm<sup>-1</sup>, leading to  $f_{aq}$  between 0.94 and 0.993. Taking also the acid dissociation of oxalic acid into account at a typical pH of 4 (Eq. 8),  $f_{aq}$  is above 0.9999. The other  $k_h$  values for oxalic acid from Table 3 are about two orders of magnitude higher, leading to an even more complete dissolution. For the other species in Table 3,  $k_h$  varies between 10<sup>8</sup> and 10<sup>11</sup> Matm<sup>-1</sup> (provided one dismisses the lowest values from Salo et al. (2010) corresponding probably to non-purely crystalline samples) orders of magnitude higher than  $k^*$ . Hence, for clouds, the diacids and hydroxy polyacids should reside almost completely in the aqueous phase.

For aqueous aerosols,  $k^*$  is typically between  $4 \times 10^8$  and  $4 \times 10^{10}$  Matm<sup>-1</sup>, which is in the range of  $k_h$  values from Table 3 for linear diacids. To the extent that the HLCs reported here are applicable, one can conclude that for linear diacids significant partitioning to the aqueous phase or the gas phase are both possible, depending on the species and the LWC. However, an aqueous aerosol is not a dilute aqueous solution, but on the contrary a concentrated solution containing both organics and inorganics. Therefore, in a more rigorous treatment, an activity coefficient model (e.g. AIOMFAC, Zuend et al., 2011) should be used, provided the mixture composition is known.

*Acknowledgements.* This work was supported by the project BIOSOA (SD/CS/05A, 2011–2014) funded by the Belgian Science Policy Office.

