Solubility Study and Thermodynamic Modelling of Succinic Acid and Fumaric Acid in Bio-based Solvents

Pablo López-Porfiri[1], Patricia Gorgojo[1,2,3], and María Gonzalez-Miquel[1,4]*

[1] Department of Chemical Engineering, Faculty of Science and Engineering, The University of Manchester, Manchester M13 9PL, UK.

[2] Nanoscience and Materials Institute of Aragón (INMA) CSIC-Universidad de Zaragoza, C/ Mariano Esquillor s/n, 50018 Zaragoza, Spain.

[3] Chemical and Environmental Engineering Department, Universidad de Zaragoza, C/ Pedro Cerbuna 12, 50009 Zaragoza, Spain.

[4] Departamento de Ingeniería Química Industrial y del Medioambiente, ETS Ingenieros Industriales, Universidad Politécnica de Madrid, C/ José Gutiérrez Abascal 2, Madrid, Spain.

[*] Corresponding author. E-mail: maria.gonzalezmiquel@upm.es

Highlights

- Solubilities of two important organic acids for the bio-refinery industry have been measured in several bio-based green solvents and correlated.
- Activity coefficients and excess energies at the saturation concentration were estimated from the solid-liquid equilibria computed by the COSMO-RS method.
- Systems molecular interactions were analyzed throughout the excess enthalpy contributions: electrostatic, hydrogen bonding, and van der Waals forces interactions.
- The free Gibbs energy, enthalpy, and entropy of solution of all systems were estimated at the experimental harmonic mean temperature.

ABSTRACT

The solubility and involved energies of organic acids in green solvents are relevant to the design of sustainable biorefinery downstream processes. In this work, the solubility of two important bio-based organic acids such as succinic acid and fumaric acid, in water and four bio-based solvents (i.e., ethyl acetate, 1,8-cineole, cyclopentyl methyl ether, and 2-methyl tetrahydrofuran) were measured within a temperature range of [283 – 313] K. A gravimetric methodology was adopted, previously validated using the organic acid aqueous solubilities available in the literature. The reported data present an average estimated uncertainty of measurement, with a level of confidence of 95%, of $5.2 \cdot 10^{-4}$ mol·mol^{-1}. Experimental results were correlated with the van't Hoff equation, and the Buchowski–Ksiazaczak λh model, where the root mean squared deviations were less than $3.9 \cdot 10^{-4}$ for all systems. From the experimental data and the COSMO-RS molecular simulation method, the solid-liquid equilibria were modelled to estimate the excess energies and the solute activity coefficients of the saturated solutions. Excess enthalpy contribution analysis shows that attractive hydrogen-bonding interactions between the organic acids and the green solvents drive the dissolution phenomena. The magnitude of the hydrogen-bonding interactions increases with temperature for all systems, agreeing with the observed solubility trends. The organic acid energies of solution were estimated from the van't Hoff equation, demonstrating an enthalpy-entropy compensation effect. The energetic analysis shows that the dissolution phenomenon is an enthalpy-driven process for fumaric acid, whereas no predominant effect is observed for succinic acid.

Keywords: Organic acid, Green solvent, Solubility, COSMO-RS, Solution energy.

1. Introduction

Nowadays, a global urge to reduce fossil fuel consumption along with the increasing governmental environmental regulations encourages green routes in chemical manufacturing [1]. In this context, the production of bio-based molecules from renewable sources in the biorefinery industry is key to addressing such a challenge [2]. Their use as chemical building blocks allows replacing petrochemical-based compounds in the production of solvents, food, textiles, cosmetics, pharmaceuticals, detergents, as well as other chemical commodities [3].

In particular, the United States Department of Energy has identified bio-based organic acids, among other target building blocks, as strategic platform chemicals [4]. Succinic acid and fumaric acid are two bio-based organic acids that can be produced by fermentation starting from lignocellulose-derived sugars [5,6]. However, despite being acknowledged as a competitive alternative to the petrochemical routes, there is still a need to improve its downstream recovery and purification processes [7]. After fermentation, separation of the bio-based organic acids from the aqueous broths is usually carried out by liquid-liquid extraction (LLX) with volatile organic compounds, such as organophosphorus compounds or aliphatic amines, which is a clear mismatch with their green purpose [6,8]. Moreover, the separation stage is critical, with a cost representing more than two-thirds of the total production cost [9].

Due to the aforementioned reasons, the search for alternative solvents presenting more benign environmental, health and safety features for biotechnological uses is receiving increasing attention [10]. A large number of hydrophobic green novel solvents have been proposed and systematically evaluated to fulfil the current gap in bio-based organic acids recovery [11]. Although several of them have proven their capacity for the extraction task, basic thermodynamic properties such as the solute solubilities and solution energies are still needed to carry out proper

separation and purification process design. It is possible to find in the literature experimental studies on the solubility of succinic acid in several organic solvents, such as isopropanol [12]; water-methanol and water-ethanol mixtures [13]; binary mixtures of cyclohexanone, cyclohexanol, and cyclohexane [14]; binary mixtures of methanol, ethanol, and propanol [15]; as well as on the solubility of fumaric acid in n-propanol, isopropanol, ethanol and acetone [16]. Nevertheless, most of the aforementioned solvents show polar nature, excluding them for LLX from aqueous matrices. Based on the molecular structural information, the quantum chemical COnductor-like Screening Model for Real Solvents (COSMO-RS) method is capable of predicting the chemical potential of the compounds and thereby further thermodynamic properties [17]. COSMO-RS method has been shown to accurately predict the temperature-dependence of the solubility of carboxylic acids in water [18]. Likewise, the method allows identifying solute-solvent affinities by computing mixing enthalpies and molecular interactions, as well as partition coefficient estimations, which have been used to evaluate solvent candidates for sustainable LLX [19,20]. COSMO-RS predictions had been thoroughly assessed by contrasting with experimental data [21]. The average absolute relative deviations (AARD) of infinite dilution activity coefficients resulted in 84.6% and 58.5% for the TZVP-COSMO and TZVPD-FINE parametrization levels, respectively. Likewise, the infinite dilution partial excess enthalpy showed an AARD of 244% and 255% for the above parametrizations. The authors noted that the higher deviations correspond to the systems composed of nitrate-based ionic liquids and heavy nonpolar solutes with high conformational flexibility. This prompted significant discrepancies between the calculated and modelled activity coefficients, which are propagated in the computation of the excess enthalpy. Better accuracy for predictions of activity coefficients at infinite dilution with COSMO-RS has been reported for thiophene in several ionic liquids modelled with a TZVP parametrization,

obtaining an AARD of 24.1% [22]. In an extensive study, the relative mean deviations of the excess enthalpies of 10,851 binary data sets were modelled using COSMO-RS with TZVP parametrization, resulting in a range of 15.5% – 161.8% [23]. On the other hand, the COSMO-RS method predictions of free energies of hydration showed an accuracy of approximately 0.5 kcal·mol^{-1} [24].

The objective of this work is to perform a solubility study of succinic acid and fumaric acid in aqueous solutions and target bio-based solvents. 1,8-Cineole, also known as eucalyptol (Cineole), cyclopentyl methyl ether (CPME), and 2-methyl tetrahydrofuran (2-MeTHF) have been selected based on their good performance in the recovery of such organic acids, as shown in our previous work [11]. We have also included ethyl acetate (EtOAc), a conventional solvent providing high organic acids extraction yields, which has been recently demonstrated that can be also produced from renewable sources [25]. The obtained experimental solubilities have been correlated with thermodynamic models including the van't Hoff equation and the Buchowski–Ksiazaczak λh model. The obtained experimental solubilities have been correlated with thermodynamic models counting with two adjustable parameters each. First, the van't Hoff equation [26], which expresses the entropy and enthalpy effect in the dissolution phenomenon; and secondly, the Buchowski–Ksiazaczak λh model that quantifies the energy of solution and the ideality degree [27].

The molecular interactions, in terms of excess energies, within the saturated mixture were modelled with the COSMO-RS method. In addition, the energies of solution have been estimated from the van't Hoff equation to better understand the organic acids dissolution process phenomenon. Overall, the results derived from this work will contribute to the development and design of green routes for bio-based organic acids recovery, promoting sustainable downstream processes in the biorefinery industry.

2. Materials and methods

2.1 Materials

Succinic acid (≥ 99.0% w/w, CAS: 110-15-6), fumaric acid (≥ 99.0% w/w, CAS: 110-17-8), cineole (99% w/w, CAS: 470-82-6), and 2-MeTHF (≥ 99.0% w/w, CAS: 96-47-9) were purchased from Sigma-Aldrich. EtOAc (99% w/w, CAS: 141-78-6) and CPME (> 99.9% w/w, CAS: 5614-37-9) were purchased from Alfa Aesar. Reagents were used without further purification. Both organic acids were kept in a desiccator to avoid surrounding water absorption due to the hygroscopic effect. High purity water (Milli-Q type 1) was produced in the laboratory. Table 1 shows some of the relevant properties for the experimental and modelling procedures of the compounds used in this work, including the moisture content in the bio-based solvents measured using a Karl Fisher titrator prior to the experiments. The molecular structures of the compounds can be found in Figure 1.

Table 1. Properties of bio-based solvents and organic acids used in this work.

Compound	Abbreviation	MW	BP[a]	MP[b]	Water content[c]
		g·mol^{-1}	T_b / K	T_m / K	wt_w / g·g^{-1}
Water	-	18.02	373.15	-	-
Ethyl acetate	EtOAc	88.11	350.15	-	0.0003
1,8-Cineole	Cineole	154.25	449.15	-	0.0024
Cyclopentyl methyl ether	CPME	100.16	379.15	-	0.0005
2-Methyl tetrahydrofuran	2-MeTHF	86.13	351.15	-	0.0001
Succinic acid	-	118.09	-	455.2	-
Fumaric acid	-	116.07	-	562.55	-

[a] Boiling point according to the suppliers.
[b] Melting point [28].
[c] Measured by Karl Fisher titration at 25ºC using a Metrohm 899 coulometer. Standard uncertainty for water content, u(wt_w) ≤ 10^{-7} g·g^{-1}.

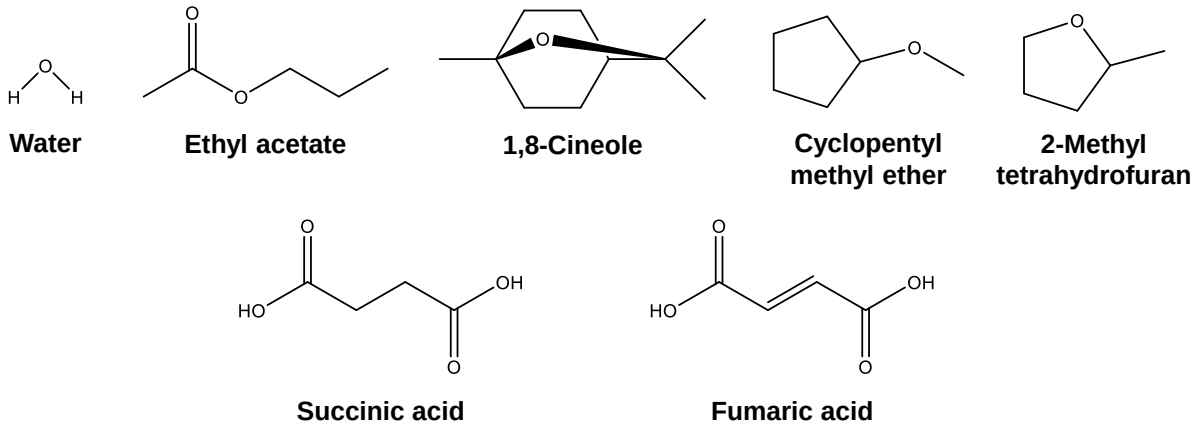


Figure 1. Molecular structures of the bio-based solvents and organic acids studied in this work.

