1. Introduction

Among numerous phenolic compounds present in vegetables, fruits and aromatic herbs, phenolic acids are an important class as they show relevant pharmacological, biological and organoleptic activities, usually presenting high-value added [1,2]. In this context, reliable solubility data of phenolic acids and derivatives in water and organic solvents are essential to the design of separation processes, such as extraction or crystallization, in the food, cosmetic, chemical and pharmaceutical industries [2,3].

Following our previous work on the solubility of hydroxybenzoic acids (gallic, α-resorcylic, gentisic and protocatechuic acids) [4], in this work our focus is on benzoic acids containing the methoxy group, namely syringic, vanillic and veratric acids. Interesting structural differences can be seen in Fig. 1, compared to vanillic acid (4-hydroxy-3-methoxybenzoic acid), in veratric acid (3,4-dimethoxybenzoic acid) a methoxy group replaces the hydroxyl group, while syringic acid (4-hydroxy-3,5-dimethoxybenzoic acid) has an additional methoxy group. Exploring their properties and applications, syringic acid presents strong antibacterial properties [5], skin carcinogenesis protective action [6], and potential effect on the prevention of diabetic cataract [7]. In addition, along with vanillic acid, syringic acid also presents potential hepatoprotective effects, suppressing hepatic fibrosis in injured livers [8]. Vanillic acid also presents biological properties: specific inhibitor of the snake venom enzyme 5′-nucleotidase [9], potential use to regulate chronic intestinal inflammation [10], and prevention of obesity by activating thermogenesis in brown and white adipose tissues [11]. Veratric acid can be used to treat skin disorders due to its ability of protecting UVB-induced skin injuries [12]. Additionally, studies show that veratric acid has significant antihypertensive, anti-inflammatory and antioxidant properties [13–15].

In this work, the solubility of the selected acids was measured in water and organic solvents (methanol, ethanol, 1-propanol, 2-propanol, 2-butanone, ethyl acetate, acetonitrile, dimethylformamide – DMF, 1,2-propanediol, 1,3-propanediol and 1,3-butanediol) at 298.2 K and 313.2 K. For many of the studied binary systems, no solubility data have been reported yet, but whenever possible, the new solubility data were
critically compared to the data available in literature. Besides a few studies concerning veratric [16,17], vanillic [18–21], and syringic acids [22–24], solubility, other works focused on aromatic acids with similar structures, such as o-anisic acid (2-methoxybenzoic acid) [25], p-anisic acid (4-methoxybenzoic acid) [25], 2,4-dimethoxybenzoic acid [26], 3,5-dimethoxybenzoic acid [27] and 3,4,5-trimethoxybenzoic acid [28]. In order to obtain a more complete picture of the solid-liquid equilibrium (SLE), the melting properties of the solutes were measured by Differential Scanning Calorimetry (DSC) as well as solid phase studies were carried out by X-Ray Diffraction (XRD).

The empirical selection of solvents only is quite restrictive, as it requires substantial experimental work and high cost [29]. An effective alternative is the use of predictive and semi-predictive thermodynamic models to estimate SLE diagrams. In this context, the NRTL-SAC model was successfully employed to predict SLE for several substituted benzoic acids [38–41]. Alternately, due to the high uncertainty of the melting enthalpies of the studied acids, the reference solvent approach (RSA) was combined with the NRTL-SAC model [32,34] under the same conditions as mentioned before. A third approach followed in this work was the application of the Abraham solvation model [35–37], which was already successfully employed to predict SLE for several substituted benzoic acids [38–47] in different organic solvents.

2. Experimental

2.1. Chemicals

In the solubility measurements of aqueous systems, ultrapure water (resistivity of 18.2 MΩ cm, free particles ≤0.22 μm and total organic carbon ≤5 μg·dm⁻³) was used. The solids were kept in a desiccator to avoid water contamination. All the compounds were used as received from suppliers and are listed in Table 1.

<table>
<thead>
<tr>
<th>Component</th>
<th>Mass purity (%)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syringic acid</td>
<td>99.1 ±0.1</td>
<td>Alfa Aesar</td>
</tr>
<tr>
<td>Vanillic acid</td>
<td>99.2 ±0.2</td>
<td>Merck KGaA</td>
</tr>
<tr>
<td>Veratric acid</td>
<td>99.2 ±0.3</td>
<td>Acros Organics</td>
</tr>
<tr>
<td>Methanol</td>
<td>≥99.0</td>
<td>Honeywell</td>
</tr>
<tr>
<td>Ethanol</td>
<td>≥99.0</td>
<td>Carlo Erba</td>
</tr>
<tr>
<td>Isopropanol</td>
<td>≥99.0</td>
<td>Honeywell</td>
</tr>
<tr>
<td>1-Propanol</td>
<td>≥99.5</td>
<td>Carlo Erba</td>
</tr>
<tr>
<td>2-Butanone</td>
<td>≥99.5</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>Ethyl acetate</td>
<td>≥99.7</td>
<td>Carlo Erba</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>≥99.9</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>DMF</td>
<td>≥99.9</td>
<td>Carlo Erba</td>
</tr>
<tr>
<td>1,2-Propanediol</td>
<td>≥99.5</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>1,3-Propanediol</td>
<td>≥99.8</td>
<td>DuPont</td>
</tr>
<tr>
<td>1,3-Butanediol</td>
<td>≥99.45</td>
<td>Sigma Aldrich</td>
</tr>
</tbody>
</table>

* The purity was obtained in the certificate of analysis issued by the manufacturer.