Edited by: G. McFiggans

## References

- Abraham, M. H., Whiting, G. S., Fuchs, R., and Chambers, E. J.: Thermodynamics of solute transfer from water to hexadecane, *J. Chem. Soc. Perk. T.*, 2, 291–300, doi:10.1039/p29900000291, 1990.
- Apelblat, A.: Dissociation constants and limiting conductances of organic acids in water, *J. Mol. Liq.*, 95, 99–145, 2002.
- Apelblat, A. and Manzurola, E.: Solubility of oxalic, malonic, succinic, adipic, maleic, malic, citric, and tartaric acids in water from 278.15 to 338.15 K, *J. Chem. Thermodyn.*, 19, 317–320, 1987.
- Apelblat, A. and Manzurola, E.: Solubility of ascorbic, 2-furancarboxylic, glutaric, pimelic, salicylic, and o-phthalic acids in water from 279.15 to 342.15 K, and apparent molar volumes of ascorbic, glutaric, and pimelic acids in water at 298.15 K, *J. Chem. Thermodyn.*, 21, 1005–1008, 1989.
- Apelblat, A. and Manzurola, E.: Solubility of suberic, azelaic, levulinic, glycolic, and diglycolic acids in water from 278.25 K to 361.35 K, *J. Chem. Thermodyn.*, 22, 289–292, 1990.
- Apelblat, A., Dov, M., Wisniak, J., and Zabicky, J.: The vapour pressure of water over saturated aqueous solutions of malic, tartaric, and citric acids, at temperatures from 288 K to 323 K, *J. Chem. Thermodyn.*, 27, 35–41, 1995a.
- Apelblat, A., Dov, M., Wisniak, J., and Zabicky, J.: Osmotic and activity coefficients of HO<sub>2</sub>CCH<sub>2</sub>C(OH)(CO<sub>2</sub>H)CH<sub>2</sub>CO<sub>2</sub>H (citric acid) in concentrated aqueous solutions at temperatures from 298.15 K to 318.15 K, *J. Chem. Thermodyn.*, 27, 347–353, 1995b.
- Bilde, M., Svenningsson, B., Mønster, J., and Rosenørn, T.: Even-odd alternation of evaporation rates and vapor pressures of C<sub>3</sub>–C<sub>9</sub> dicarboxylic acid aerosols, *Environ. Sci. Technol.*, 37, 1371–1378, 2003.
- Booth, A. M., Barley, M. H., Topping, D. O., McFiggans, G., Garforth, A., and Percival, C. J.: Solid state and sub-cooled liquid vapour pressures of substituted dicarboxylic acids using Knudsen Effusion Mass Spectrometry (KEMS) and Differential Scanning Calorimetry, *Atmos. Chem. Phys.*, 10, 4879–4892, doi:10.5194/acp-10-4879-2010, 2010.
- Bretti, C., Crea, F., Foti, C., and Sammartano, S.: Solubility and activity coefficients of acidic and basic nonelectrolytes in aqueous salt solutions. 2. Solubility and activity coefficients of suberic, azelaic, and sebacic acids in NaCl(aq), (CH<sub>3</sub>)<sub>4</sub>NCl(aq), and (C<sub>2</sub>H<sub>5</sub>)<sub>4</sub>NI(aq) at different ionic strengths and at *t* = 25 °C, *J. Chem. Eng. Data*, 51, 1660–1667, doi:10.1021/je060132t, 2006.
- Cappa, C. D., Lovejoy, E. R., and Ravishankara, A. R.: Determination of evaporation rates and vapor pressures of very low volatility compounds: a study of the C<sub>4</sub>–C<sub>10</sub> and C<sub>12</sub> dicarboxylic acids, *J. Phys. Chem. A*, 111, 3099–3109, doi:10.1021/jp068686q, 2007.
- Cappa, C. D., Lovejoy, E. R., and Ravishankara, A. R.: Evidence for liquid-like and nonideal behavior of a mixture of organic aerosol components, *P. Natl. Acad. Sci. USA*, 105, 18687–18691, doi:10.1073/pnas.0802144105, 2008.