2.2 Experimental procedure

The solubilities of succinic acid and fumaric acid in the bio-based solvents and water were obtained using a gravimetric method, by weighting the solids deposited after the complete evaporation of a saturated solution of [organic acid (1) + solvent (2)]. First, a feed oversaturated solution, i.e., a solution with visible suspended crystals, was prepared at a high temperature (> 60 °C). Then, 2 ml of the feed solution were centrifuged for 3 h in a Sigma 4-16KS refrigerated centrifuge at a

controlled fixed temperature of the data point measurement and a relative centrifugal force (RCF) of 23,506g. Lastly, 1 ml of the saturated supernatant was weighed in a Mettler Toledo MS1045/01 analytical scale in a previously weighted glass vial, and let dry overnight in a vacuum oven at 10 Pa and a temperature set at the boiling point of the solvent. The dry crystallized solid (i.e., just the organic acid) was weighed again to obtain the organic acid mass contained within the saturated supernatant sample. The solubility expressed as the saturated acid mole fraction, x_1 / mol·mol^{-1}, was calculated following Eq. 1, where M_{vial}, M_{dry}, and $M_{solution}$ correspond to the vial mass, the dry organic acid plus the vial mass, and the saturated supernatant plus the vial mass, respectively. The mass standard uncertainty is u(mass) = 1.48·10^{-4} g. Measurements were performed in triplicate, reporting the mean value along with their respective uncertainty of measurement with a level of confidence of 95% ($LC_{95\%}$) calculated according to NIST directions [29].

$$x_1 = \frac{M_{dry}-M_{vial}}{MW_1} \cdot \left[\frac{M_{dry}-M_{vial}}{MW_1} + \frac{M_{solution}-M_{dry}}{MW_2}\right]^{-1} \quad (1)$$

2.3 Solubility correlations

The solid-liquid equilibria of a system composed of a mixture miscible in the liquid phase but completely immiscible in the solid phase can be expressed throughout Eq. 2. According to it, two main factors are relevant in the compound solubility: the saturated solute activity coefficient in the liquid phase, γ_1 / mol·mol^{-1}, to quantify the non-ideality within the system; and the solute enthalpy of fusion, ΔH_m / kJ·mol^{-1}, a temperature-dependent property, which is integrated between the system temperature, T / K, and the solute melting point, T_m / K. Eq. 2 also utilizes the gas constant, R / kJ·mol^{-1}·K^{-1}. Different correlations have been proposed in the literature for an easier approach to expressing the solubility data. These correlations are derived from the analysis and simplifications of Eq. 2, employing empirical parameters (p_i) fitted from experimental data.

$$ln(x_1 \cdot \gamma_1) = \int_T^{T_m} \frac{\Delta H_m(T)}{R} d\left(\frac{1}{T}\right) \quad (2)$$

The **van't Hoff equation** (Eq. 3) is one of the more broadly thermodynamic models used to correlate solubility with temperature. It has two non-dependent temperature empirical parameters, named *A* (p_1) and *B* (p_2) in this work, respectively related to the entropy and enthalpy of solution, hence giving a good overview of the system non-ideality.

$$ln(x_1) = A + \frac{B}{T} \quad (3)$$

The **Buchowski–Ksiazaczak λh model** (Eq. 4) is a two-parameter solubility correlation that also includes the solute melting point, T_m / K. It states λ (p_1) as the system non-ideality measure and *h* (p_2) as the enthalpy of solution. The correlation assumes both parameters as approximately constant with the temperature.

$$ln\left(1 + \frac{\lambda \cdot (1 - x_1)}{x_1}\right) = \lambda \cdot h\left(\frac{1}{T} - \frac{1}{T_m}\right) \quad (4)$$

To fit the experimental solubilities of organic acids reported in this work to the above-described correlations, the minimization of the root mean squared deviation, $\delta(x_1)$, of the N-data points was used as the objective function (OF), showed in Eq. 5, where x_1^{exp} and x_1^{cal} correspond to the experimental data and the data calculated from each solubility correlation, respectively. A MatLab® script has been developed to obtain the respective fitted values of the correlation parameters.

$$OF: min[\delta(x_1)] = min\left[\sqrt{\frac{1}{N}\sum\left(x_1^{exp} - x_1^{cal}\right)^2}\right] \quad (5)$$

2.4 COSMO-RS approach

Thermodynamic phase equilibria can be modelled from the screening charge density, σ / e·nm^{-2}, on the molecular surface by the quantum chemical COnductor-like Screening Model for Real Solvents (COSMO-RS) method [30]. As a statistical thermodynamic-based theory, the COSMO-

RS method improves the COSMO polarization charge densities approach and had proved to be capable to overcome several limitations of the dielectric continuum models [31]. Through this property, the model can compute the chemical potential of the organic acid (1) as a pure compound, μ_1 / kJ·mol^{-1}, and mixed in the α-phase, μ_1^a / kJ·mol^{-1}. For a multi-component system, the solid-liquid equilibrium (SLE) is determined by Eq. 6, considering the free energy of the specie in each phase. Herein, the ΔG_{fus} / kJ·mol^{-1} denotes the compound Gibbs free energy of fusion which can be estimated or provided. To enhance the model predictions, ΔG_{fus} might be obtained from the experimental data given the saturated molar concertation, x_1 / mol·mol^{-1}, at a fixed temperature, T / K. The value of $\Delta G_{fus}(T)$ is iterated until it meets the SLE governed by the experimental saturated mole fraction, x_1, and the chemical potentials, μ_1 and μ_1^a, computed by the COSMO-RS method.

$$\mu_1 + \Delta G_{fus}(T) = \mu_1^a + RTln(x_1) \quad (6)$$

By combining the COSMO-RS method and the experimental results, the equilibria of the system can be solved, allowing to determine the solute activity coefficient, γ_1^{α} / mol·mol^{-1}, of the saturated solution by Eq. 7. Additionally, it is possible to identify specific molecular interactions, i.e., attractive or repulsive forces within the mixtures, by computing the saturated mixture excess energies: excess Gibbs free energy, G^E / kJ·mol^{-1}, excess enthalpy, H^E / kJ·mol^{-1}, and excess entropy expressed as $-TS^E$ / kJ·mol^{-1}, according to Eq. 8. Furthermore, a deeper understanding of these interactions can be obtained from the contributions to H^E. These energies are divided into electrostatic energy (MF), hydrogen bonding (HB), and van der Waals forces (vdW), according to Eq. 9.

$$RTln(\gamma_i^a) = \mu_i^a - \mu_i \quad (7)$$

$$G^E = H^E - TS^E \quad (8)$$

$$H^E = H^E(\mathrm{MF}) + H^E(\mathrm{HB}) + H^E(\mathrm{vdW}) \quad (9)$$

Computational calculations were performed using the *COSMOtherm* software, version C30, release 18.0.2, at the parametrization BP_TZVP_18.

3. Results and discussion

3.1 Methodology validation

The experimental methodology proposed in this work was validated by comparing the solubility of organic acids in water measured herein with the data found in the literature for both succinic acid [15,32–46] and fumaric acid [33,38,42,47,48] within a temperature range of [283 – 333] K. Compiled literature data for each acid were treated as single data sets and fed into the van't Hoff equation (Eq. 3). Experimental data, $ln\left(x_1^{exp}\right)$, was compared with the mean literature values, $ln\left(x_1^{lit}\right)$, evaluated at the corresponding temperature, by the root mean squared deviation, $\delta[ln(x_1)]$. Deviation between the values obtained in this work and those reported in literature are $\delta[ln(x_1)] = 0.19$ and $\delta[ln(x_1)] = 0.20$, for succinic acid and fumaric acid, respectively. Hence, the proposed methodology is in good agreement with the previously reported data. The validation results details can be found in SI.

3.2 Experimental organic acid solubilities

Experimental results obtained in this work for the solubility of succinic acid and fumaric acid using the gravimetric methodology described in section 2.2 are exhibited in Figure 2. The reported data is expressed as the saturated solute mole fraction with the respective data point uncertainty of measurement: $x_1 \pm U_{Comb,95\%}(x_1)$ / mol·mol^{-1}. Succinic and fumaric acids display solubilities ranging [0.0025 – 0.0439] mol·mol^{-1} and [0.0005 – 0.0290] mol·mol^{-1}, respectively. Succinic acid shows a greater solubility than fumaric acid in water, EtOAc, and 2-MeTHF. As will be discussed later, the solubility phenomenon is driven by enthalpic and entropic processes, where the broken and formed molecular interactions define the system behavior. Although both organic acids have

two carboxyl groups (see Figure 1) to form interactions with the solvent molecules, the molecular structure of succinic acid presents three rotatable bonds, making it more flexible. In contrast, fumaric acid has a carbon−carbon double bond in its chain, resulting in a more rigid molecule. The molecular capability to rotate might favour the formation of solute-solvent interactions, hence increasing the solubility.

Saturated succinic acid mole fractions in water and bio-based solvents measured in this work at temperatures in the range [283 – 313] K are plotted in Figure 2.a, along with the solubility in water data compiled from the literature. The solubility values show an overall trend following: EtOAc ≈ CPME < Cineole ≈ Water < 2-MeTHF. Likewise, Figure 2.b displays the solubility values for fumaric acid, which exhibits greater solubility in 2-MeTHF, CPME, and cineole than in water, following the trend: EtOAc ≈ Water < CPME < Cineole < 2-MeTHF. Note that similar trends were obtained in our previous work for the experimental liquid-liquid extraction of both organic acids from model aqueous solutions using the bio-based solvents studied herein [11].

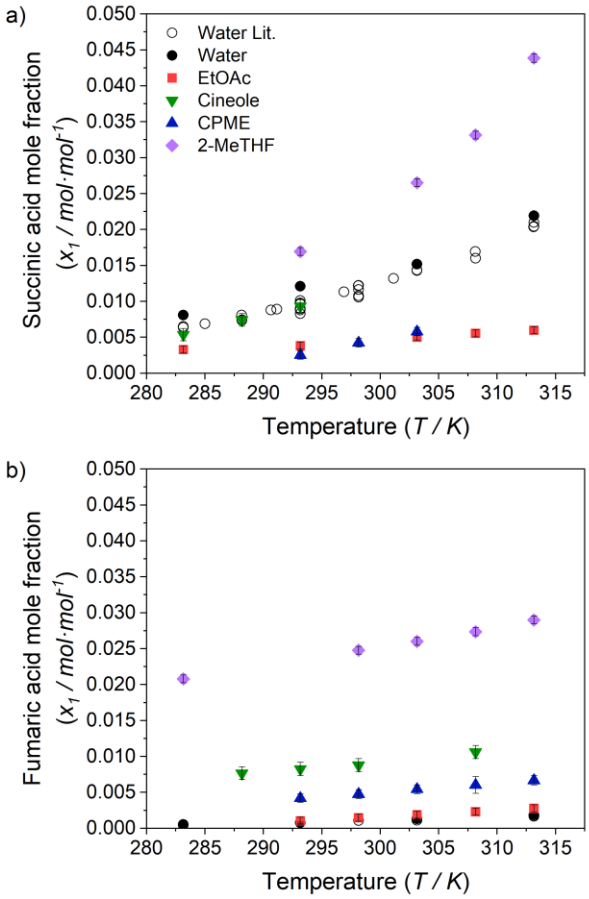


Figure 2. Experimental and literature [15,32,41–48,33–40] solubilities of organic acids in water and bio-based solvents at temperatures ranging [283 – 313] K: a) Succinic acid; b) Fumaric acid. Average combined expanded uncertainty of measurement, $U_{Comb,95\%}(x_1) = 5.2 \cdot 10^{-4}$ mol·mol^{-1}.