2.2. Solubility experiments

The solubility experiments were carried out by the isothermal shake-flask method, which is described thoroughly elsewhere [4,48]. Saturated solutions containing a small amount of solid in excess were prepared in a flask with around 50 ml of solvent. The flasks, covered with aluminum foil to avoid light degradation, were placed on plate stirrers inside a thermostatic bath (maximum temperature deviation of ±0.1 K). From preliminary experiments, it was found that the minimum stirring and settling times to reach the equilibrium were 30 h and 12 h, respectively.

After reaching equilibrium, three samples of around 4 cm³ were extracted from the mother solution by using plastic syringes coupled to a polypropylene filter (0.45 μm pore size). The gravimetric method was employed to quantify the solubilities in water, methanol, ethanol, 1-propanol, 2-propanol, 2-butanone, ethyl acetate, acetonitrile and DMF. In these cases, the solubilities were determined by weighting (Denver Instruments, precision of ±0.1 mg) the collected samples and the remissent solid after solvent evaporation.

For systems with 1,2-propanediol, 1,3-propanediol and 1,3-butanediol, solvents with higher boiling temperatures, the solubilities were determined by refractometry (Abbebat 500, Anton Paar) with a reproducibility within ±5 × 10⁻⁵. Each collected sample was first diluted in the same solvent, and the refractive index measured at least three times to calculate the solubility. The calibration curves (R² ≥ 0.998) used to correlate the concentration with the refractive index were obtained using six standard solutions with known compositions.

Since the solubility of veratric acid in water at 298.2 K is lower than 1 × 10⁻³ g/g of solvent, the gravimetric method is not suitable to determine the solubility. Alternatively, the solubilities were determined by UV–Vis spectroscopy at 258 nm (T70, PG Instruments). Each sample, after being diluted in a mixture of ethanol-water (65:35 mass ratio), was placed in cuvettes (5 mm optical path) and the absorbance read at least three independent times. As before, a calibration curve (R² ≥ 0.999) was used to calculate the solubility of veratric acid in water.

2.3. Melting properties

Melting temperatures and enthalpies were determined by DSC (204 F1 Phoenix, NETZSCH) using a nitrogen flowing system. Samples from 2 up to 10 mg (±0.1 mg) were hermetically sealed into aluminum crucibles and placed, along with a reference cell, to be heated or cooled at a rate of 1 K/min or 2 K/min, respectively. The experiments were performed from 293.15 K to 493.15 K for syringic and vanillic acids, and from 293.15 K to 473.15 K for veratric acid, where at least three runs were considered for the final results. An external calibration using 11 eleven compounds (ultra-pure water, 4-nitrotoluene, naphthalene, benzoic acid, diphenylacetic acid, indium, anthracene, tin, caffeine, bismuth and zinc) was performed for the two transition properties. In all cases, the onset value was considered as the melting temperature.

Fig. 1. Chemical structures of: (a) syringic acid; (b) vanillic acid and (c) veratric acid.
The phase change was also investigated by a visual method, which makes easier to identify any eventual degradation. Each sample was placed in an automatic glass capillary device (M-565 Büchi, 50–60 Hz, 150 W, temperature resolution: 0.1 K) and heated at the same conditions as employed for the DSC measurements. The melting temperatures were registered when the last crystal disappeared and the observed visual changes were recorded. For each compound, three independent measurements were also carried out.

2.4. Solid-phase studies

2.4.1. Samples

The pure acids from the manufacturer as well as the solids crystallized after evaporation of a set of selected solvents were analyzed by powder or single X-Ray diffraction.

2.4.2. Powder and single X-ray diffraction

Powder XRD data were collected on a X’Pert MPD Philips diffractometer, using Cu-Kα radiation (λ = 1.5406 Å), with a curved graphite monochromator, a set incident area of 10 mm², and a flat plate sample holder, in a Bragg–Brentano parafocusing optics configuration. Intensity data were collected by the step counting method (step 0.02° and time 5 s) in the range 5° to 50°. The pure acids from the manufacturer as well as the solids crystallized after evaporation of a set of selected solvents were analyzed by powder or single X-Ray diffraction.

The Abraham solvation model is based on two linear free energy relationships (LFERS), as described in detail by Abraham and co-workers [35–37]. One LFER quantifies the solute partition between two condensed phases (Eq. (4)) and, the other, the partition between a gas phase and an organic solvent (Eq. (5)):

\[
\log(P_S) = c + eE + sS + aA + bB + rV
\]

\[
\log(K_S) = c + eE + sS + aA + bB + lL
\]

In these equations, the uppercase descriptors \(E, S, A, B, V\) and \(L\) represent the Abraham solute descriptors, where \(E\) is the solute excess molar refraction, \(S\) refers to the solute dipolarity/polarizability, \(A\) and \(B\) account for the overall solute hydrogen bond acidity and basicity, \(V\) is the solute’s McGowan characteristic molecular volume and \(L\) is the logarithm of the gas-to-hexadecane partition coefficient at 298 K. The lower case regression coefficients and constants represent condensed phase properties, already established for a large number of solvents.

For the solubility calculations, only the first LFER quantifying the solute transfer between the organic solvent and water was considered in this work. In that case, the partition between water and a solvent \(P_S\) is defined as the ratio between the molar solubilities in the organic solvent \(S_o\) and in water \(S_w\) as follows:

\[
P_S = \frac{S_o}{S_w}
\]

As discussed by Abraham et al. [37], the application of Eq. (6) is subject to some constraints, namely: (a) the solid phase in equilibrium with both water and organic solvent is the same; (b) the secondary medium activity coefficient of the solute in the two phases is near unity; the same (undissociated, if ionizable) chemical species should be present in each phase.