- Carlo, M. J.: Thermodynamic quantities of some biochemically important organic acids in aqueous solutions at 25 °C, Ph.D. thesis, Texas A&M University, 1971.
- Carlson, H. C. and Colburn, A. P.: Vapor-liquid equilibria of nonideal solutions, *Ind. Eng. Chem.*, 34, 581–589, doi:10.1021/ie50389a013, 1942.
- Chattopadhyay, S. and Ziemann, P. J.: Vapor pressures of substituted and unsubstituted monocarboxylic and dicarboxylic acids measured using an improved thermal Desorption Particle Beam Mass Spectrometry Method, *Aerosol Sci. Tech.*, 39, 1085–1100, doi:10.1080/02786820500421547, 2005.
- Clegg, S. L. and Seinfeld, J. H.: Thermodynamic models of aqueous solutions containing inorganic electrolytes and dicarboxylic acids at 298.15 K. 1. The acids as nondissociating components., *J. Phys. Chem. A*, 110, 5692–5717, doi:10.1021/jp056149k, 2006b.
- Clegg, S. L. and Seinfeld, J. H.: Thermodynamic models of aqueous solutions containing inorganic electrolytes and dicarboxylic acids at 298.15 K. 2. Systems including dissociation equilibria., *J. Phys. Chem. A*, 110, 5718–5734, doi:10.1021/jp056150j, 2006b.
- Compennolle, S., Ceulemans, K., and Müller, J.-F.: Estimating fusion properties for functionalised acids, *Atmos. Chem. Phys.*, 11, 8385–8394, doi:10.5194/acp-11-8385-2011, 2011.
- Ervens, B., Turpin, B. J., and Weber, R. J.: Secondary organic aerosol formation in cloud droplets and aqueous particles (aqSOA): a review of laboratory, field and model studies, *Atmos. Chem. Phys.*, 11, 11069–11102, doi:10.5194/acp-11-11069-2011, 2011.
- Davies, M. and Thomas, D. K.: Isopiestic Studies of Aqueous Dicarboxylic Acid Solutions, *J. Phys. Chem.*, 60, 41–44, doi:10.1021/j150535a011, 1956.
- Fredenslund, A., Jones, R. L., and Prausnitz, J. M.: Group-contribution estimation of activity-coefficients in nonideal liquid-mixtures, *AIChE J.*, 21, 1086–1099, 1975.
- Gaffney, J. S. and Senum, G. I.: Peroxides, peracids, aldehydes, and PANs and their links to natural and anthropogenic organic sources, in: *Gas-Liquid Chemistry of Natural Waters*, edited by: Newman, L., vol. 1, NTIS TIC-4500, UC-11, BNL 51757 Brookhaven National Laboratory, 5–1–5–7, 1984.
- Gaffney, J. S., Streit, G. E., Spall, W. D., and Hall, J. H.: Beyond acid rain. Do soluble oxidants and organic toxins interact with SO<sub>2</sub> and NO<sub>x</sub> to increase ecosystem effects?, *Environ. Sci. Technol.*, 21, 519–524, doi:10.1021/es00160a001, 1987.
- Hansen, H. K., Rasmussen, P., Fredenslund, A., Schiller, M., and Gmehling, J.: Vapor-liquid-equilibria by UNIFAC group contribution. 5. Revision and extension, *Ind. Eng. Chem. Res.*, 30, 2352–2355, 1991.
- Hilal, S. H., Ayyampalayam, S. N., and Carreira, L. A.: Air-liquid partition coefficient for a diverse set of organic compounds: Henry's Law constant in water and hexadecane, *Environ. Sci. Technol.*, 42, 9231–9236, doi:10.1021/es8005783, 2008.
- Huisman, A. J., Krieger, U. K., Zuend, A., Marcolli, C., and Peter, T.: Vapor pressures of substituted polycarboxylic acids are much lower than previously reported, *Atmos. Chem. Phys.*, 13, 6647–6662, doi:10.5194/acp-13-6647-2013, 2013.
- Jordan, T.: *Vapor Pressure of Organic Compounds*, Interscience Publishers, inc., 250 Fifth Avenue, New York, NY, 1954.