Experimental error is composed of uncertainties arising from random and systematic effects. The first one can be related to the variability of the results, while the second is to the measurement procedure. Its proper determination by statistical methods is a subject of discussion among authors, and nonstatistical approaches, such as the mathematical gnostic, have been proposed for the estimation of thermophysical/chemical properties uncertainty [49]. Due to the complexity of the procedure methodology, no more replicate experiments by data point were possible to measure. This results in large confidence bound given by the basic statistical methods estimation of the random variability, which no genuinely reflect the experiment precisions, and the uncertainty should be estimated from other scientific judgment. Since the proposed methodology is based on

reliable independent measurements under highly controlled conditions, the reported errors for the organic acid solubility values are expressed as the combined expanded uncertainty of measurement, $U_{Comb,95\%}(x_1)$ / mol·mol^{-1}. They were estimated from the measured variables and the error propagation with a level of confidence of 95% [29]. The respective variable standard uncertainties and the expression for the combined uncertainty of the organic acid saturated mole fraction derived from Eq. 1 are available in SI. Solubilities of succinic acid and fumaric acid have an $U_{Comb,95\%}(x_1) < 8.7 \cdot 10^{-4}$ mol·mol^{-1} and $U_{Comb,95\%}(x_1) < 9.3 \cdot 10^{-4}$ mol·mol^{-1} respectively, where the average relative error, $100 \cdot U_{Comb,95\%}(x_1)/x_1$, is 11%. Experimental solubility results and uncertainty estimation details can be found in SI.

3.3 Correlation of organic acid solubility data

The following thermodynamic models for solubility: i) the van't Hoff equation (Eq. 3) and ii) the Buchowski–Ksiazaczak λh model (Eq. 4) were used to correlate the solid-liquid equilibria data as a function of the temperature. The empirical parameters were obtained by fitting the equation to the experimental data points and minimizing the objective function (Eq. 5). Table 2 presents the fitted parameters for each [organic acid (1) + solvent (2)] system along with their root mean square deviation, $\delta(x_1)$. Although the modified Apelblat equation shows a slightly better fit, this might be due to overfitting, considering the few data points available within the system.

The percentage of the individual relative deviations, ε_i / % (Eq. 10), for each data point and the predicted saturated solute mole fraction obtained using the solubility correlations can be found in SI. Succinic acid systems show an average relative deviation of 4.6%, with a maximum of 6.6% in the CPME solution correlated with the Buchowski–Ksiazaczak λh model. On the other hand, the fumaric acid solubilities present an overall better fit, with an average relative deviation of 3.9%.

The largest data point deviations are found for the fumaric acid in EtOAc with the van't Hoff equation and Buchowski–Ksiazaczak λh model, where their respective maximum relative deviations are 9.1% and 9.2%.

$$\varepsilon_i = 100 \cdot \left(x_{1,i}^{exp} - x_{1,i}^{cal}\right)/x_{1,i}^{exp} \qquad (10)$$

Table 2. Fitted values of parameters for the van't Hoff equation (Eq. 3), and Buchowski–Ksiazaczak λh model (Eq. 4) for the organic acid solubilities in water and bio-based solvents.

	Succinic acid			Fumaric acid		
Eq.	p_1	p_2	**100·δ(x_1)**	p_1	p_2	**100·δ(x_1)**
			Water			
van't Hoff	5.407	-2894	0.025	6.408	-3973	0.002
λh model	0.3419	8220	0.024	0.5147	7707	0.002
			EtOAc			
van't Hoff	0.8937	-1880	0.005	7.545	-4202	0.003
λh model	0.0227	68940	0.005	1.074	3909	0.003
			Cineole			
van't Hoff	10.64	-4485	0.008	0.2407	-1477	0.006
λh model	2.236	2012	0.008	0.0602	21451	0.005
			CPME			
van't Hoff	17.71	-6927	0.010	1.839	-2145	0.002
λh model	12.41	559.3	0.010	0.1158	17683	0.002
			2-MeTHF			
van't Hoff	11.26	-4509	0.037	-0.4378	-972.8	0.004
λh model	4.410	1045	0.039	0.0424	14738	0.003

3.4 Thermodynamic analysis of organic acid – solvent systems using COSMO-RS

The observed solubility behaviour depends on the molecular interactions present within the system. Such interactions can be quantified from the chemical potential of the species determining the solid-liquid equilibria. To estimate the chemical potential of the compounds, the COSMO-RS method uses the charge density probability distribution of a molecular surface segment, defined as the σ-profile, as the main descriptor. The molecular charge density, or σ-surface, and the respective σ-profile of the compounds used in this work are depicted in Figure 3. Three core regions are recognized: two polar-nature regions, hydrogen bond donor (HBD) at $\sigma < -0.82$ e·nm^{-2} and hydrogen bond acceptor (HBA) at $\sigma > 0.82$ e·nm^{-2}, and a non-polar region within. Based on the experimental data and the σ – profiles of the organic acid solutes and the solvents, the solid-liquid equilibria of each studied system were determined along with their intermolecular affinities.

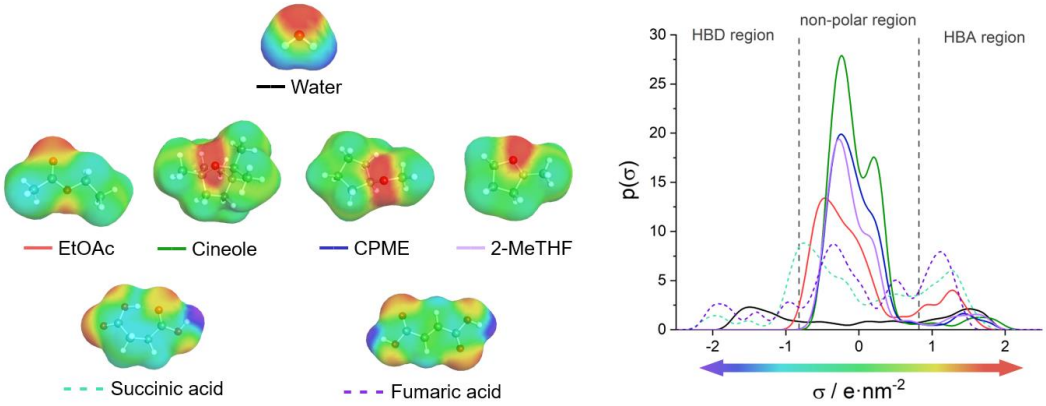


Figure 3. Representation of the σ – surfaces (left) and σ – profiles (right) of the bio-based solvents and organic acids studied in this work.

As one of the main thermodynamic properties of a multicomponent solution, the activity coefficient of each specie provides key insights into the molecular affinity, $\ln(\gamma_1) < 0$, or repulsion, $\ln(\gamma_1) > 0$, i.e., to understand the non-ideal behaviour of a solution. The activity coefficients of the organic acids at the saturated liquid condition are presented in Figure 4 as a temperature function. Note that the organic acid solubility increases with the temperatures, so the γ_1-values are not at the

same concentration. However, the infinite dilution activity coefficients of each acid present similar trends with the temperature than the saturated mixtures. The details of the computed activities coefficients results can be found in SI. The activity coefficients of both organic acids are lower in bio-based solvents than in water, supporting bio-based solvents as good extractants for the recovery of the acids from fermentation broths. From the solvents σ – profiles can be inferred that molecules with a greater non-polar area provide a relatively lower solubility, while those showing greater HBA capacity enhance the solubility of the target solutes. As mentioned above, the molecular rotation capacity of succinic acid allows breaking the internal hydrogen bond depicted in Figure 3, hence promoting interactions with the solvent molecules. The computed γ_1-values results do not follow the experimental solubility, suggesting that more complex interactions are present in the solid solvation phenomena. Regarding the solubility of organic acids in aqueous solutions, the molecule of water presents both hydrogen bond donor and acceptor capacities, which can promote strong interactions with the carboxyl groups of the organic acids. Nevertheless, the limited solubility of fumaric acid in water is explained by the "hydrophobic effect"; this is an entropic phenomenon [50] where the water molecules can form complex tetrahedral networks with water-water interactions stronger than acid-water interactions, requiring high energy to disturb them [51].

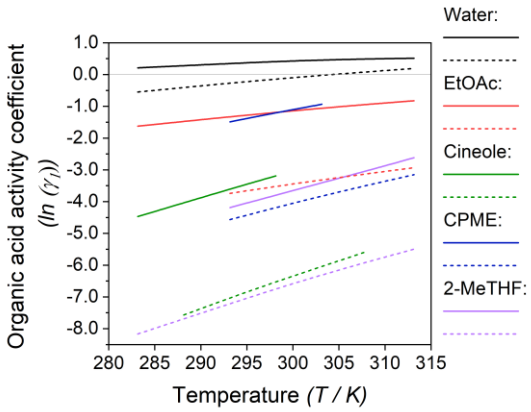


Figure 4. Natural logarithm activity coefficient, ln(γ_1), of saturated succinic acid (solid lines) and fumaric acid (dash lines) in water and bio-based solvents computed by the COSMO-RS method within the experimental temperature range of each [organic acid (1) + solvent (2)] system.

At the solid-liquid equilibria, the specific interactions of the saturated liquid phase can be described through an excess energies analysis. Figure 5 shows the saturated organic acid-solvent mixtures excess energies and their respective contributions to the H^E-value, according to Eq. 9 and Eq. 10, respectively. Within each system, the temperature dependence of the H^E and $-TS^E$ determining the excess G^E remained relatively constant. Thus, the analysis was made at a fixed temperature of 293.15 K to allow a proper comparison. Overall results demonstrate that it is possible to establish a proportional trend between G^E values and the solubility results for each organic acid.

Except for succinic acid in water, G^E values are negative, reflecting the spontaneity of the solvation process. This agrees with the ln(γ_1)-value > 0 for the former system, evidencing its low unlike-molecules affinities. Succinic acid tends to form internal hydrogen bonds, exposing the non-polar surface charges which repulse the water molecules. Most systems show a proportional $-TS^E$ compensation effect with the H^E, suggesting an increasing order within the system due to the new interactions formed. Such compensation effect is predominant for aqueous systems where G^E values tend to zero, denoting the hydrophobic effect described above. In general, the main H^E contribution corresponds to hydrogen bonding interactions between the carboxyl groups of the

solutes and the solvents; this is followed by a slight contribution of electrostatic interactions and a negligible influence of the van der Waals forces. Solute-solvent interactions are stronger for succinic acid than fumaric acid in water and EtOAc, while stronger for fumaric acid in Cineole, CPME, and 2-MeTHF. Both solvent groups differ mainly in the non-polar region and the HBD/HBA capacity. Notably, for the systems composed of CPME with both acids, greater ratios of HB/MF and HB/vdW contributions are observed. This might explain the low affinity identified from the activity coefficient analysis, yet the similar HBA capacity than other bio-based solvents, such as cineole and 2-MeTHF. The low MF contribution suggests a poor molecular interstitial accommodation, reducing the interaction capacity and thus limiting the solubility of both acids in CPME. The energetic analysis revealed the complex interaction network that can be formed between the target organic acids and the bio-based solvents, which is critical to consider for the appropriate selection of the extraction media. The computed excess energies' results details can be found in SI.