Excepting two solvents (1,3-propanediol and 1,3-butanediol), the Abraham descriptors of the solvents studied in this work are available or present high uncertainty, which lead us to combine the NRTL-SAC model to the Reference Solvent Approach (RSA) proposed by Abildskov and O’Connell [34,52]. This methodology can be briefly described by the following equation:

\[
\ln x_{Si} = \ln x_{Sj} + \ln y_{Si}(T_{\{x_{Si}\}}) - \ln y_{Sj}(T_{\{x_{Sj}\}})
\]

where \(x_{Si}\) is the mole fraction solubility of the solute \(S\) in a pure solvent \(i\), \(x_{Sj}\) is the solubility of the same solute in a pure reference solvent \(j\), \(y_{Si}\) is the activity coefficient of the solute in solvent \(i\), while \(y_{Sj}\) is the activity coefficient of the solute in the reference solvent \(j\) and \(T\) is the temperature of the system.

In this methodology, the experimental solubility of a solute in a chosen reference solvent is used along with the NRTL-SAC model to determine the other variables in Eq. (2). For a given set of data, the optimal reference solvent is found by:

\[
\min \sum_{i=\text{data}} \delta\ln x_{Si,j} = \min \sum_{i=\text{data}} (\ln x_{Si} + \ln y_{Si}) - N(\ln x_{Sj} + \ln y_{Sj})
\]

where \(\sum \delta\ln x_{Si,j}\) is the error associated to the mole fraction solubilities of solute \(S\) in all the solvents assuming a reference solvent \(j\) and \(N\) is the number of the experimental data points in a given set.

3. Thermodynamic modeling

3.1. The NRTL-SAC model

Due to its semi-predictive nature and successful application in previous studies [29–32,48–50], the NRTL-SAC model was chosen to describe the solubility of the solutes studied in this work in pure solvents. A detailed description of the model, with its fundamentals and equations, is presented elsewhere [30,49]. Briefly, each molecule is described by four conceptual segments (hydrophobic \(X\), hydrophilic \(Z\), polar attractive \(Y^+\), and polar repulsive \(Y^-\)), representing different molecular surface interactions. For 63 organic solvents, these parameters are already available in the literature [30,49]. Therefore, only the molecular descriptors of the solute are required to calculate its solubility in a pure solvent.

Assuming pure solid phase and that the triple point can be calculated by the melting point, and neglecting the heat capacities change, the solubility of a solid solute in a liquid solvent can be determined by the following equation [51]:

\[
\ln x_S = \frac{\Delta_m H}{R T_m} \left(1 - \frac{T}{T_m}\right) - \ln y_S
\]

where \(x_S\) is the mole fraction solubility of the solute \(S\), \(R\) is the ideal gas constant, \(T\) is the absolute temperature, \(T_m\) is the melting temperature of the solute, \(\Delta_m H\) its melting enthalpy, and \(y_S\) is the activity coefficient of the solute \(S\), calculated in this case using the NRTL-SAC model [30].

In Eq. (1), the accuracy of the melting properties plays a crucial role for the reliability of the estimates. However, in many cases these properties are unavailable or present high uncertainty, which lead us to combine the NRTL-SAC model to the Reference Solvent Approach (RSA) proposed by Abildskov and O’Connell [34,52]. This methodology can

\[
\text{be briefly described by the following equation:}
\]

\[
\ln x_{Si} = \ln x_{Sj} + \ln y_{Si}(T_{\{x_{Si}\}}) - \ln y_{Sj}(T_{\{x_{Sj}\}})
\]

where \(x_{Si}\) is the mole fraction solubility of the solute \(S\) in a pure solvent \(i\), \(x_{Sj}\) is the solubility of the same solute in a pure reference solvent \(j\), \(y_{Si}\) is the activity coefficient of the solute in solvent \(i\), while \(y_{Sj}\) is the activity coefficient of the solute in the reference solvent \(j\) and \(T\) is the temperature of the system.

In this methodology, the experimental solubility of a solute in a chosen reference solvent is used along with the NRTL-SAC model to determine the other variables in Eq. (2). For a given set of data, the optimal reference solvent is found by:

\[
\min \sum_{i=\text{data}} \delta\ln x_{Si,j} = \min \sum_{i=\text{data}} (\ln x_{Si} + \ln y_{Si}) - N(\ln x_{Sj} + \ln y_{Sj})
\]

where \(\sum \delta\ln x_{Si,j}\) is the error associated to the mole fraction solubilities of solute \(S\) in all the solvents assuming a reference solvent \(j\) and \(N\) is the number of the experimental data points in a given set.

3.2. The Abraham solvation model
In addition, parameter $E$ can also be calculated from the solute's refractive index, which can be experimentally obtained for liquid solutes or estimated using ACD free software for solid solutes, as described in detail by Abraham et al. [36].

4. Results and discussion

4.1. Experimental Solubilities

The measured solubilities of syringic acid, veratric acid and vanillic acid in water, methanol, ethanol, 1-propanol, 2-propanol, ethyl acetate, 2-butanoate, acetonitrile, DMF, 1,2-propanediol, 1,3-propanediol and 1,3-butanediol at 298.2 K and 313.2 K are presented in Table 2.