- Levien, B. J.: A Physicochemical Study of Aqueous Citric Acid Solutions., *J. Phys. Chem.*, 59, 640–644, doi:10.1021/j150529a016, 1955.
- NIST Chemistry WebBook, NIST Standard Reference Database Number 69, edited by: Linstrom, P. J. and Mallard, W. G., National Institute of Standards and Technology, Gaithersburg MD, 20899, <http://webbook.nist.gov>.
- Maffia, M. C. and Meirelles, A. J. A.: Water activity and pH in aqueous polycarboxylic acid systems, *J. Chem. Eng. Data*, 46, 582–587, doi:10.1021/jc0002890, 2001.
- Mansoori, G. A.: Classical thermodynamic basis of activity coefficients: predictive and consistency rules for binary and ternary mixtures based on the relation between excess Gibbs free energies of (c)- and (c-1)-component mixtures, *Fluid Phase Equilib.*, 4, 197–209, 1980.
- Marcolli, C. and Peter, Th.: Water activity in polyol/water systems: new UNIFAC parameterization, *Atmos. Chem. Phys.*, 5, 1545–1555, doi:10.5194/acp-5-1545-2005, 2005.
- Marcolli, C., Luo, B., and Peter, T.: Mixing of the organic aerosol fractions: liquids as the thermodynamically stable phases, *J. Phys. Chem. A*, 108, 2216–2224, doi:10.1021/jp036080l, 2004.
- Mentel, T. F., Folkers, M., Tillmann, R., Henk H., Wahner, A., Otjes, R., Blom, M., and ten Brink, H. M.: Determination of the Henry coefficients for organic aerosol components, *Geophys. Res. Abstr.*, 6, 1525, 2004.
- Meylan, W. M. and Howard, P. H.: Src's epi suite, v3.20, Syracuse Research Corporation, Syracuse, New York, 2000.
- Modarresi, H., Modarress, H., and Dearden, J. C.: QSPR model of Henry's law constant for a diverse set of organic chemicals based on genetic algorithm-radial basis function network approach, *Chemosphere*, 66, 2067–2076, 2007.
- Peng, C., Chan, M. N., and Chan, C. K.: The hygroscopic properties of dicarboxylic and multifunctional acids: measurements and UNIFAC predictions, *Environ. Sci. Technol.*, 35, 4495–4501, 2001.
- Prausnitz, J. M., Lichtenthaler, R. N., and de Azevedo, E. G.: *Molecular Thermodynamics of Fluid-Phase Equilibria*, 3rd Edn., Prentice-Hall PTR, Upper Saddle River, New Jersey 07458, 1999.
- Raatikainen, T. and Laaksonen, A.: Application of several activity coefficient models to water-organic-electrolyte aerosols of atmospheric interest, *Atmos. Chem. Phys.*, 5, 2475–2495, doi:10.5194/acp-5-2475-2005, 2005.
- Raventos-Duran, T., Camredon, M., Valorso, R., Mouchel-Vallon, C., and Aumont, B.: Structure-activity relationships to estimate the effective Henry's law constants of organics of atmospheric interest, *Atmos. Chem. Phys.*, 10, 7643–7654, doi:10.5194/acp-10-7643-2010, 2010.
- Ribeiro da Silva, M. A. V., Monte, M. J. S., and Ribeiro, J. R.: Vapour pressures and the enthalpies and entropies of sublimation of five dicarboxylic acids, *J. Chem. Thermodyn.*, 31, 1093–1107, 1999.
- Ribeiro da Silva, M. A. V., Monte, M. J. S., and Ribeiro, J. R.: Thermodynamic study on the sublimation of succinic acid and of methyl- and dimethyl-substituted succinic and glutaric acids, *J. Chem. Thermodyn.*, 33, 23–31, 2001.
- Robinson, R. A., Smith, P. K., and Smith, E. R. B.: The osmotic coefficients of some organic compounds in relation to their chemical constitution, *Trans. Faraday Soc.*, 38, 63–70, doi:10.1039/TF9423800063, 1942.
- Roux, M. V., Temprado, M., and Chickos, J. S.: Vaporization, fusion and sublimation enthalpies of the dicarboxylic acids from C<sub>4</sub> to C<sub>14</sub> and C<sub>16</sub>, *J. Chem. Thermodyn.*, 37, 941–953, 2005.
- Salo, K., Jonsson, Å. M., Andersson, P. U., and Hallquist, M.: Aerosol volatility and enthalpy of sublimation of carboxylic acids, *J. Phys. Chem. A*, 114, 4586–4594, doi:10.1021/jp910105h, 2010.
- Sander R.: *Compilation of Henry's Law Constants for Inorganic and Organic Species of Potential Importance in Environmental Chemistry (Version 3)*, <http://www.henrys-law.org>, 1999.
- Saxena, P. and Hildemann, L. M.: Water-soluble organics in atmospheric particles: a critical review of the literature and application of thermodynamics to identify candidate compounds, *J. Atmos. Chem.*, 24, 57–109, 1996.
- Soonsin, V., Zardini, A. A., Marcolli, C., Zuend, A., and Krieger, U. K.: The vapor pressures and activities of dicarboxylic acids reconsidered: the impact of the physical state of the aerosol, *Atmos. Chem. Phys.*, 10, 11753–11767, doi:10.5194/acp-10-11753-2010, 2010.
- Taniewska-Osinska, S., Tkaczyk, M., and Apelblat, A.: Enthalpies of dilution and solution of succinic, malonic, and glutaric acids in water at 298.15 K, *J. Chem. Thermodyn.*, 22, 715–720, 1990.
- Thalladi, V. R., Nüsse, M., and Boese, R.: The Melting Point Alternation in  $\alpha$ ,  $\omega$ -Alkanedicarboxylic Acids, *J. Am. Chem. Soc.*, 122, 9227–9236, doi:10.1021/ja0011459, 2000.
- Velezmoro, C. E. and Meirelles, A. J. A.: Water activity in solutions containing organic acids, *Dry. Technol.*, 16, 1789–1805, doi:10.1080/07373939808917496, 1998.
- Wise, M. E.: Hygroscopic growth of ammonium sulfate/dicarboxylic acids, *J. Geophys. Res.*, 108, 4638, doi:10.1029/2003JD003775, 2003.
- Yalkowsky, S. H.: Estimation of entropies of fusion of organic compounds, *Ind. Eng. Chem. Fundam.*, 18, 108–111, doi:10.1021/i160070a003, 1979.
- Yalkowsky, S. H. and Dannenfelser, R. M.: AQUASOL database of aqueous solubility. Version 5., College of Pharmacy, University of Arizona, Tucson, AZ, 1992.
- Yalkowsky, S. H. and Wu, M.: Estimation of the ideal solubility (crystal-liquid fugacity ratio) of organic compounds., *J. Pharm. Sci.*, 99, 1100–1106, doi:10.1002/jps.21897, 2010.
- Yaws, C. L.: *Handbook of Vapor Pressure*, Gulf Publishing Company, Houston, 1994.
- Yu, Q., Ma, X., and Gao, W.: Determination of the solubility, dissolution enthalpy and entropy of suberic acid in different solvents, *Fluid Phase Equilib.*, 330, 44–47, 2012.
- Zuend, A., Marcolli, C., Booth, A. M., Lienhard, D. M., Soonsin, V., Krieger, U. K., Topping, D. O., McFiggans, G., Peter, T., and Seinfeld, J. H.: New and extended parameterization of the thermodynamic model AIOMFAC: calculation of activity coefficients for organic-inorganic mixtures containing carboxyl, hydroxyl, carbonyl, ether, ester, alkenyl, alkyl, and aromatic functional groups, *Atmos. Chem. Phys.*, 11, 9155–9206, doi:10.5194/acp-11-9155-2011, 2011.