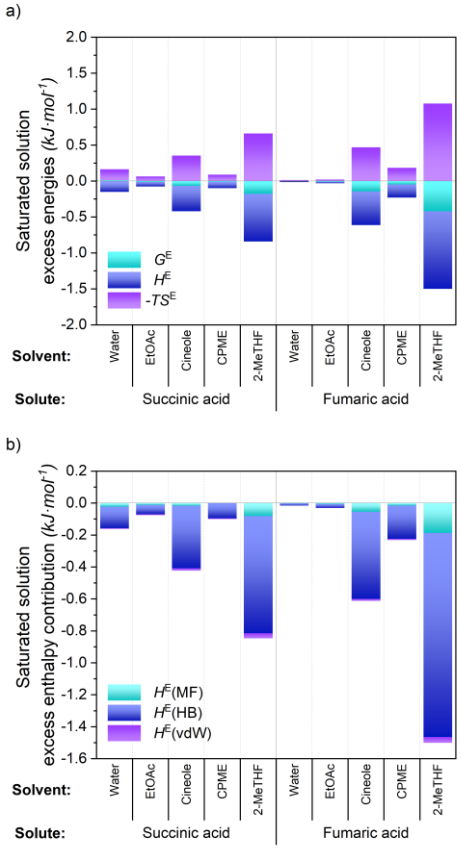


Figure 5. Analysis of the excess mixing energies of saturated organic acid – bio-based solvent systems computed by the COSMO-RS method at 293.15 K. a) Excess Gibbs free energy (G^E / kJ·mol^{-1}), Excess enthalpy (H^E / kJ·mol^{-1}), and Excess entropy ($-TS^E$ / kJ·mol^{-1}); b) Excess enthalpy contributions (kJ·mol^{-1}): electrostatic energy (MF), hydrogen bonding (HB), and van der Waals forces (vdW).

3.5 Energies of solution of organic acid – solvent systems

Given the solubility data and the van't Hoff equation (Eq. 3) is possible to estimate the energies of solution from the variation of the equilibrium constants with the temperature following the modification proposed by Krug et al. [52,53]. The method normalizes the temperature by accounting for the harmonic temperature, i.e., the harmonic mean of the N experimental temperature data points, according to Eq. 11. This results in a solubility function in the form of $\ln(x_1) = f(1/T\text{-}1/T_{hm})$, as shown in Figure 6, where the Gibbs free energy of solution, ΔG_{soln} / kJ·mol^{-1}, corresponds to the function intercept times the gas constant, R / kJ·mol^{-1}·K^{-1}, and the

harmonic temperature, as shown in Eq. 12; the enthalpy of solution, ΔH_{soln} / kJ·mol^{-1}, is related to the function slope times the gas constant, as per Eq. 13; and the entropy of solution, ΔS_{soln} / kJ·mol^{-1}·K^{-1}, is derived from the Gibbs free excess energy equation (Eq. 14). Uncertainties of the energies of solution are derived from the respective intercept and slope confident bounds of the van't Hoff equation linear regression. The full derivation of the linear regression parameters, as well as the energies of solution uncertainties, can be found in SI.

$$T_{hm} = \frac{N}{\sum {}^{1}/_{T}} \tag{11}$$

$$\Delta G_{soln} = -R \cdot T_{hm} \cdot \ln(x_1)_{T=T_{hm}} \tag{12}$$

$$\Delta H_{soln} = -R \frac{dln(x_1)}{d\left(\frac{1}{T} - \frac{1}{T_{hm}}\right)} \tag{13}$$

$$\Delta S_{soln} = \frac{\Delta H_{soln} - \Delta G_{soln}}{T_{hm}} \tag{14}$$

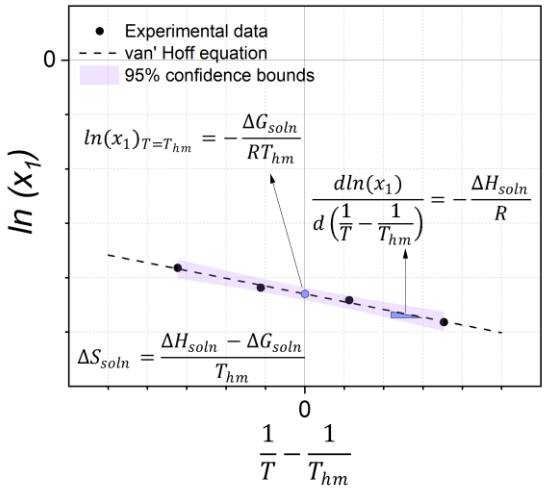


Figure 6. Schematic representation of energies of solution estimation from the solubility experimental data in the van't Hoff plot modified by Krug et al. [52,53]. The free energy of solution, ΔG_{soln} / kJ·mol^{-1}, is proportional to the function intercept, i.e., the function evaluated at $T = T_{hm}$, the enthalpy of solution, ΔH_{soln} / kJ·mol^{-1}, to the function slope. The entropy of solution is obtained from the Gibbs free energy relationship: ΔG=ΔH-$T\Delta S$.

The energies of solution of the organic acids in water and the bio-based solvents studied in this work were estimated at a fitted temperature of T_{hm} = 299 K, i.e. the harmonic mean of the temperature of all data reported, to allow for a proper comparison between the systems. The estimated energies of solution for succinic acid and fumaric acid are presented in Table 3. As expected, the dissolution of the organic acids is an endothermic process ($\Delta H_{soln} > 0$) in all solvents. Gibbs free energy results follow a similar trend to that of the overall organic acid solubilities: the higher the solubility values, the lower the Gibbs free energy. Positive values of the entropy of solution in all cases, except for fumaric acid in 2-MeTHF, suggest a possible contribution of this factor to the dissolution. Note that due to the few point measured for the systems [succinic acid + cineole] and [succinic acid + CPME], the uncertainty of the linear regression parameter causes a large estimated uncertainty of their respective energies of solution. Thus those data points are presented only as a qualitative insight into the energies trend.

To corroborate the predominant driving force in the organic acid dissolution process within the studied solvents, the relationship between the enthalpy and entropy of solution was assessed. A slope of the enthalpy respect to the entropy variation (Eq. 15) bigger than one denotes an enthalpy-driven dissolution process, and otherwise an entropy-driven process.[54] The results demonstrate an enthalpy-entropy compensation effect where a proportional linear trend is found between those factors for both organic acids, as depicted in Figure 7. The slopes of the enthalpy to the entropy are 1.01 and 1.30 for succinic acid and fumaric acid, respectively. This indicates an enthalpy-driven process for fumaric acid, i.e. its dissolution mechanism is based on the absorbed energy to break the solid molecular structure over the entropy change, which agrees with the observed low solubilities. On the other hand, no predominated effect can be observed for succinic acid, both enthalpy and entropy mechanisms are balanced.

$$Slope = \frac{d(\Delta H_{soln})}{d(T_{hm} \cdot \Delta S_{soln})} \quad (15)$$

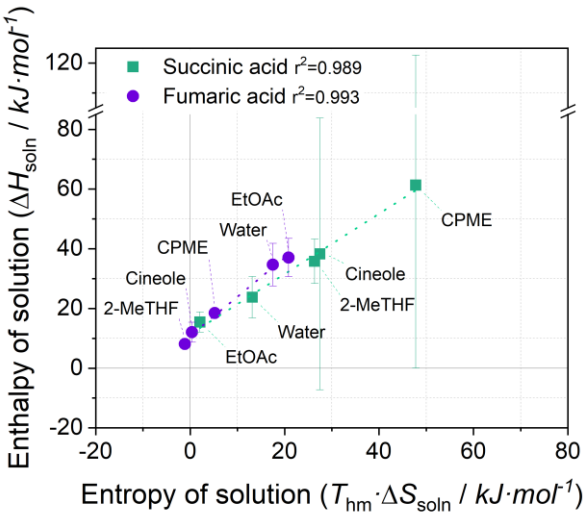


Figure 7. Enthalpy-entropy compensation effect for organic acids in water and bio-based solvents at a harmonic mean temperature of T_{hm} = 299 K for succinic acid and fumaric acid. Dotted lines represent the linear regression for each organic acid.

Table 3. Estimated Gibbs free energy (ΔG_{soln}), enthalpy (ΔH_{soln}), and entropy ($T_{hm} \cdot \Delta S_{soln}$) of solution, and their respective uncertainty with a confidence limit of 95%($U_{Comb,95\%}$), for succinic acid and fumaric acid in water and the bio-based solvents studied in this work. Estimations were performed following the methodology detailed in section 3.5 at T_{hm}=299 K.

Solvent	ΔG_{soln} kJ·mol^{-1}		ΔH_{soln} kJ·mol^{-1}		$T_{hm} \cdot \Delta S_{soln}$ kJ·mol^{-1}	
			Succinic acid			
Water	10.62	±0.20	23.77	±6.97	13.15	±7.17
EtOAc	13.41	±0.09	15.46	±3.41	2.05	±3.49
Cineole	10.81	±1.83	38.70	±45.51	27.89	±47.34
CPME	13.59	±1.14	61.35	±61.30	47.76	±62.44
2-MeTHF	9.48	±0.13	35.81	±7.46	26.33	±7.59
			Fumaric acid			
Water	17.14	±0.49	34.65	±7.13	17.51	±0.52
EtOAc	16.23	±0.37	37.06	±6.42	20.83	±6.06
Cineole	11.69	±0.17	12.04	±3.35	0.36	±3.18
CPME	13.25	±0.11	18.45	±1.05	5.20	±0.94
2-MeTHF	9.18	±0.06	8.08	±0.54	-1.10	±0.48

4. Conclusions

The saturated succinic acid and fumaric acid mole fraction solubility in water and bio-based solvents were measured in the temperatures range of [283 – 313] K, with an experimental uncertainty of measurement less than $9.3 \cdot 10^{-4}$ mol·mol^{-1}. Succinic acid shows to be the most soluble in 2-MeTHF, while fumaric acid also presents high solubility in 2-MeTHF, followed by cineole and CPME, rather than EtOAc and water. Results were successfully correlated with three solubility models, i.e., the van't Hoff equation, and the Buchowski–Ksiazaczak λh model, with relative deviations of less than 10%. The two-parameter van't Hoff equation gives the proper description of the reported data, allowing for further energetic interpretation of the dissolution phenomenon. The saturated liquid mixture interactions were analysed by computing the solid-liquid equilibria using the COSMO-RS method. Results reveal a complex unlike molecule interactions, where hydrogen bonding dominates the organic acid dissolution process. The Gibbs free energy, enthalpy, and entropy of solution were estimated for all systems at a mean harmonic temperature of the experimental data. While the dissolving process of the organic acids is endothermic, the thermodynamic analysis shows an enthalpy-entropy compensation effect for both solutes within the bio-based solvents studied in this work, providing insights into the predominant driving force for organic acid dissolution. This work provides new experimental data and insights into the bio-organic acid solubility in green hydrophobic solvents. Overall, these findings will promote the design of more benign biorefinery downstream processes for organic acids recovery.

AUTHOR INFORMATION

Corresponding Author

*E-mail: maria.gonzalezmiquel@upm.es

Funding Sources

This work was funded by the CONICYT PFCHA/DOCTORADO BECAS CHILE/2017 – 72180306. P. Gorgojo acknowledges the Spanish Ministry of Economy and Competitiveness and the European Social Fund through the Ramon y Cajal programme (RYC2019-027060-I/AEI/10.13039/501100011033). M. González-Miquel also acknowledges Comunidad Autónoma de Madrid (Spain) for funding through "Programa de Excelencia para el Profesorado Universitario".

