Modeling the Solubility of Monoterpenoids with Hybrid and Predictive Thermodynamic Tools

Sérgio M. Vilas-Boas, Isabella W. Cordova, Dinis O. Abranches, João A. P. Coutinho, Olga Ferreira,* and Simão P. Pinho*

Cite This: https://doi.org/10.1021/acs.iecr.2c03991

1. INTRODUCTION

Monoterpenoids constitute a structurally diverse class of compounds abundantly found in essential oils. Due to their often pleasant aromas and diverse biological and pharmacological activities, monoterpenoids have been increasingly exploited in the fragrance, cosmetic, food, and pharmaceutical industries. To design and optimize extraction and purification processes to obtain monoterpenoids from their natural matrices, the knowledge of equilibrium properties, such as solubility and partition coefficient data, in different solvent systems is extremely valuable. From the pool of relevant monoterpenoids, four commercially important representatives were selected: (−)-borneol, (1R)-(−)-camphor, 1-(−)-menthol, and thymol.

Camphor is one of the chief representatives of bicyclic monoterpenes, having several applications in the pharmaceutical field. Present in Cinnamomum camphora species, camphor has been widely used in traditional medicine to treat several diseases, such as rheumatism, chest congestion, muscle pain, asthma, and bronchitis. In modern medicine, this substance is an ingredient of analgesics and rubefacients, although its overdose ingestion or skin penetration might cause severe toxic effects or even death. Another bicyclic monoterpenoid, borneol, is a monoterpenic alcohol used as a fragrance ingredient in personal care and cleaning products. This compound also presents relevant pharmaceutical and biological properties, such as analgesic, antibacterial, anti-inflammatory, and antioxidant activities. It has been extensively investigated as a potential permeation enhancer to facilitate drug delivery through different biological barriers.

Menthol is widely employed as an aroma ingredient in cosmetics, perfumes, household cleaners, and detergents. It is a cyclic monoterpenic alcohol produced by numerous plants from the Lamiaceae family. It is broadly used in medicine to relieve localized pain and respiratory disorders due to its “cool” characteristic sensation triggered by its interactions with thermoreceptors in the human skin. Recently, the potential of menthol and thymol for the formulation of some innovative deep eutectic solvents has been demonstrated. Thymol, a phenolic monoterpenoid, presents various applications in medicine, dentistry, food, and agrochemical industries. Because of its numerous bioactivities (antioxidant, anti-inflammatory, antimicrobial, and antifungal), thymol has been extensively studied as a therapeutic agent for wound healing. Besides, thymol brings several nutritional and...
Table 1. Chemical Structure, CAS Number, Source, Mass Purity, Melting Temperature ($T_m$), and Melting Enthalpy ($\Delta_m H$) of the Solutes and Organic Solvents Studied in This Work