## Appendix A

### Obtaining the activity coefficient ratios $\gamma^\infty/\gamma^{\text{sat}}$

#### A1 Procedure

The water activity data of Peng et al. (2001), in tabulated form, was taken from <http://ihome.ust.hk/~keckchan/hygroscopic.html>. The data of Levien (1955) for citric acid were also taken from this site. The data from Velezmoro and Meirelles (1998); Apelblat et al. (1995a, b); Marcolli et al. (2004), and Maffia and Meirelles (2001) were taken directly from the original works. All other data (Davies and Thomas, 1956; Wise, 2003; Robinson et al., 1942) were taken not from the original work but from the E-AIM site (<http://www.aim.env.uea.ac.uk/aim/accnt2/tables.php>), with the exception of the tartaric acid data of Robinson et al. (1942), which was taken from the original work. We tested the following functional expressions to fit  $\ln \gamma_w$  (see e.g. Prausnitz et al., 1999):

$$\text{Margules: } f(x_1, A_{12}, A_{21}) = (A_{12} + 2(A_{21} - A_{12})x_1)x_2^2, \quad (\text{A1})$$

$$\text{Van Laar: } f(x_1, A_{12}, A_{21}) = A_{12} \left( \frac{A_{21}x_2}{A_{12}x_1 + A_{21}x_2} \right)^2, \quad (\text{A2})$$

$$\text{Wilson: } f(x_1, A_{12}, A_{21}) \\ = -\ln(x_1 + A_{12}x_2) + x_2 \left( \frac{A_{12}}{x_1 + A_{12}x_2} - \frac{A_{21}}{A_{21}x_1 + x_2} \right) \quad (\text{A3})$$

$$\text{UNIQUAC: } f(x_1, A_{12}, A_{21}) = \ln \frac{\Phi_1}{x_1} + \frac{z}{2} q_1 \ln \frac{\theta_1}{\Phi_1} + \Phi_2 \left( l_1 - \frac{r_1}{r_2} l_2 \right) \\ - r_1 \left[ \ln(\theta_1 + \theta_2 e^{-A_{21}/T}) + \theta_2 \left( \frac{e^{-A_{21}/T}}{\theta_1 + \theta_2 e^{-A_{21}/T}} - \frac{e^{-A_{12}/T}}{\theta_2 + \theta_1 e^{-A_{12}/T}} \right) \right] \quad (\text{A4})$$

with  $A_{12}$  and  $A_{21}$  the parameters to fit. For a detailed overview of the quantities used in the UNIQUAC equation, we refer to Prausnitz et al. (1999). The best fitting function  $f$  was then chosen, with the lowest SD from the  $\ln \gamma_w$  data. If the  $\ln \gamma_w$  data would cover the entire concentration range and  $f$  fits the data well, the resulting parameters could be used to estimate  $\ln \gamma_s^\infty$  as  $f(x_s = 0, A_{21}, A_{12})$ . However, as we only used data below and up to the solubility limit, the parameters will not be well enough constrained to this end. Instead, we use the parameters to estimate the activity coefficient ratio  $\gamma_s^\infty/\gamma_s^{\text{sat}}$ :

$$\ln \frac{\gamma_s^\infty}{\gamma_s^{\text{sat}}} = f(x_s = 0, A_{21}, A_{12}) - f(x_s = x_s^{\text{sat}}, A_{21}, A_{12}). \quad (\text{A5})$$

Clegg and Seinfeld (2006a, b) already performed an extensive analysis on the activity of dicarboxylic acids (C2–C5) and malic acid. Moreover, the collected data and fitted activity models are presented at the E-AIM site (<http://www.aim.env.uea.ac.uk/aim/accnt2/>). These models allow us to calculate  $\gamma_s^\infty/\gamma_s^{\text{sat}}$  and are in good agreement with our analysis. This gives us confidence in the validity of

our approach to obtain the  $\gamma_s^\infty/\gamma_s^{\text{sat}}$  values for tartaric acid and citric acid, which were not considered by Clegg and Seinfeld (2006a).