Notes

The authors declare no competing financial interest.

Notation

ΔG_{fus}	Gibbs free energy of fusion	$kJ \cdot mol^{-1}$
ΔG_{soln}	Gibbs free energy of solution	$kJ \cdot mol^{-1}$
ΔH_m	Enthalpy of fusion	$kJ \cdot mol^{-1}$
ΔH_{soln}	Enthalpy of solution	$kJ \cdot mol^{-1}$
ΔS_{soln}	Entropy of solution	$kJ \cdot mol^{-1} \cdot K^{-1}$
A	Entropy factor (p_1) in van't Hoff equation	-
$AARD$	Average absolute relative deviation	
B	Enthalpy factor (p_2) in van't Hoff equation	-
G	Gibbs free energy	$kJ \cdot mol^{-1}$
H	Enthalpy	$kJ \cdot mol^{-1}$
HB	Hydrogen-bonding energy contribution	$kJ \cdot mol^{-1}$
HBA	Hydrogen bond acceptor	
HBD	Hydrogen bond donor	
h	Enthalpic factor (p_2) in λh model	$kJ \cdot mol^{-1}$
k_L	Coverage factor	
LC	Level of confidence	
LLX	Liquid-liquid extraction	
$M_{solution}$	Saturated supernatant plus the vial mass	g
M_{vial}	Vial mass	g
MF	Electrostatic energy contribution (misfit)	$kJ \cdot mol^{-1}$
MW	Molar weight	$g \cdot mol^{-1}$
M_{wet}	Dry organic acid plus the vial mass	g
N	Number of data points	
p	Fitted parameter i of solubility correlation	
p(σ)	σ-profile	
S	Entropy	$kJ \cdot mol^{-1} \cdot K^{-1}$
T	Temperature	K
T_b	Compound boiling point	K
T_{hm}	Harmonic mean temperature	K
T_m	Compound melting point	K
u	Standard uncertainty	
$U_{Comb,95\%}$	Combined expanded uncertainty ($LC_{95\%}$)	
vdW	van der Waals forces energy contribution	$kJ \cdot mol^{-1}$
wt_w	Solvent water content	$g \cdot g^{-1}$
x	Saturated solute mole fraction	$mol \cdot mol^{-1}$

Greek letters:

γ	Activity coefficient	$mol \cdot mol^{-1}$
δ	Root mean squared deviation	
ε	Relative deviation	%
λ	Non-ideality (p_1) in λh model	
μ	Chemical potential	$kJ \cdot mol^{-1}$
σ	Molecular charge density	$e \cdot nm^{-2}$

Subscripts:

1	Organic acid
2	Bio-based green solvent
i	Individual data point *i*

Superscripts:

α	Phase
cal	Calculated from solubility correlation
E	Excess property
exp	Experimental value
lit	Literature average value

References

[1] V.R. Veleva, B.W. Cue, The role of drivers, barriers, and opportunities of green chemistry adoption in the major world markets, Curr. Opin. Green Sustain. Chem. 19 (2019) 30–36. https://doi.org/10.1016/j.cogsc.2019.05.001.

[2] J.J. Bozell, G.R. Petersen, Technology development for the production of biobased products from biorefinery carbohydrates - The US Department of Energy's "top 10" revisited, Green Chem. 12 (2010) 539–554. https://doi.org/10.1039/b922014c.

[3] L. Nattrass, C. Biggs, A. Bauen, C. Parisi, E. Rodríguez-Cerezo, M. Gómez-Barbero, The EU bio-based industry: Results from a survey, 2016. https://doi.org/10.2791/806858.

[4] T.A. Werpy, J.E. Holladay, J.F. White, Top Value Added Chemicals From Biomass: I. Results of Screening for Potential Candidates from Sugars and Synthesis Gas, Richland, WA, 2004. https://doi.org/10.2172/926125.

[5] M. Jiang, J. Ma, M. Wu, R. Liu, L. Liang, F. Xin, W. Zhang, H. Jia, W. Dong, Progress of succinic acid production from renewable resources: Metabolic and fermentative strategies, Bioresour. Technol. 245 (2017) 1710–1717. https://doi.org/10.1016/j.biortech.2017.05.209.

[6] Q. Xu, S. Li, H. Huang, J. Wen, Key technologies for the industrial production of fumaric acid by fermentation, Biotechnol. Adv. 30 (2012) 1685–1696. https://doi.org/10.1016/j.biotechadv.2012.08.007.

[7] J.M. Pinazo, M.E. Domine, V. Parvulescu, F. Petru, Sustainability metrics for succinic acid production: A comparison between biomass-based and petrochemical routes, Catal. Today. 239 (2015) 17–24. https://doi.org/10.1016/j.cattod.2014.05.035.

[8] T. Brouwer, M. Blahusiak, K. Babic, B. Schuur, Reactive extraction and recovery of levulinic acid, formic acid and furfural from aqueous solutions containing sulphuric acid, Sep. Purif. Technol. 185 (2017) 186–195. https://doi.org/10.1016/j.seppur.2017.05.036.

[9] H. Song, S.Y. Lee, Production of succinic acid by bacterial fermentation, Enzyme Microb. Technol. 39 (2006) 352–361. https://doi.org/10.1016/j.enzmictec.2005.11.043.

[10] M. Gonzalez-Miquel, J. Esteban, Novel Solvents for Biotechnology Applications, in: M. Moo-Young (Ed.), Compr. Biotechnol., 3rd ed., Elsevier: Pergamon, 2019: pp. 790–806. https://doi.org/10.1016/B978-0-444-64046-8.00459-6.

[11] P. López-Porfiri, P. Gorgojo, M. Gonzalez-Miquel, Green Solvent Selection Guide for Biobased Organic Acid Recovery, ACS Sustain. Chem. Eng. 8 (2020) 8958–8969. https://doi.org/10.1021/acssuschemeng.0c01456.

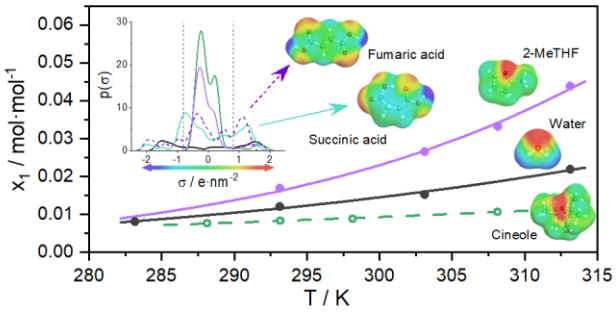
[12] R.J. Davey, J.W. Mullin, M.J.L. Whiting, Habit modification of succinic acid crystals grown from different solvents, J. Cryst. Growth. 58 (1982) 304–312. https://doi.org/10.1016/0022-0248(82)90277-9.

[13] X. Jiang, Y. Hu, Z. Meng, W. Yang, F. Shen, Solubility of succinic acid in different aqueous solvent mixtures: Experimental measurement and thermodynamic modeling, Fluid Phase Equilib. 341 (2013)

7–11. https://doi.org/10.1016/j.fluid.2012.12.018.
[14] X. Sheng, W. Luo, Q. Wang, Determination and Correlation for the Solubilities of Succinic Acid in Cyclohexanol + Cyclohexanone + Cyclohexane Solvent Mixtures, J. Chem. Eng. Data. 63 (2018) 801–811. https://doi.org/10.1021/acs.jced.7b00956.
[15] W.D. Bancroft, F.J.C. Butler, Solubility of Succinic Acid in Binary Mixtures, J. Phys. Chem. 36 (1932) 2515–2520. https://doi.org/10.1021/j150339a013.
[16] L. Dang, W. Du, S. Black, H. Wei, Solubility of Fumaric Acid in Propan-2-ol, Ethanol, Acetone, Propan-1-ol, and Water, J. Chem. Eng. Data. 54 (2009) 3112–3113. https://doi.org/10.1021/je9001637.
[17] F. Eckert, A. Klamt, Fast Solvent Screening via Quantum Chemistry: COSMO-RS Approach, AIChE J. 48 (2002) 369–385. https://doi.org/10.1002/aic.690480220.
[18] B. Schröder, L.M.N.B.F. Santos, I.M. Marrucho, J.A.P. Coutinho, Prediction of aqueous solubilities of solid carboxylic acids with COSMO-RS, Fluid Phase Equilib. 289 (2010) 140–147. https://doi.org/10.1016/j.fluid.2009.11.018.
[19] J. Esteban, A.J. Vorholt, W. Leitner, An overview of the biphasic dehydration of sugars to 5-hydroxymethylfurfural and furfural: a rational selection of solvents using COSMO-RS and selection guides, Green Chem. 22 (2020). https://doi.org/10.1039/c9gc04208c.
[20] M. Gonzalez-Miquel, M. Massel, A. Desilva, J. Palomar, F. Rodriguez, J.F. Brennecke, Excess enthalpy of monoethanolamine + ionic liquid mixtures: How good are COSMO-RS predictions?, J. Phys. Chem. B. 118 (2014) 11512–11522. https://doi.org/10.1021/jp507547q.
[21] K. Paduszyński, An overview of the performance of the COSMO-RS approach in predicting the activity coefficients of molecular solutes in ionic liquids and derived properties at infinite dilution, Phys. Chem. Chem. Phys. 19 (2017) 11835–11850. https://doi.org/10.1039/C7CP00226B.
[22] P. Matheswaran, C.D. Wilfred, K.A. Kurnia, A. Ramli, Overview of Activity Coefficient of Thiophene at Infinite Dilution in Ionic Liquids and their Modeling Using COSMO-RS, Ind. Eng. Chem. Res. 55 (2016) 788–797. https://doi.org/10.1021/acs.iecr.5b04152.
[23] D. Constantinescu, J. Rarey, J. Gmehling, Application of COSMO-RS Type Models to the Prediction of Excess Enthalpies, Ind. Eng. Chem. Res. 48 (2009) 8710–8725. https://doi.org/10.1021/ie900315p.
[24] A. Klamt, M. Diedenhofen, Blind prediction test of free energies of hydration with COSMO-RS, J. Comput. Aided. Mol. Des. 24 (2010) 357–360. https://doi.org/10.1007/s10822-010-9354-4.
[25] A. Zirahi, H. Sadeghi Yamchi, A. Haddadnia, M. Zirrahi, H. Hassanzadeh, J. Abedi, Ethyl acetate as a bio-based solvent to reduce energy intensity and CO 2 emissions of in situ bitumen recovery, AIChE J. 66 (2020). https://doi.org/10.1002/aic.16828.
[26] J.M. Prausnitz, R.N. Lichtenthaler, E.G. de Azevedo, Molecular Thermodynamics of Fluid-Phase Equilibria, 3rd ed., Prentice- Hall: Englewood Cliffs, NJ, 1999.
[27] H. Buchowski, A. Ksiazczak, S. Pietrzyk, Solvent activity along a saturation line and solubility of hydrogen-bonding solids, J. Phys. Chem. 84 (1980) 975–979. https://doi.org/10.1021/j100446a008.
[28] P.J. Linstrom, W.G. Mallard, eds., NIST Chemistry WebBook, NIST Standard Reference Database Number 69, National Institute of Standards and Technology, Gaithersburg MD, 20899, n.d. https://doi.org/10.18434/T4D303.
[29] B.N. Taylor, C.E. Kuyatt, Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, Gaithersburg, MD, 1994. http://physics.nist.gov/TN1297.
[30] A. Klamt, F. Eckert, COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids, Fluid Phase Equilib. 172 (2000) 43–72. https://doi.org/10.1016/S0378-3812(00)00357-5.
[31] A. Klamt, The COSMO and COSMO-RS solvation models, WIREs Comput. Mol. Sci. 1 (2011) 699–709. https://doi.org/10.1002/wcms.56.
[32] A. Apelblat, E. Manzurola, 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 (1987) 317–320. https://doi.org/10.1016/0021-9614(87)90139-X.
[33] J.W. Mullin, Crystallisation, Butterworths, London, 1972.
[34] G. Massol, F. Lamouroux, Sur la solubilite dans l'eau des acides maloniques substitues, Comptes Rendus Hebd. Des Seances l'Academie Des Sci. 128, (1899) 1000–1002.
[35] J.W. Mullin, M.J.L. Whiting, Succinic Acid Crystal Growth Rates in Aqueous Solution, Ind. Eng.