Each data point presented in Table 2 is an average of at least 3 independent measurements. The low variation coefficients, always inferior to 3.8%, indicate that the standard deviations are much lower than the solubility values. The solubility of the acids is, in general, considerably higher in organic solvents than in water. In nine of the studied binary systems (water, methanol, ethanol, 1-propanol, 2-propanol, DMF, 1,2-propanediol, 1,3-propanediol and 1,3-butanediol), the solubility at 298.2 K ranks: vanillic acid > syringic acid > veratric acid. With the exception of DMF, all those solvents present at least one hydroxyl group, capable of forming stronger hydrogen bonds with the carboxylic group, and also to the hydroxyl groups, which are more accessible in vanillic acid, while veratric acid only presents two hydrogen bond acceptor methoxy substituents.

For ethyl acetate and 2-butanoate, vanillic acid is still the most soluble one, but less polar veratric acid presents higher solubilities than syringic acid. In acetonitrile, the most soluble compound is veratric acid, followed by syringic acid, and finally vanillic acid. In fact, the importance of hydrogen bonds is now weaker (the solvents contain only hydrogen bond acceptor groups) than in the previous set of solvents, contributing to the changes observed in the solubility order.

In the case of systems containing alcohols, the solubility decreases with an increase of the alkyl chain of the alcohol. Surprisingly, much higher solubilities were found for systems with DMF, behavior that was also observed for dihydroxybenzoic acids [4]. In all cases, the solubility increases with temperature.

4.2. Data analysis

Whenever possible, the solubilities of the selected acids were compared to literature data. In Fig. 2, a comparison of the solubilities of syringic, vanillic and veratric acids in water and ethanol is presented. A complete overview of this coherence analysis is presented in Table S1 and in Figs. S1–S4 of Supporting Information. Fig. 2 shows that in general the solute solubility in water and ethanol agrees with the data found in literature, presenting deviations lower than 1.0 g/100 g of solvent, with the exception of the system syringic acid/ethanol, where the solubility measured by Noubigh et al. [24] is much higher, while in water [24] is 3 to 4 times higher than the solubility measured in this work and by Queimada et al. [24]. Due to the higher disagreement observed in the systems containing ethyl acetate and ethanol, solubility experiments were here performed twice, reaching always very similar solubility values between the two independent measurements.

For vanillic acid, the solubilities in water (0.128 g/100 g of water at 298.2 K and 0.269 g/100 g of water at 313.2 K) measured in this work are quite similar to the literature average values (0.151 g/100 g of water and 0.271 g/100 g of water at 298.2 K and 313.2 K, respectively). However, the differences concerning the solubility in some organic solvents are relevant, especially at 313.2 K. Solubility experiments were repeated for systems containing ethyl acetate, methanol and 2-propanol and no significant changes were observed.

Particularly, the solubility of vanillic acid in methanol published by Noubigh and Abderrabba [20] is much higher than the values found in this work. In that study, the solubility of vanillic acid in methanol at 298.2 K is 7.1 times higher than the solubility of the same compound in ethanol, which is quite unexpected. For instance, in previous studies [24,56] the same group reported, for syringic and protocatechuic acids, solubility values 1.8 and 1.6 times higher in methanol than in ethanol, respectively.

For veratric acid, the solubility in water at 298.2 K (0.051 g/100 g of solvent) measured in this work is comparable to the value reported by Bowen and coworkers [17] (0.058 g/100 g of solvent). In organic solvents, the solubilities obtained in this work are also in good agreement with the data reported by Li et al. [16] and Bowen et al. [17]. In the first case, the authors employed the robust laser monitoring method, that dynamically detects the temperature at which the last solid particles disappeared, eliminating the issues related to the chemical analysis, stirring and settling times.

4.3. Melting properties

The melting temperatures and enthalpies found in literature [16,21,23,57–60] along with the values measured in this work by DSC, and by the visual method, are summarized in Table 3. Exemplificative thermograms for the DSC measurements are shown in Fig. S5 and Fig. S6 of the Supporting Information.

The melting temperatures obtained by DSC in this work agree very well with the values from the literature. The melting temperatures measured by the visual method are slightly higher than the values measured by DSC, probably due to the employed methodologies. While in the visual method the melting temperature was considered as the point in which the last crystal disappears, the onset temperatures were registered in the DSC measurements.

The melting enthalpies from this work, however, present less consistency with the values reported by other authors, being larger for all the