#### A2 Importance of acid dissociation

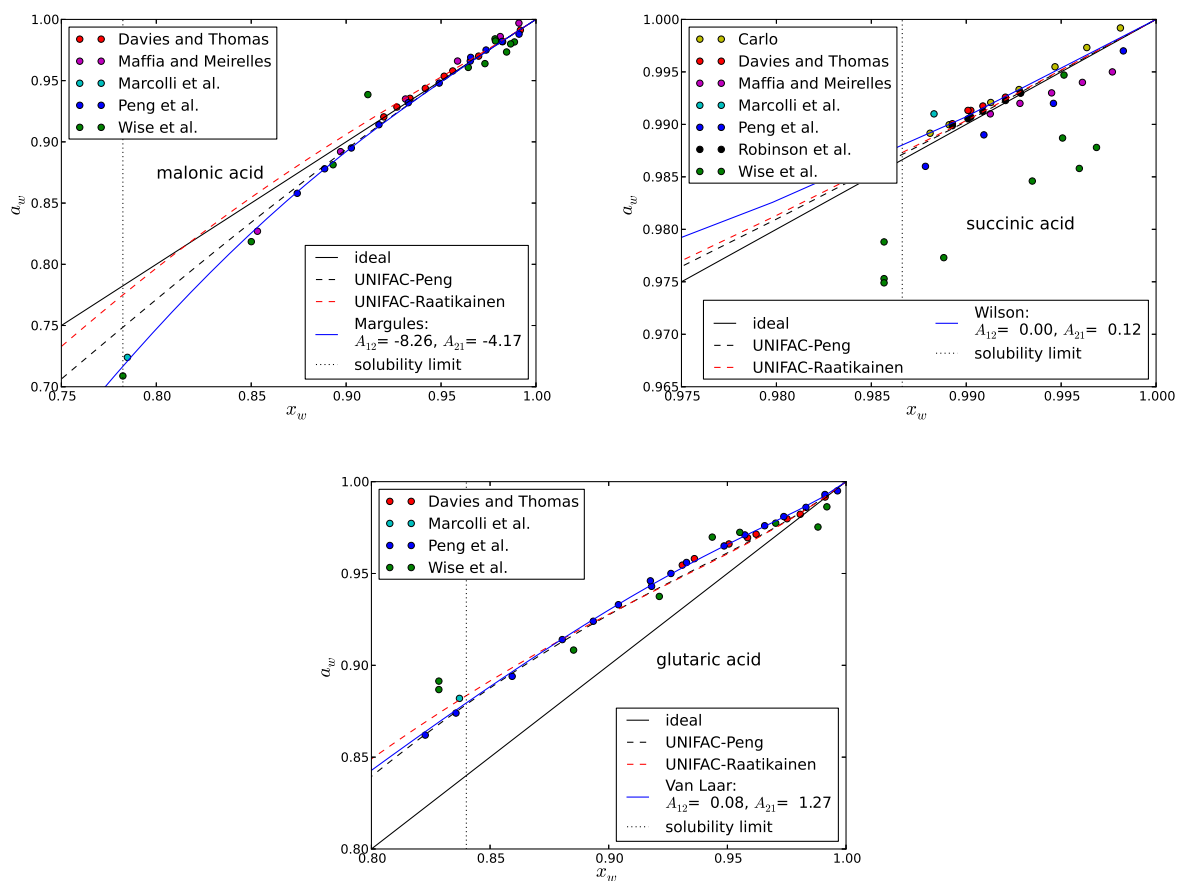
The dissociation of the acids could in principle affect the water activity data. To derive the intrinsic  $k_h$ , one needs  $\gamma_s^\infty/\gamma_s^{\text{sat}}$  of the undissociated acid. However, for most acid–water mixtures considered here, even the most dilute ones, acid dissociation is a minor phenomenon. This was confirmed with the activity calculator at the E-AIM site, where one can choose to take the acid dissociation explicitly into account or not; the effect on  $\gamma_s^\infty/\gamma_s^{\text{sat}}$  is very small. An exception is oxalic acid, for which we followed the analysis of Clegg and Seinfeld (2006a). In all other cases, we neglected acid dissociation, but still the derived  $k_h$  should reflect reasonably intrinsic Henry's law constants.