Chem. Fundam. 19 (1980) 117–121. https://doi.org/10.1021/i160073a020.
[36] J.H.C. Merckel, Die loslichkeit der dicarbonsauren, Recl. Des Trav. Chim. Des Pays-Bas. 56 (1937) 811–814.
[37] R. Wright, CLXXXVII.—Selective solvent action. Part VI. The effect of temperature on the solubilities of semisolutes in aqueous alcohol, J. Chem. Soc. (1927) 1334–1337. https://doi.org/10.1039/JR9270001334.
[38] S.H. Yalkowsky, Y. He, P. Jain, Handbook of Aqueous Solubility Data, 2nd ed., Boca Raton, Fla: CRC., 2010.
[39] R.M. Dawson, Data for Biochemical Research, Clarendon Press, Oxford, 1959.
[40] S.S. Doosaj, W.V. Bhagwat, Solubilities of weak acids in salts of weak acids at very high concentrations, J. Indian Chem. Soc. 10 (1933) 225–232.
[41] G.S. Forbes, A.S. Coolidge, Relations between distribution ratio, temperature and concentration in system: water, ether, succinic acid, J. Am. Chem. Soc. 41 (1919) 150–167. https://doi.org/10.1021/ja01459a004.
[42] R.K. Freier, Aqueous Solutions Volume 1: Data for Inorganic and Organic Compounds, Walter de Gruyter, New York, 1976.
[43] A.-P. Hyvärinen, H. Lihavainen, A. Gaman, L. Vairila, H. Ojala, M. Kulmala, Y. Viisanen, Surface Tensions and Densities of Oxalic, Malonic, Succinic, Maleic, Malic, and cis -Pinonic Acids, J. Chem. Eng. Data. 51 (2006) 255–260. https://doi.org/10.1021/je050366x.
[44] F. Lamouroux, Sur la solubilite dans l'eau des acides normaux de la serie oxalique, Comptes Rendus Hebd. Des Seances l'Academie Des Sci. 128 (1899) 998–1000.
[45] K. Linderstrom-Lang, Solubility of hydroquinone, Comptes Rendus Des Trav. Du Lab. Carlsbreg. 15, (1924) 4–28.
[46] H. Marshall, D. Bain, Sodium succinates, J. Chem. Soc. 97 (1910) 1074– 1085,.
[47] R.M.C. Dawson, D.C. Elliott, W.H. Elliott, K.M. Jones, Data for Biochemical Research, Oxford University Press, Pergamon, 1969.
[48] J.M. Weiss, C.R. Downs, The physical properties of maleic, fumaric and malic acids, J. Am. Chem. Soc. 45 (1923) 1003–1008. https://doi.org/10.1021/ja01657a018.
[49] Z. Wagner, M. Bendová, J. Rotrekl, A. Sýkorová, M. Čanji, N. Parmar, Density and sound velocity measurement by an Anton Paar DSA 5000 density meter: Precision and long-time stability, J. Mol. Liq. 329 (2021). https://doi.org/10.1016/j.molliq.2021.115547.
[50] K. Liu, J.D. Cruzan, R.J. Saykally, Water clusters, Science (80-.). 271 (1996) 929–933. https://doi.org/10.1126/science.271.5251.929.
[51] T. Lazaridis, Hydrophobic Effect, ELS. John Wiley Sons, Ltd Chichester. (2013). https://doi.org/10.1002/9780470015902.a0002974.pub2.
[52] R.R. Krug, W.G. Hunter, R.A. Grieger, Enthalpy-entropy compensation. 1. Some fundamental statistical problems associated with the analysis of van't Hoff and Arrhenius data, J. Phys. Chem. 80 (1976) 2335–2341. https://doi.org/10.1021/j100562a006.
[53] R.R. Krug, W.G. Hunter, R.A. Grieger, Enthalpy-entropy compensation. 2. Separation of the chemical from the statistical effect, J. Phys. Chem. 80 (1976) 2341–2351. https://doi.org/10.1021/j100562a007.
[54] A. Aydi, C. Ayadi, K. Ghachem, A. Al-Khazaal, D. Delgado, M. Alnaief, L. Kolsi, Solubility, Solution Thermodynamics, and Preferential Solvation of Amygdalin in Ethanol + Water Solvent Mixtures, Pharmaceuticals. 13 (2020) 395. https://doi.org/10.3390/ph13110395.

TOC GRAPHIC



SYNOPSIS

Solubilities of succinic and fumaric acids in water and bio-based solvents were experimentally measured and correlated, and thermodynamically assessed.

Solubility Study and Thermodynamic Modelling of Succinic Acid and Fumaric Acid in Bio-based Solvents

Pablo López-Porfiri[1], Patricia Gorgojo[1,2,3], and María Gonzalez-Miquel [1,4]*

[1] Department of Chemical Engineering, Faculty of Science and Engineering, The University of Manchester, Manchester M13 9PL, UK.

[2] Nanoscience and Materials Institute of Aragón (INMA) CSIC-Universidad de Zaragoza, C/ Mariano Esquillor s/n, 50018 Zaragoza, Spain.

[3] Chemical and Environmental Engineering Department, Universidad de Zaragoza, C/ Pedro Cerbuna 12, 50009 Zaragoza, Spain.

[4] Departamento de Ingeniería Química Industrial y del Medioambiente, ETS Ingenieros Industriales, Universidad Politécnica de Madrid, C/ José Gutiérrez Abascal 2, Madrid, Spain.

[*] Corresponding author. E-mail: maria.gonzalezmiquel@upm.es

Table of content

Notation

a	Resolution of the instrument	
G	Gibbs free energy	kJ·mol^{-1}
H	Enthalpy	kJ·mol^{-1}
HB	Hydrogen-bonding energy contribution	kJ·mol^{-1}
k_L	Coverage factor	-
LC	Level of confidence	-
$M_{solution}$	Saturated supernatant plus the vial mass	g
M_{vial}	Vial mass	g
MW	Molar weight	g·mol^{-1}
MF	Electrostatic energy contribution (misfit)	kJ·mol^{-1}
M_{wet}	Dry organic acid plus the vial mass	g
N	Number of data points	-
p	Fitted parameter i of solubility correlation	-
S	Entropy	kJ·mol^{-1}·K^{-1}
s	Parameter standard deviations of the least-squared linear regression method	
$s_{Y/X}$	Random errors in the *y*-direction of the least-squared linear regression method	
T	Temperature	K
$t_{(N-2)}$	*t*- factor	
u	Standard uncertainty	-
$U_{Comb,95\%}$	Combined expanded uncertainty ($LC_{95\%}$)	-
v	Degree of freedom	
vdW	van der Waals forces energy contribution	kJ·mol^{-1}
wt_w	Solvent water content	g·g^{-1}
X	Independent variable of the least-squared linear regression method	
x	Saturated solute mole fraction	mol·mol^{-1}
Y	Dependent variable of the least-squared linear regression method	

Greek letters:

γ	Activity coefficient	mol·mol^{-1}
ε	Relative deviation	%

Subscripts:

1	Organic acid
2	Bio-based green solvent
eff	Effective
i	Individual data point *i*
j	Variable *j*

Superscripts:

cal	Calculated from solubility correlation
E	Excess property
exp	Experimental value
lit	Literature average value
∞	Infinite dilution

1. Methodology validation

Solubility data found in the literature for succinic acid [1–16] and fumaric acid [3,8,12,17,18] in water was used to validate the proposed methodology. Each solubility data set were fed into the van't Hoff equation (Eq. S1) and compared with the respective correlated data from experimental values by the root mean squared deviation, $\delta[ln(x_1)]$, according to Eq. S2. The validation results are shown in Table S1.

$$ln(x_1) = p_1 + \frac{p_2}{T} \quad \text{(S1)}$$

$$\delta[ln(x_1)] = \sqrt{\frac{1}{N}\sum\left(ln\left(x_1^{exp}\right) - ln\left(x_1^{lit}\right)\right)^2} \quad \text{(S2)}$$

Table S1. Comparison of experimental and mean values found in the literature for aqueous solubility of succinic and fumaric acids for methodology validation. The estimated combined expanded uncertainty for the mean literature values: $U_{comb,95\%}\left(ln\left(x_1^{lit}\right)\right)$ = 0.10 and 0.14, for succinic and fumaric acids, respectively.

	Succinic acid		**Fumaric acid**	
$\mathbf{1/T}$ 1/K	$\mathbf{ln\left(x_1^{exp}\right)}$ mol·mol^{-1}	$\mathbf{ln\left(x_1^{lit}\right)}$ mol·mol^{-1}	$\mathbf{ln\left(x_1^{exp}\right)}$ mol·mol^{-1}	$\mathbf{ln\left(x_1^{lit}\right)}$ mol·mol^{-1}
0.0035	-4.82	-5.08	-7.86	-7.48
0.0034	-4.42	-4.66	-7.19	-7.09
0.0033	-4.19	-4.27	-6.63	-6.72
0.0032	-3.82	-3.90	-6.30	-6.38
	$\boldsymbol{\delta[ln(x_1)]}$:	0.19	$\boldsymbol{\delta[ln(x_1)]}$:	0.20

2. Experimental uncertainty

Reported solubility values, $\hat{x}_1$, are presented as the measurement result, x_1, along with the expanded combined uncertainty, $U_{\text{Comb}}(x_1)$, as is shown in Eq. S3. The saturated acid mole fraction was calculated following Eq. 4, where M_{vial}, M_{dry}, and M_{solution} correspond to the vial mass, the dry organic acid plus the vial mass, and the saturated supernatant plus the vial mass, respectively. Two approaches were adopted to evaluate $U_{\text{Comb}}(x_1)$, according to the NIST guidelines for expressing results uncertainty [19]. First, the type A evaluation by statistical methods of the variability of the results, where the error arising from a random effect on the repeatability of the measurement is obtained assuming a *t*-distribution and the degree of freedom, *v*, according to Eq. S5. The results standard uncertainty, $u(x_1)$, is obtained from the measurement deviation, while the *t*-factor, $t_{(N-2)}$, for a level of confidence (*LC*) of 95% of the measured data in triplicate, i.e., degree freedom v = 1, results in 12.71. Due to the complexity of the procedure methodology, no more replicate experiments by data point were possible to measure. This results in large confidence bound given by the basic statistical methods, which no genuinely reflect the experiment precisions, and the uncertainty should be estimated from other scientific judgment.