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Syringic acid</th>
<th>Vanillic acid</th>
<th>Veratric acid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>298.2 K</td>
<td>313.2 K</td>
<td>298.2 K</td>
</tr>
<tr>
<td>Water</td>
<td>0.142 ± 0.001</td>
<td>0.231 ± 0.004</td>
<td>0.128 ± 0.002</td>
</tr>
<tr>
<td>Methanol</td>
<td>11.480 ± 0.036</td>
<td>16.237 ± 0.028</td>
<td>18.264 ± 0.035</td>
</tr>
<tr>
<td>Ethanol</td>
<td>5.562 ± 0.008</td>
<td>8.006 ± 0.019</td>
<td>11.947 ± 0.015</td>
</tr>
<tr>
<td>2-Propanol</td>
<td>2.294 ± 0.010</td>
<td>3.739 ± 0.002</td>
<td>7.009 ± 0.016</td>
</tr>
<tr>
<td>1-Propanol</td>
<td>2.593 ± 0.008</td>
<td>4.083 ± 0.008</td>
<td>6.435 ± 0.009</td>
</tr>
<tr>
<td>2-Butanoate</td>
<td>2.658 ± 0.012</td>
<td>3.665 ± 0.006</td>
<td>5.516 ± 0.026</td>
</tr>
<tr>
<td>Ethyl acetate</td>
<td>1.006 ± 0.008</td>
<td>1.446 ± 0.005</td>
<td>2.193 ± 0.007</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>0.951 ± 0.009</td>
<td>1.665 ± 0.003</td>
<td>1.383 ± 0.018</td>
</tr>
<tr>
<td>DMF</td>
<td>66.963 ± 0.009</td>
<td>81.527 ± 0.087</td>
<td>88.945 ± 0.220</td>
</tr>
<tr>
<td>1,2-Propanediol</td>
<td>3.894 ± 0.047</td>
<td>5.808 ± 0.054</td>
<td>7.514 ± 0.129</td>
</tr>
<tr>
<td>1,3-Propanediol</td>
<td>4.353 ± 0.028</td>
<td>6.193 ± 0.045</td>
<td>7.377 ± 0.096</td>
</tr>
<tr>
<td>1,3-Butanediol</td>
<td>3.757 ± 0.026</td>
<td>5.784 ± 0.033</td>
<td>6.643 ± 0.149</td>
</tr>
</tbody>
</table>

*Temperature and pressure standard uncertainties are $u(T) = 0.10 K$ and $u(p) = 0.05$, respectively. Standard deviations are placed after plus-minus sign.
studied solutes. This fact could be related to possible degradation, as discussed below.

DSC experiments detected that syringic acid presents a solid-solid transition at \((462.6 \pm 1.1) K\), before reaching the fusion transition. This behavior is reproducible for runs with independent samples, but it was not observed in experiments with successive runs. The visual method gives some insights on possible degradation of syringic acid upon melting, due to a slight modification in the solid coloration from white to a light yellow.

In the case of vanillic acid, the melting enthalpy reported by Manic et al. [57] is closer to the value obtained by us. Nevertheless, possible decomposition of the sample was identified by the visual method starting at \(481.6 K\). An illustrative image of this phenomena is shown in Fig. S7 of Supporting Information.

From the three solutes, veratric acid was the only one which did not present any modification on the fusion peak shapes over successive runs or any visual degradation. For this compound, a transition was identified at \((422.9 \pm 0.4) K\), reproducible when the heating rate was set as 1 K/min. Even so, the melting enthalpy measured is 5 kJ/mol higher than that measured by Li et al. [16].

4.4. Solid phase studies

Solid samples of syringic, vanillic and veratric acids as received from suppliers and those recrystallized from the saturated solutions of water, methanol, 2-butanone, acetonitrile and dimethylformamide were analyzed by powder and single crystal X-ray diffraction.

From a single crystal, it was found that syringic acid received from the supplier as well syringic acid recrystallized from 2-butanone, acetonitrile and dimethylformamide crystalized in monoclinic \(C\) system with cell parameters of 
\[ a = 26.64 \text{ Å}, \quad b = 4.14 \text{ Å}, \quad c = 15.84 \text{ Å}, \quad \beta = 96.30°, \]
which are comparable to those published in CCDC database (CCDC

---

**Fig. 2.** Comparison of the experimental solubility data obtained in this work with literature data [16,17,19–24]: (a) syringic acid water; (b) vanillic acid water; (c) veratric acid water; (d) syringic acid ethanol; (e) vanillic acid ethanol; (f) veratric acid ethanol.
The comparison of the powder X-ray diffraction patterns of all the vanillic acid samples shown in Fig. S9 of SI reveals that they are this work.

Table 3
The melting temperatures and enthalpies of the studied acids from the literature [16,21,23,57,58,61] and measured in this work.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$T_m$/°C</th>
<th>$\Delta H$/kJ mol$^{-1}$</th>
<th>Methodology</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syringic acid</td>
<td>480.3 ± 0.6</td>
<td>33.7 ± 1.8</td>
<td>DSC</td>
<td>[23]</td>
</tr>
<tr>
<td></td>
<td>482.5</td>
<td>28.1</td>
<td>DSC</td>
<td>[59]</td>
</tr>
<tr>
<td></td>
<td>481.6</td>
<td>–</td>
<td>DSC</td>
<td>[60]</td>
</tr>
<tr>
<td></td>
<td>482.3 ± 0.1</td>
<td>–</td>
<td>Visual Method</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td>480.9 ± 0.3$^a$</td>
<td>40.3 ± 0.6</td>
<td>DSC</td>
<td>This work</td>
</tr>
<tr>
<td>Vanillic acid</td>
<td>484.9 ± 0.2</td>
<td>32.8 ± 0.1</td>
<td>DSC</td>
<td>[21]</td>
</tr>
<tr>
<td></td>
<td>484.7</td>
<td>25.6</td>
<td>DSC</td>
<td>[59]</td>
</tr>
<tr>
<td></td>
<td>480.7 ± 0.2</td>
<td>29.1 ± 0.6</td>
<td>DSC</td>
<td>[57]</td>
</tr>
<tr>
<td></td>
<td>481.15</td>
<td>–</td>
<td>DSC</td>
<td>[58]</td>
</tr>
<tr>
<td></td>
<td>484.9 ± 0.12</td>
<td>–</td>
<td>Visual Method</td>
<td>This work</td>
</tr>
<tr>
<td>Veratric acid</td>
<td>481.3 ± 0.3$^a$</td>
<td>34.63 ± 0.8</td>
<td>DSC</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td>453.1 ± 0.2</td>
<td>29.6 ± 0.1</td>
<td>DSC</td>
<td>[16]</td>
</tr>
<tr>
<td></td>
<td>455.75 ± 0.17</td>
<td>–</td>
<td>Visual Method</td>
<td>This work</td>
</tr>
<tr>
<td></td>
<td>453.5 ± 1.3$^a$</td>
<td>34.64 ± 0.3</td>
<td>DSC</td>
<td>This work</td>
</tr>
</tbody>
</table>

$^a$ The experimental onset temperatures were considered as melting temperatures in this work.