#### A3 Linear diacids

*Oxalic acid.* Acid dissociation is significant in the subsaturation range. There is considerable uncertainty in the water activity data and Clegg and Seinfeld (2006a) concluded that Raoult's law could be assumed for the undissociated acid. Therefore,  $\gamma_s^\infty/\gamma_s^{\text{sat}} = 1$ .

*Malonic acid.* Water activity data for malonic acid was taken from Peng et al. (2001); Davies and Thomas (1956); Maffia and Meirelles (2001); Wise (2003), and Marcolli et al. (2004) (Fig. A1). Not all data are of similar precision; especially at high dilution, scatter can become important. The isopiestic data set of Davies and Thomas (1956) has a relatively high precision and is therefore fully retained. The data of Wise (2003) and of Peng et al. (2001) for  $x_w > 0.95$ , and of Maffia and Meirelles (2001) for  $x_w > 0.93$  was not used. The data point of Wise (2003) at  $x_w = 0.91$  is considered an outlier and was therefore also not used. The Margules function could best fit the data. The resulting activity coefficient ratio is  $\gamma_s^\infty/\gamma_s^{\text{sat}} = 0.59$ . Note that, after oxalic acid, malonic acid is the strongest acid considered here. According to the dicarboxylic acid activity calculator at the E-AIM site, even for the most dilute mixtures (molality  $\sim 0.5$ ) 94% of the acid is in undissociated form, and  $\gamma_s^\infty/\gamma_s^{\text{sat}} = 0.69$  regardless if the dissociation is taken into account or not.

*Succinic acid.* Water activity data was taken from Peng et al. (2001); Davies and Thomas (1956); Maffia and Meirelles (2001); Wise (2003); Marcolli et al. (2004), and Robinson et al. (1942). This diacid has a much lower solubility compared to malonic acid. The data of Wise (2003) deviates considerably and was omitted. Most data indicate  $a_w \approx x_w$ , implying that  $\gamma_s^\infty/\gamma_s^{\text{sat}}$  will be close to unity. If we select the data as recommended by E-AIM, i.e. the more reliable isopiestic data of Davies and Thomas (1956) and of Robinson et al. (1942), and part of the data of Carlo



**Fig. A1.** Water activity of malonic, succinic and glutaric acids in the subsaturation region. Mixture-specific fitted expressions and group-contribution activity estimations are also indicated.

(1971), a fitting with the Wilson function finally results in  $\gamma_s^\infty/\gamma_s^{\text{sat}} = 1.2$ .

*Glutaric acid.* The data of Wise (2003) are very scattered and are therefore not used. The other data show clearly that  $a_w > x_w$ . After fitting with the Van Laar formula, an activity coefficient ratio of  $\gamma_s^\infty/\gamma_s^{\text{sat}} = 3.1$  is obtained.

*Adipic acid.* The solubility of adipic acid is very low, such that  $\gamma_s^\infty/\gamma_s^{\text{sat}} \approx 1$  can be anticipated. This is confirmed by the data point of Marcolli et al. (2004), where  $a_w \approx x_w$  at the solubility limit.