$$\hat{x}_1 = x_1 \pm U_{comb}(x_1) \tag{S3}$$

$$x_1 = \frac{M_{dry} - M_{vial}}{MW_1} \cdot \left[\frac{M_{dry} - M_{vial}}{MW_1} + \frac{M_{solution} - M_{dry}}{MW_2}\right]^{-1} \tag{S4}$$

$$U_{comb}(x_1) = u(x_1) \cdot t_{(N-2)} \tag{S5}$$

The type B approach corresponds to the uncertainty of measurement, which comes from the error of the measured properties variation, due to the instruments' resolution and error propagation in the data treatment. Since the proposed methodology is based on reliable independent measurements under highly controlled conditions, a type B approach represents a better approximation to the experimental uncertainty. Assuming all the variable errors as small and independent, the combined standard uncertainty, $u_{comb}(x_1)$, is determined by the law of propagation of uncertainty: Eq. S6. The list of the variable uncertainty estimation derived from the organic acid saturated mole fraction measurements (Eq. S4) is given in Table S2. Then, $U_{\text{Comb}}(x_1)$ is calculated by including a coverage factor, k_{L}, (Eq. S7) given by the *t*-distribution for a specific confidence level, *p*, and effective degrees of freedom, v_{eff}, as shown in Eq. S8. The effective degrees of freedom are calculated by the Welch-Satterthwaite formula (Eq. S9), where v_{j} =r-1, is the independent variable *j* degree of freedom and *r* is the size of the sample. The mass weight

uncertainty was modelled by a normal distribution of the scale resolution (±a=0.0001 g) according to the instruments manufacturer's specifications and assuming a 50 per cent probability that the value of the quantity lies in the interval a_- to a_+, as shown in Eq. S10. Thus, for the measurement procedure of this work and r=3 replicates for the M_{vial}, M_{dry}, and $M_{solution}$ measurements give a ν_{eff} = 4. Lastly, taking LC = 95% for $t(4)$ results in a k_L of 2.78.

$$u_{comb}^2(x_1) = \sum_{j=1}^{n} \left(\frac{\partial x_1}{\partial M_j}\right)^2 \cdot u^2(M_j) \tag{S6}$$

$$U_{comb}(x_1) = k_L \cdot u_{comb}(x_1) \tag{S7}$$

$$k_L = t_p(\nu_{eff}) \tag{S8}$$

$$\nu_{eff} = \frac{u_{comb}^4(x_1)}{\sum_{j=1}^{n} \frac{\left(\frac{\partial x_1}{\partial M_j}\right)^4 \cdot u^4(M_j)}{\nu_j}} \tag{S9}$$

$$u(M_j) \approx \frac{a_- - a_+}{2} \cdot 1.48 \tag{S10}$$

Table S2. Estimation of the combined expanded uncertainty for the saturated organic acid mole fraction in bio-based solvents, $U_{Comb,95\%}(x_1)$, based on the contributions of properties and variables measured and calculated.

Property, Variable	**Type of uncertainty** Coverage factor, k_L (CL / %)	**Uncertainty**
Water content in the bio-based solvent ($10^{-4} \leq wt_w \leq 10^{-2}$) g·g^{-1}	Standard	$u(wt_w) = 10^{-7}$ g·g^{-1}
Temperature ($283 \leq T \leq 313$) K	Standard	$u(T) = 1$ K
Mass ($0.0000 \leq M \leq 10.0000$) g	Standard	$u(M_j) = 1.48 \cdot 10^{-4}$ g
Saturated mole fraction of organic acid ($0 \leq x_1 \leq 1$) mol·mol^{-1}	Combined $k_L = 2.78$ (95%)	$\frac{U_{comb}(x_1)}{k_L} = \sqrt{\left[\left\lvert\frac{\partial x_1}{\partial M_{vial}}\right\rvert^2 + \left\lvert\frac{\partial x_1}{\partial M_{solution}}\right\rvert^2 + \left\lvert\frac{\partial x_1}{\partial M_{dry}}\right\rvert^2\right] \cdot u(M)^2}$

$$\frac{\partial x_1}{\partial M_{vial}} = \frac{M_{dry} - M_{vial}}{(MW_1)^2 \cdot \left[\frac{M_{dry} - M_{vial}}{MW_1} - \frac{M_{solution} - M_{dry}}{MW_2}\right]^2} - \frac{1}{\left[M_{dry} - M_{vial} - \frac{MW_1}{MW_2}(M_{solution} - M_{dry})\right]}$$

$$\frac{\partial x_1}{\partial M_{solution}} = \frac{M_{dry} - M_{vial}}{MW_1 \cdot MW_2 \cdot \left[\frac{M_{dry} - M_{vial}}{MW_1} - \frac{M_{solution} - M_{dry}}{MW_2}\right]^2}$$

$$\frac{\partial x_1}{\partial M_{dry}} = \frac{1}{\left[M_{dry} - M_{vial} - \frac{MW_1}{MW_2}(M_{solution} - M_{dry})\right]} - \frac{(M_{dry} - M_{vial}) \cdot \left(1/MW_1 + 1/MW_2\right)}{\left[M_{dry} - M_{vial} - \frac{MW_1}{MW_2}(M_{solution} - M_{dry})\right]^2}$$

3. Experimental solubility data

Table S3 summarizes the experimental data obtained from the proposed methodology along with the respective estimated measurement (type B) and statistical (type A) combined expanded uncertainty, $U_{Comb,95\%}(x_1)$ / mol·mol^{-1}.

Table S3. Experimental saturated organic acid mole fraction (x_1 / mol·mol^{-1}) in water and bio-based solvents and its combined expanded uncertainties.

Succinic acid				Fumaric acid			
T^*	x_1	$U_{Comb,95\%}(x_1)$ /mol·mol^{-1}		T^*	x_1	$U_{Comb,95\%}(x_1)$ /mol·mol^{-1}	
K	mol·mol^{-1}	Type B**	Type A***	K	mol·mol^{-1}	Type B**	Type A***
			Water				
283	0.0081	$8.878\cdot10^{-5}$	$4.056\cdot10^{-4}$	**283**	0.0005	$9.040\cdot10^{-5}$	$4.130\cdot10^{-4}$
293	0.0121	$9.673\cdot10^{-5}$	$4.419\cdot10^{-4}$	**293**	0.0008	$9.129\cdot10^{-5}$	$4.171\cdot10^{-4}$
303	0.0152	$9.849\cdot10^{-5}$	$4.499\cdot10^{-4}$	**303**	0.0013	$9.253\cdot10^{-5}$	$4.227\cdot10^{-4}$
313	0.0219	$1.022\cdot10^{-4}$	$4.667\cdot10^{-4}$	**313**	0.0018	$9.121\cdot10^{-5}$	$4.167\cdot10^{-4}$
			EtOAc				
283	0.0033	$4.946\cdot10^{-4}$	$2.260\cdot10^{-3}$	**293**	0.0010	$5.018\cdot10^{-4}$	$2.292\cdot10^{-3}$
293	0.0038	$4.927\cdot10^{-4}$	$2.251\cdot10^{-3}$	**298**	0.0015	$4.974\cdot10^{-4}$	$2.272\cdot10^{-3}$
303	0.0050	$4.966\cdot10^{-4}$	$2.269\cdot10^{-3}$	**303**	0.0019	$5.018\cdot10^{-4}$	$2.292\cdot10^{-3}$
308	0.0056	$4.969\cdot10^{-4}$	$2.270\cdot10^{-3}$	**308**	0.0023	$4.956\cdot10^{-4}$	$2.264\cdot10^{-3}$
313	0.0060	$4.814\cdot10^{-4}$	$2.199\cdot10^{-3}$	**313**	0.0028	$5.064\cdot10^{-4}$	$2.314\cdot10^{-3}$
			Cineole				
283	0.0054	$8.577\cdot10^{-4}$	$3.918\cdot10^{-3}$	**288**	0.0077	$8.907\cdot10^{-4}$	$4.069\cdot10^{-3}$
288	0.0074	$8.731\cdot10^{-4}$	$3.989\cdot10^{-3}$	**293**	0.0083	$9.317\cdot10^{-4}$	$4.256\cdot10^{-3}$
293	0.0094	$8.708\cdot10^{-4}$	$3.978\cdot10^{-3}$	**298**	0.0088	$8.908\cdot10^{-4}$	$4.069\cdot10^{-3}$
				308	0.0106	$9.157\cdot10^{-4}$	$4.183\cdot10^{-3}$
			CPME				
293	0.0025	$5.988\cdot10^{-4}$	$2.736\cdot10^{-3}$	**293**	0.0042	$6.105\cdot10^{-4}$	$2.789\cdot10^{-3}$
298	0.0042	$6.086\cdot10^{-4}$	$2.780\cdot10^{-3}$	**298**	0.0047	$6.130\cdot10^{-4}$	$2.800\cdot10^{-3}$
303	0.0058	$5.889\cdot10^{-4}$	$2.690\cdot10^{-3}$	**303**	0.0054	$6.097\cdot10^{-4}$	$2.785\cdot10^{-3}$
				308	0.0060	$7.568\cdot10^{-4}$	$3.457\cdot10^{-3}$
				313	0.0067	$6.268\cdot10^{-4}$	$2.863\cdot10^{-3}$
			2-MeTHF				
293	0.0169	$5.421\cdot10^{-4}$	$2.476\cdot10^{-3}$	**283**	0.0208	$5.374\cdot10^{-4}$	$2.455\cdot10^{-3}$
303	0.0265	$5.407\cdot10^{-4}$	$2.470\cdot10^{-3}$	**298**	0.0248	$5.354\cdot10^{-4}$	$2.446\cdot10^{-3}$
308	0.0331	$5.458\cdot10^{-4}$	$2.494\cdot10^{-3}$	**303**	0.0260	$5.463\cdot10^{-4}$	$2.496\cdot10^{-3}$
313	0.0439	$5.712\cdot10^{-4}$	$2.610\cdot10^{-3}$	**308**	0.0274	$5.527\cdot10^{-4}$	$2.525\cdot10^{-3}$
				313	0.0290	$5.443\cdot10^{-4}$	$2.487\cdot10^{-3}$

* Standard uncertainty for temperature, $u(T)$ = 1 K.

** Combined expanded uncertainty of measurement (type B) estimated with a level of confidence of 95% of saturated organic acid mole fraction in water and bio-based solvents.

*** Statistical combined expanded uncertainty (type A) estimated with a level of confidence of 95% of saturated organic acid mole fraction in water and bio-based solvents.

4. Correlation relative deviations

Individual relative deviations percentages (Eq. S11) of the empirical correlation used in this work are presented the Figure S1 below.

$$\varepsilon_i = 100 \cdot \left(x_{1,i}^{exp} - x_{1,i}^{cal}\right)/x_{1,i}^{exp} \quad \text{(S11)}$$

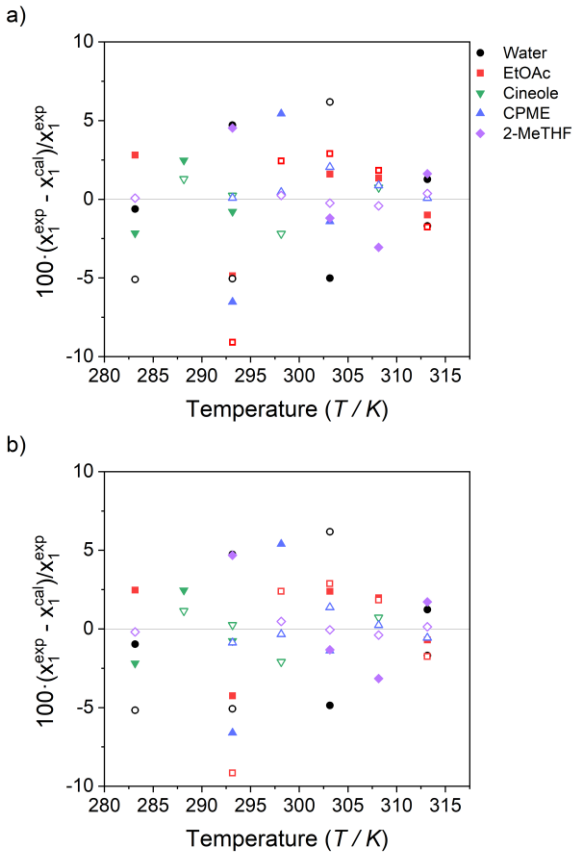


Figure S1. Percentage of the individual relative deviation from a) van't Hoff equation; b) Buchowski–Ksiazaczak λh model for experimental organic acid solubility in water and bio-based solvents. Filled symbols: succinic acid; empty symbols: fumaric acid.