To begin, the experimental solubilities in seven solvents (water, methanol, ethanol, 2-propanol, ethyl acetate, acetonitrile and 2-butanol) were used to correlate the four conceptual segment parameters ($X, Y, Z, Y^o - Z$) for each solute. After, those were used to predict the solubility in 2-propanol and DMF. The average relative deviation (ARD) was calculated for each binary system as:

$$ARD(\%) = \frac{1}{NP} \sum_{i} \left| \frac{x_{i}^{exp} - x_{i}^{calc}}{x_{i}^{exp}} \right| \times 100$$

where $NP$ is the number of data points.

As mentioned before, the high uncertainty observed in the fusion properties, particularly regarding the enthalpies, can lead to difficulties or even hamper the parameter estimation. For that reason, the RSA proposed by Abildskov and O’Connell [34,52] was combined with the NRTL-SAC model. The optimized molecular descriptors of the solutes obtained for both correlation approaches (NRTL-SAC or NRTL-SAC + RSA) as well as the reference solvent (when applicable), the outlier solvent (presenting the highest ARD) and the global ARD are presented in Table 4.

The results above show that NRTL-SAC is an adequate model to correlate the solubility of the studied compounds, presenting ARDs varying between 20 and 41%, which is very satisfactory for a semi-predictive model with a reduced number of parameters to be estimated for each solute. The combination of the NRTL-SAC model with RSA also showed reliable correlation, resulting in lower ARD for syringic and vanillic acids. In previous works, the authors have applied the NRTL-SAC model to predict the solubility of drug molecules in water and organic solvents [4,31,32], obtaining ARD similar to those found in this work. Fig. 3 relates the calculated and experimental solubility data for the correlations performed by NRTL-SAC and NRTL-SAC + RSA.

The NRTL-SAC segment descriptors obtained by the conventional approach and with the use of RSA are in closer agreement for veratric acid, the only compound that showed good thermal stability upon melting.

As can be seen in Fig. 3, the NRTL-SAC and NRTL-SAC + RSA describe well the solubility in water (maximum ARD of 22%). The results obtained using NRTL-SAC only present higher ARDs for the systems containing methanol and 2-butane. For a few cases, the model is limited in giving the exact order of magnitude of the solubilities, leading to lower calculated values. On the other hand, the correlations performed using NRTL-SAC + RSA represent better the solubilities in alcohols, but delivering higher ARDs in the cases of ethyl acetate and 2-butane.

After, solubility predictions in 1-propanol and DMF were carried out. The obtained ARDs were 22% and 15% for 1-propanol, and 94% and 73% for DMF, applying NRTL-SAC or NRTL-SAC + RSA, respectively. Although the difference was not so pronounced in the correlation step, the ARDs found are considerably lower when the NRTL-SAC + RSA estimated parameters are used.

The prediction analysis was also extended to the binary systems previously studied by different authors [17,20,21,24]. In Fig. 4, a summary of predictions found using NRTL-SAC and NRTL-SAC + RSA is presented.

Regarding the NRTL-SAC predictions, the highest ARDs are observed for ethylene glycol, isobutanol and dimethylformamide, whereas the model provides good estimates for the solubilities in propyl acetate and for alcohols. In the case of NRTL-SAC RSA, the three highest predictions are the systems with tetrahydrofuran, 1,4-dioxane and 2-butanol. All the systems with other alcohols and propyl acetate are also well described by this methodology. For both, NRTL-SAC and NRTL-SAC + RSA approaches, the ARD for all the predictions is 59%, and the models generally underestimate the calculated solubilities.

4.6. Abraham solvation model

Since no modification in the crystalline structures was observed by X-ray diffraction and the measured solubilities were generally not too high, the Abraham solvation model was applied to describe the SLE of

$$F = \sum_{i} \left( \frac{x_{i}^{exp} - x_{i}^{calc}}{x_{i}^{exp}} \right)^2$$

where $x_i$ is the mole fraction solubility in the solvent $i$ and the upper scripts “exp” and “calc” mean the experimental and calculated values, respectively.
the studied acids. To simultaneously solve the set of LFERs (Eq. (4)) for each solute, the multiple linear regression model was applied. Initially, the same set of solvents (water, methanol, ethanol, 2-propanol, 2-butanol, ethyl acetate and acetonitrile) employed in NRTL-SAC was selected to correlate the solutes parameters. All experimental solubilities were converted from mole fraction \( (X_s) \) to molar solubilities \( (S_s) \) applying:

\[
S_s^{\exp} = \frac{X_s^{\exp} V_{\text{Solute}}}{(1 - X_s^{\exp}) V_{\text{Solvent}}}
\]  

(9)

The molar volumes of the solutes were calculated from the solid densities obtained by crystallographic studies [60,62,63]. The values for syringic, vanillic and veratric acids are 134.0 cm\(^3\) mol\(^{-1}\), 115.5 cm\(^3\) mol\(^{-1}\) and 126.9 cm\(^3\) mol\(^{-1}\), respectively. For the solvents, the molar volumes were calculated based on the density values found in literature (Table S2 of Supporting information). The obtained ARDs for the correlation step were 21% for syringic acid, 15% for vanillic acid and 5% for veratric acid. Predictions were also performed for the solubility data measured in this work (1-propanol, DMF, 1,2-propanediol) and for the data available in literature [17,20,21,24], showing a global ARD of 43%. For these first set of simulations, the parameters obtained are in Table S3 of SI.