#### A4 Hydroxy polyacids

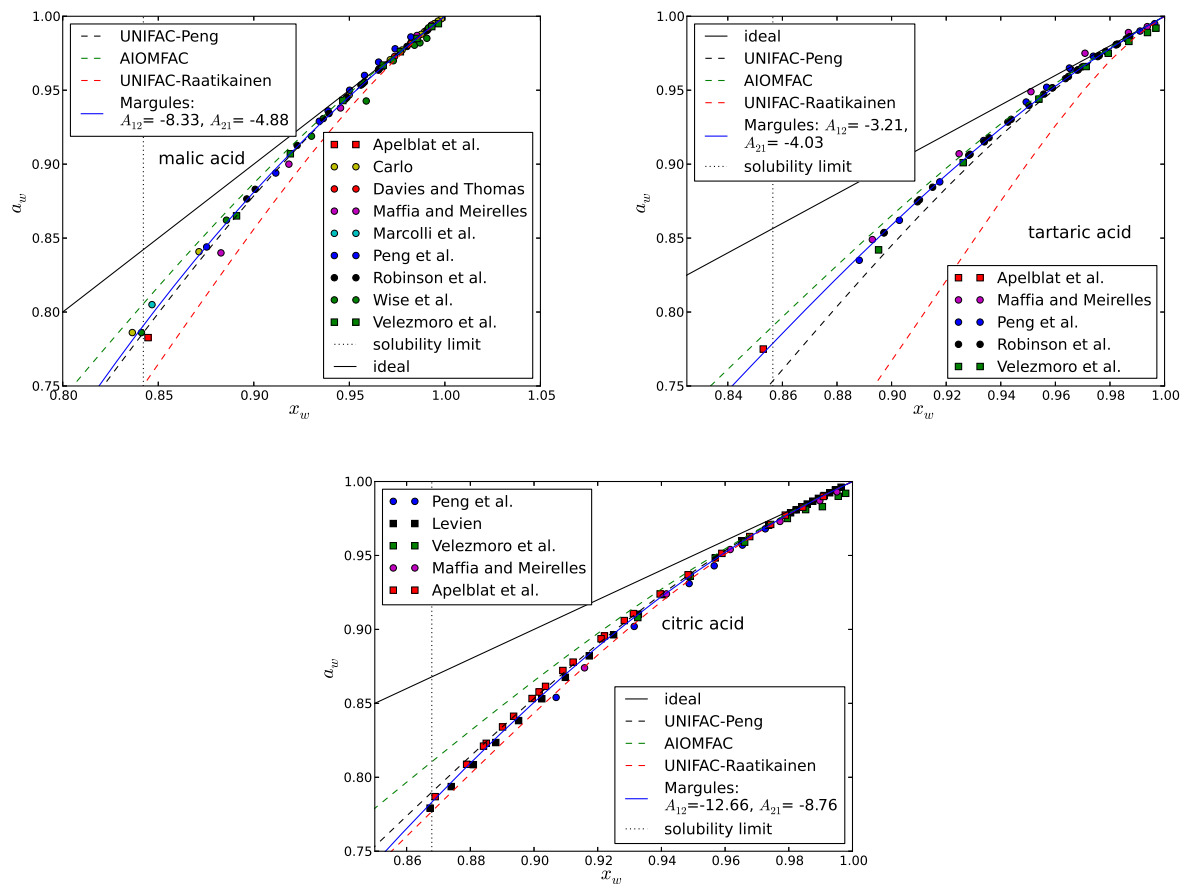
For malic, tartaric and citric acids, one has  $a_w \leq x_w$  (Fig. A2). Note that only malic acid was considered by Clegg and Seinfeld (2006a).

*Malic acid.* We selected all the data of Davies and Thomas (1956); Carlo (1971); Apelblat et al. (1995a), and Robinson et al. (1942), while from the data of Wise (2003); Peng et al. (2001); Velezmoro and Meirelles (1998) we selected only the  $a_w \leq 0.95$  points. The data of Maffia and Meirelles (2001) was excluded as the  $a_w$  data was lower than for the other

data sources. Fitting with the Margules function resulted in  $\gamma_s^\infty/\gamma_s^{\text{sat}} = 0.52$ .

*Tartaric acid.* We selected all the data of Apelblat et al. (1995a), and Robinson et al. (1942). We selected only the  $a_w \leq 0.95$  points for the Maffia and Meirelles (2001) data and the  $a_w \leq 0.97$  points for the Velezmoro and Meirelles (1998) data. The data of Velezmoro and Meirelles (1998) was excluded as the  $a_w$  data was lower than for the other data sources.

*Citric acid.* The data of the different data sources (Levien, 1955; Peng et al., 2001; Maffia and Meirelles, 2001; Velezmoro and Meirelles, 1998) are in good agreement with each other. We only excluded the data points of Velezmoro and Meirelles (1998) at  $a_w \geq 0.983$ .



**Fig. A2.** Water activity of malic, tartaric and citric acids in the subsaturation region. Mixture-specific fitted expressions and group-contribution activity estimations are also indicated.