5. Computed organic acid activity coefficients

Table S4 summarizes the organic acid activity coefficients at the saturated solution condition, γ_1 / mol·mol^{-1}, and at infinite dilution, γ_1^{∞} / mol·mol^{-1}, obtained from solid-liquid equilibria computed by the COSMO-RS method within the experimental temperature range.

Table S4. Organic acid activity coefficients at the saturated liquid condition obtained from SLE computed by the COSMO-RS method within the experimental temperature range.

Succinic acid			Fumaric acid		
T / K	γ_1 / mol·mol^{-1}	γ_1^{∞} / mol·mol^{-1}	T / K	γ_1 / mol·mol^{-1}	γ_1^{∞} / mol·mol^{-1}
			Water		
283	1.232	1.351	283	0.5770	0.5827
293	1.419	1.648	293	0.7676	0.7803
303	1.596	1.947	303	0.9814	1.009
313	1.672	2.236	313	1.214	1.263
			EtOAc		
283	0.1969	0.1925	293	0.0238	0.0236
293	0.2657	0.2598	298	0.0296	0.0292
303	0.3472	0.3387	303	0.0363	0.0358
308	0.3919	0.3821	308	0.0442	0.0434
313	0.4386	0.4277	313	0.0533	0.0522
			Cineole		
283	0.0114	0.0104	288	0.0005	0.0004
288	0.0272	0.0242	293	0.0009	0.0008
293	0.0413	0.0357	298	0.0015	0.0012
			308	0.0039	0.0032
			CPME		
293	0.2251	0.2180	293	0.0104	0.0096
298	0.3003	0.2857	298	0.0152	0.0140
303	0.3927	0.3685	303	0.0219	0.0201
			308	0.0308	0.0282
			313	0.0427	0.0388
			2-MeTHF		
293	0.0152	0.0117	283	0.0003	0.0002
303	0.0332	0.0228	298	0.0012	0.0008
308	0.0486	0.0312	303	0.0018	0.0012
313	0.0732	0.0419	308	0.0028	0.0018
			313	0.0041	0.0026

6. Saturated organic acid – bio-based solvent excess mixing energies

Table S5 summarizes the excess mixing energies of saturated organic acid – bio-based solvent systems computed by the COSMO-RS method at 293.15 K: excess Gibbs free energy (G^E / kJ·mol^{-1}), excess enthalpy (H^E / kJ·mol^{-1}), and excess entropy ($-TS^E$ / kJ·mol^{-1}); electrostatic energy excess enthalpy contribution (H^E(MF) / kJ·mol^{-1}), hydrogen bonding energy excess enthalpy contribution (H^E(HB) / kJ·mol^{-1}), and van der Waals forces energy excess enthalpy contribution (H^E(vdW) / kJ·mol^{-1}).

Table S5. Excess mixing energies of saturated organic acid – bio-based solvent systems and excess enthalpy contributions: electrostatic energy (MF), hydrogen bonding (HB), and van der Waals forces (vdW), computed by the COSMO-RS method at 293.15 K.

Organic acid	**Solvent**	H^E kJ·mol^{-1}	G^E kJ·mol^{-1}	$-TS^E$ kJ·mol^{-1}	H^E(MF) kJ·mol^{-1}	H^E(HB) kJ·mol^{-1}	H^E (vdW) kJ·mol^{-1}
Succinic acid	Water	-0.1514	0.0125	0.1638	-0.0223	-0.1352	-0.0046
	EtOAc	-0.0758	-0.0124	0.0634	-0.0111	-0.0586	-0.0057
	Cineole	-0.4208	-0.0661	0.3548	-0.0170	-0.3910	-0.0148
	CPME	-0.0990	-0.0092	0.0898	0.0005	-0.0953	-0.0051
	2-MetHF	-0.8401	-0.1783	0.6619	-0.0809	-0.7333	-0.0324
Fumaric acid	Water	-0.0156	-0.0005	0.0151	-0.0016	-0.0145	0.0002
	EtOAc	-0.0314	-0.0091	0.0222	-0.0054	-0.0244	-0.0013
	Cineole	-0.6125	-0.1439	0.4686	-0.0541	-0.5452	-0.0130
	CPME	-0.2307	-0.0471	0.1835	-0.0152	-0.2087	-0.0072
	2-MetHF	-1.4982	-0.4201	1.0781	-0.1866	-1.2772	-0.0365

7. Energies of solution uncertainties estimation

The van't Hoff equation (Eq. S1) can be expressed as a linear function $Y=f(X)$ as shown in Eq. S12. In this way, the parameters can be obtained by the least-squared linear regression method of the N data points (X, Y) [20]. The least-squared line intercept, p_1, as well as slope, p_2, are determined by the Eq. S13 and Eq. S14, respectively, where $\bar{X}$ is average of $1/T$ data point the and $\bar{Y}$ is the average of the $\ln(x_1)$ data points.

$$\mathrm{Y} = p_1 + p_2 \cdot X \tag{S12}$$

$$p_1 = \bar{Y} - p_2 \cdot \bar{X} \tag{S13}$$

$$p_2 = \frac{\sum_i \{(X_i - \bar{X})(Y_i - \bar{Y})\}}{\sum_i (X_i - \bar{X})^2} \tag{S14}$$

The standard deviations of the intercept, s_{p1}, and the slope, s_{p2}, are obtained from Eq. S15 and Eq. S16, respectively. Both are based on the random errors in the y-direction, $s_{Y/X}$, which is calculated by Eq. S17. Lastly, the expanded combined uncertainty of each parameter is estimated from the t-distribution with N-2 degree of freedom, as Eq. S18 and Eq. S19 shows.

$$s_{p_1} = \frac{s_{Y/X}}{\sqrt{\sum_i (X_i - \bar{X})^2}} \tag{S15}$$

$$s_{p_2} = s_{Y/X} \sqrt{\frac{\sum_i (X_i)^2}{N \cdot \sum_i (X_i - \bar{X})^2}} \tag{S16}$$

$$s_{Y/X} = \sqrt{\frac{\sum_i (Y_i - \bar{Y})^2}{N - 2}} \tag{S17}$$

$$U_{comb}(p_1) = s_{p_1} \cdot t_{(N-2)} \tag{S18}$$

$$U_{comb}(p_2) = s_{p_2} \cdot t_{(N-2)} \tag{S19}$$

Based on the linear regression of the van't Hoff equation parameters and their respective confidence bounds, the expanded combined uncertainties of the energies of the solution are determined from the Eq. S20 to Eq. S22.

$$U_{comb}(\Delta G_{soln}) = R \cdot T_{hm} \cdot U_{comb}(p_1) \tag{S22}$$

$$U_{comb}(\Delta H_{soln}) = R \cdot U_{comb}(p_2) \tag{S21}$$

$$U_{comb}(\Delta S_{soln}) = \frac{U_{comb}(\Delta G_{soln}) + U_{comb}(\Delta H_{soln})}{T_{hm}} \tag{S22}$$

References

[1] A. Apelblat, E. Manzurola, 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 (1987) 317–320. https://doi.org/10.1016/0021-9614(87)90139-X.

[2] W.D. Bancroft, F.J.C. Butler, Solubility of Succinic Acid in Binary Mixtures, J. Phys. Chem. 36 (1932) 2515–2520. https://doi.org/10.1021/j150339a013.

[3] J.W. Mullin, Crystallisation, Butterworths, London, 1972.

[4] G. Massol, F. Lamouroux, Sur la solubilite dans l'eau des acides maloniques substitues, Comptes Rendus Hebd. Des Seances l'Academie Des Sci. 128, (1899) 1000–1002.

[5] J.W. Mullin, M.J.L. Whiting, Succinic Acid Crystal Growth Rates in Aqueous Solution, Ind. Eng. Chem. Fundam. 19 (1980) 117–121. https://doi.org/10.1021/i160073a020.

[6] J.H.C. Merckel, Die loslichkeit der dicarbonsauren, Recl. Des Trav. Chim. Des Pays-Bas. 56 (1937) 811–814.

[7] R. Wright, CLXXXVII.—Selective solvent action. Part VI. The effect of temperature on the solubilities of semisolutes in aqueous alcohol, J. Chem. Soc. (1927) 1334–1337. https://doi.org/10.1039/JR9270001334.

[8] S.H. Yalkowsky, Y. He, P. Jain, Handbook of Aqueous Solubility Data, 2nd ed., Boca Raton, Fla: CRC., 2010.

[9] R.M. Dawson, Data for Biochemical Research, Clarendon Press, Oxford, 1959.

[10] S.S. Doosaj, W.V. Bhagwat, Solubilities of weak acids in salts of weak acids at very high concentrations, J. Indian Chem. Soc. 10 (1933) 225–232.

[11] G.S. Forbes, A.S. Coolidge, Relations between distribution ratio, temperature and concentration in system: water, ether, succinic acid, J. Am. Chem. Soc. 41 (1919) 150–167. https://doi.org/10.1021/ja01459a004.

[12] R.K. Freier, Aqueous Solutions Volume 1: Data for Inorganic and Organic Compounds, Walter de Gruyter, New York, 1976.

[13] A.-P. Hyvärinen, H. Lihavainen, A. Gaman, L. Vairila, H. Ojala, M. Kulmala, Y. Viisanen, Surface Tensions and Densities of Oxalic, Malonic, Succinic, Maleic, Malic, and cis -Pinonic Acids, J. Chem. Eng. Data. 51 (2006) 255–260. https://doi.org/10.1021/je050366x.

[14] F. Lamouroux, Sur la solubilite dans l'eau des acides normaux de la serie oxalique, Comptes Rendus Hebd. Des Seances l'Academie Des Sci. 128 (1899) 998–1000.

[15] K. Linderstrom-Lang, Solubility of hydroquinone, Comptes Rendus Des Trav. Du Lab. Carlsbreg. 15, (1924) 4–28.

[16] H. Marshall, D. Bain, Sodium succinates, J. Chem. Soc. 97 (1910) 1074– 1085,.

[17] R.M.C. Dawson, D.C. Elliott, W.H. Elliott, K.M. Jones, Data for Biochemical Research, Oxford University Press, Pergamon, 1969.

[18] J.M. Weiss, C.R. Downs, The physical properties of maleic, fumaric and malic acids, J. Am. Chem. Soc. 45 (1923) 1003–1008. https://doi.org/10.1021/ja01657a018.

[19] B.N. Taylor, C.E. Kuyatt, Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, Gaithersburg, MD, 1994. http://physics.nist.gov/TN1297.

[20] J.N. Miller, J.C. Miller, Statistics and Chemometrics for Analytical Chemistry, Prentice Hall, Gosport, UK, 2010.