The average deviations are in the expected range for semi-predictive models, but for the solubility in DMF the estimates are really very poor. Therefore, it was decided to test the impact of including the solubility data in DMF during the correlation step. After re-estimating the solute descriptors, the ARD found for each solute was 20% for syringic acid, 13% for vanillic acid and 11% for veratric acid, while for predictions the global ARD, also considering the literature data, was 26%. Table 5 presents the parameters obtained in this second correlation round. The correlation deviations for veratric acid including DMF is slightly higher, but its reduction in the predictions is much more significant (from 43% to 26%). Since DMF is structurally different from the other solvents used, its presence in the correlation step provides more robust solute parameters, improving the accuracy of the predictions for a larger number of solvents.

The importance of the chemical diversity of the solvents in the reliability of the fitted parameters has been shown previously [17,37,44]. Following this idea, a similar approach was performed for the NRTL-SAC and NRTL-SAC RSA simulations, incorporating DMF in the parameter fitting. However, no considerable changes were observed neither in the optimized solute descriptors or global ARDs. In Table 5, the estimated Abraham solute descriptors (including DMF), global correlation ARD and the outliers are presented.

To the best of our knowledge, among the acids here studied the Abraham solvation model was only used to describe the SLE of veratric acid [17,64]. In this case, the authors employed a set of 54 partition data including condensed phase and gas-liquid data, a considerable larger number than we used in this work. The solute parameters reported by Bowen et al. [17] are very close to the values presented in Table 5. In Figs. 5 and 6, a summary of the correlation and prediction results is presented, relating the calculated solubility to the experimental solubility data.

The outliers in the predictions were 2-butanol and dibutyl ether (ARD of 79% and 53% respectively). For all the other predicted solvents, the ARD found were inferior to 40%. Particularly, for 1,2-propanediol, the Abraham solvation model provides reasonable predictions (ARD of 39%) even though no diol was used in the fit. In general, the model describes well the SLE of binary systems containing alcohols (ARD of 20%) and esters (ARD of 19%).

Despite the Abraham solute parameters describe quite well the SLE of a large number of binary systems, it is relevant to analyze the descriptor’s values to check if they reflect the chemical properties of the solute. Hoover and coworkers [38] previously correlated the Abraham solute’s descriptors for 2-methoxybenzoic acid \((S = 1.410, A = 0.450 \text{ and } B = 0.620)\) and 4-methoxybenzoic acid \((S = 1.250, A = 0.620 \text{ and } B = 0.520)\) by using a similar approach employed in this

---

**Table 4**

NRTL-SAC optimized parameters, reference solvent, outlier solvent and ARD (%) for each acid using water and six organic solvents in the fitting.

<table>
<thead>
<tr>
<th>Compound</th>
<th>X</th>
<th>Y-</th>
<th>Y+</th>
<th>Z</th>
<th>Reference solvent</th>
<th>Outlier</th>
<th>ARD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syringic acid</td>
<td>0.287</td>
<td>0.840</td>
<td>0.642</td>
<td>0.000</td>
<td>–</td>
<td>Methanol</td>
<td>41</td>
</tr>
<tr>
<td>Vanillic acid</td>
<td>0.424</td>
<td>0.885</td>
<td>0.304</td>
<td>0</td>
<td>–</td>
<td>Methanol</td>
<td>38</td>
</tr>
<tr>
<td>Veratric acid</td>
<td>0.495</td>
<td>0.000</td>
<td>0.202</td>
<td>0.531</td>
<td>–</td>
<td>2-Butanone</td>
<td>20</td>
</tr>
<tr>
<td>Syringic acid + RSA</td>
<td>0.495</td>
<td>0.000</td>
<td>0.400</td>
<td>1.625</td>
<td>Methanol</td>
<td>Ethyl acetate</td>
<td>35</td>
</tr>
<tr>
<td>Vanillic acid</td>
<td>0.587</td>
<td>0.059</td>
<td>0.000</td>
<td>1.405</td>
<td>Acetonitrile</td>
<td>Ethyl acetate</td>
<td>33</td>
</tr>
<tr>
<td>Veratric acid</td>
<td>0.491</td>
<td>0.000</td>
<td>0.239</td>
<td>0.561</td>
<td>Ethanol</td>
<td>2-Butanone</td>
<td>22</td>
</tr>
</tbody>
</table>

---

**Fig. 3.** Comparison between experimental and calculated solubility for correlation: a) NRTL-SAC; b) NRTL-SAC RSA.
work (the authors used solubility data in alcohols and in ethers to fit the solute parameters).

For syringic, vanillic and veratric acids, the parameters $S$, $A$ and $B$ are higher than the parameters reported for 2-methoxybenzoic and 4-methoxybenzoic acids, even though the value order changes for each parameter. In the case of the acidity descriptor ($A$), the presence of hydroxyl groups increases the H-bond acidity of the solute, whereas intramolecular hydrogen bonds tend to reduce it. Vanillic acid (one hydroxyl and one methoxy substituents attached to the aromatic ring) presents the highest value for $A$, followed by syringic acid (one hydroxyl and two methoxy substituents) and veratric acid (two methoxy groups). The close proximity of the hydroxyl and methoxy groups in syringic and vanillic acids likely leads to the formation of intramolecular H-bonds, being stronger for syringic acid.

The absence of a hydroxyl-substituent in veratric acid reduces its acidity compared to the other two studied phenolic acids. The correlated value for this compound ($A = 0.630$) is quite close to the value found for 4-methoxybenzoic acid but considerably higher than the value for 2-methoxybenzoic-acid. The methoxy-substituents in the aromatic ring should increase the electron density of the aromatic ring through resonance, increasing the H-bond acidity of the carboxyl group. The presence of a methoxy group in the position 2 of the ring, however, can also lead to intramolecular hydrogen-bond formation, reducing the acidity of this acid compared to veratric acid and 4-methoxybenzoic acid.

Regarding the solute’s basicity, the values of $B$ progressively increase in the following order: syringic acid $>$ veratric acid $>$ vanillic acid. In the case of syringic acid, the presence of two methoxy-substituent and one hydroxyl-substituent groups leads to a higher number of available lone electron pairs on the oxygen atoms than it is observed for the other two acids. In the case of vanillic acid, the presence of intramolecular H-bonds should decrease the basicity of the substituents (methoxy and hydroxyl groups). As expected, the basicity parameters reported for 2-methoxybenzoic acid and 4-methoxybenzoic acid were lower than the values correlated to syringic, vanillic and veratric acids, which is probably related to the lower electron density present in the monosubstituted methoxybenzoic acids.

Even if the parameters give good consistency with the chemical structure of the solutes, it is relevant to mention that they are an average representation of the different conformations of the solute in each solvent, which are also different among all the solvents used for correlation [38].

5. Conclusions

In this work, the solubility of syringic acid, vanillic acid and veratric acid was measured in water and in eleven organic solvents at 298.2 and 313.2 K using the isothermal shake-flask method. For all the studied binary systems, an increment on the solubility was observed when increasing the temperature. On all the possible occasions, the measured solubility was compared to the data available in literature. In the case

<table>
<thead>
<tr>
<th>Compound</th>
<th>$E$</th>
<th>$S$</th>
<th>$A$</th>
<th>$B$</th>
<th>$V$</th>
<th>Outlier</th>
<th>ARD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syringic acid</td>
<td>1.123</td>
<td>1.757</td>
<td>0.808</td>
<td>0.878</td>
<td>1.390</td>
<td>Methanol</td>
<td>20</td>
</tr>
<tr>
<td>Vanillic acid</td>
<td>1.144</td>
<td>1.452</td>
<td>0.846</td>
<td>0.647</td>
<td>1.190</td>
<td>Methanol</td>
<td>13</td>
</tr>
<tr>
<td>Veratric acid</td>
<td>0.881</td>
<td>1.646</td>
<td>0.630</td>
<td>0.733</td>
<td>1.331</td>
<td>2-Butanone</td>
<td>11</td>
</tr>
</tbody>
</table>

Fig. 4. Comparison between experimental and calculated solubility data for prediction: a) NRTL-SAC; b) NRTL-SAC RSA.

Fig. 5. Comparison between experimental and calculated solubility in the organic solvents used in the estimation of the Abraham solute parameters.
of veratric acid, the solubilities measured in this work are in good agreement with the literature data. For vanillic and syringic acids, some inconsistencies were found.

Melting properties measurements of the three acids were also carried out by DSC and by a visual capillary method. For all acids, the melting temperatures determined by DSC strongly agree with the literature. The visual method gave some indications about degradation of syringic and vanillic acids upon the melting, which probably explains the higher uncertainties observed for the melting enthalpies. Additionally, solid-solid transitions, with small phase change enthalpies were identified and vanillic acids upon the melting, which probably explains the higher temperatures determined by DSC strongly agree with the literature.

Consistencies were found.

With the literature data. For vanillic and syringic acids, some inconsistencies were found. The global ARD in the correlation and prediction steps were 15% and 26%, respectively.

Generally, the NRTL-SAC and Abraham models provided satisfactory results, allowing to estimate the solubility order of magnitude in several solvents by using a limited set of experimental data. However, Abraham solvation model provides more accurate solubility values than NRTL-SAC. On the other hand, the NRTL-SAC ability to calculate the variation of the solubility with temperature is a very important feature.

Acknowledgments

We acknowledge the support of the project “AlPocMat@N2020 – Advanced Industrial Processes and Materials for a Sustainable Northern Region of Portugal 2020”, with the reference NORTE-01-0145-FEDER-000006, supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the Portugal 2020 Partnership Agreement, through the European Regional Development Fund (ERDF); Associate Laboratory LSRE-UCM – UID/EQUI/50020/2019 – funded by national funds through FCT/MCTES; and project AllNat – POCI-01-0145-FEDER-030463 (PTDC/EQUI-EPQ/30463/2017), financed by national funds through FEDER funds through COMPETE and Portugal2020 and national funds through FCT - Fundação para a Ciência e a Tecnologia.

The authors also thank FCT for financial support to S. M. Vilas-Boas grant (SFRH/BD/138149/2018) and V. Vieira grant (SFRH/BD/108487/2015).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.molliq.2019.111089.

References