<table>
<thead>
<tr>
<th>Compound</th>
<th>Structure</th>
<th>CAS number</th>
<th>Source</th>
<th>Purity (mass fraction)</th>
<th>Water content (mass fraction)</th>
<th>$T_m$ (K)</th>
<th>$\Delta_m H$ (kJ·mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(−)-borneol</td>
<td><img src="structure1.png" alt="Structure" /></td>
<td>464-45-9</td>
<td>Alfa Aesar</td>
<td>≥ 0.970</td>
<td>0.0022</td>
<td>480±3</td>
<td>7.3±0.3</td>
</tr>
<tr>
<td>(1R)-(−)-camphor</td>
<td><img src="structure2.png" alt="Structure" /></td>
<td>464-49-3</td>
<td>Alfa Aesar</td>
<td>≥ 0.980</td>
<td>0.0024</td>
<td>451±2</td>
<td>6.2±0.6</td>
</tr>
<tr>
<td>L-(−)-menthol</td>
<td><img src="structure3.png" alt="Structure" /></td>
<td>2216-51-5</td>
<td>Alfa Aesar</td>
<td>≥ 0.990</td>
<td>0.0013</td>
<td>315.6±0.02</td>
<td>13.47±0.06</td>
</tr>
<tr>
<td>thymol</td>
<td><img src="structure4.png" alt="Structure" /></td>
<td>89-83-8</td>
<td>Alfa Aesar</td>
<td>≥ 0.980</td>
<td>0.0006</td>
<td>322.0±0.1</td>
<td>17.4±0.6</td>
</tr>
<tr>
<td>acetonitrile</td>
<td><img src="structure5.png" alt="Structure" /></td>
<td>75-05-8</td>
<td>Fisher</td>
<td>≥ 0.999</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-butanol</td>
<td><img src="structure6.png" alt="Structure" /></td>
<td>71-36-3</td>
<td>Aldrich</td>
<td>≥ 0.995</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ethanol</td>
<td><img src="structure7.png" alt="Structure" /></td>
<td>64-17-5</td>
<td>Fisher</td>
<td>≥ 0.998</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ethyl acetate</td>
<td><img src="structure8.png" alt="Structure" /></td>
<td>141-78-6</td>
<td>Carlo Erba</td>
<td>≥ 0.999</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hexane</td>
<td><img src="structure9.png" alt="Structure" /></td>
<td>110-54-3</td>
<td>Carlo Erba</td>
<td>≥ 0.99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-(−)-limonene</td>
<td><img src="structure10.png" alt="Structure" /></td>
<td>5989-27-5</td>
<td>Aldrich</td>
<td>≥ 0.970</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,2-propanediol</td>
<td><img src="structure11.png" alt="Structure" /></td>
<td>57-55-6</td>
<td>Aldrich</td>
<td>≥ 0.995</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values correspond to the minimum mass fraction purities granted by the supplier. $^b$Water content (mass fraction) measured by Karl–Fisher titration (Metrohm, 831 KF Coulometer). $^c$(−)-Borneol presents a solid–solid transition (II $\rightarrow$ I) at $T_{II \rightarrow I} = 347.9 \pm 0.3$ K; $\Delta H_{II \rightarrow I} = 3.2 \pm 0.3$ kJ·mol$^{-1}$. $^d$ (+)-Camphor presents two low-temperature solid–solid transitions (III $\rightarrow$ II and II $\rightarrow$ I) at $T_{III \rightarrow II} = 244.4 \pm 1.0$ K; $\Delta H_{III \rightarrow II} = 10.3 \pm 3$ kJ·mol$^{-1}$ and $T_{II \rightarrow I} = 369.6 \pm 6$ K; $\Delta H_{II \rightarrow I} = 0.19 \pm 0.07$ kJ·mol$^{-1}$. $^e$Melting properties correspond to the stable $\alpha$ phase of l-menthol. $^{25}$ Biological benefits when incorporated as a feed additive in fish production. $^{22}$ Following our previous studies on water solubilities $^{23}$ and octanol-water partition coefficients $^{24}$ of a diverse set of monoterpenoids, this work aims to investigate the solubility of (−)-borneol, (1R)-(−)-camphor, 1-(−)-menthol, and thymol in seven organic solvents (acetonitrile, 1-butanol, ethanol, ethyl acetate, hexane, R-(−)-limonene, and 1,2-propanediol). For (−)-borneol, (1R)-(−)-camphor, and thymol, solubility measurements were carried out at 298.2 and 313.2 K, while for 1-(−)-menthol with a melting point of 315.6 K, $^{25}$ experiments were only performed at 298.2 K. The solubility of (−)-borneol in water, not examined in our previous works, was also obtained here at 298.2 and 313.2 K. Finally, three well-established thermodynamic models were applied to describe the solubility change with solvent and temperature. The Abraham solvation model $^{26,27}$ was applied to calculate the monoterpenoids solubility in organic solvents at 298.2 K, $^{28-30}$ while the Nonrandom Two-Liquid Segment Activity Coefficient (NRTL-SAC) model, $^{31,32}$ and the Conductor-like Screening Model for Real Solvents (COSMO-RS) $^{33-35}$ were employed to represent the available solubility data in water and organic solvents at different temperatures. $^{30,36,37}$

2. EXPERIMENTAL SECTION

2.1. Chemicals. All of the monoterpenoids and organic solvents used in the solubility experiments are listed in Table 1, along with their chemical structure, CAS number, source, and purity. For the monoterpenoids, melting temperature ($T_m$) and melting enthalpies ($\Delta_m H$) are also provided.

2.2. Solubility Experiments. The solubility experiments were conducted using the isothermal shake-flask method, which was described in detail previously. $^{40}$ In brief, around 50 mL of saturated solutions of monoterpenoid and solvent, containing small amounts of solid in excess, were prepared in all glass flasks covered with aluminum foil to prevent light...
degradation. The flasks were placed on magnetic stirrers (Cimarec I micro stirrer, Thermo Fisher) inside a heated water circulating bath (TC120, Grant) equipped with a precise temperature control system (±0.1 K). Preliminary experiments demonstrated that the equilibrium was achieved after 24 h of continuous stirring, and 16 h of settling was enough for most suspended solid particles to deposit.

At least three samples, between 0.2 and 0.5 cm³, were collected using previously heated glass syringes coupled to nylon or polytetrafluoroethylene filters (0.45 μm pore diameter). After, the samples were analyzed by UV–vis spectroscopy or gas chromatography (GC). The analytical procedures are presented in Section S1 in the Supporting Information (SI). Each reported solubility value is the average of at least three independent readings.

3. MODELING

3.1. Abraham Solvation Model. The Abraham solvation model describes the ratio of the solubilities (in molar concentration basis) of a solute in an organic solvent (S) and in water (S₀) through the following linear free energy relationship (LFER) ²⁷

\[
\log \left( \frac{S}{S_0} \right) = \epsilon + eE + sS + aA + bB + vV
\]

where the uppercase descriptors (E, S, A, B, and V) and the lowercase descriptors (ε, e, s, a, b, and v) represent the Abraham descriptors for the solute and solvent, respectively. Abraham’s research group has already reported the model parameters for several solvents.²⁶,²⁷,⁴¹,⁴² Regarding the solute descriptors, V, the solute’s McGowan characteristic molecular volume, and E, the solute excess molar refractivity, can be directly obtained from the solute chemical structure and refractive index.³⁷,⁴³ The other parameters, S (dipolarity/polarizability), A (overall hydrogen-bond acidity), and B (overall hydrogen-bond basicity), are frequently regressed using a set of experimental solubility and/or partition coefficient data.

3.2. NRTL-SAC Model. The semipredictive NRTL-SAC model has been extensively applied to describe the solubility of biomolecules in water and organic solvents³⁰,⁴⁴–⁵⁰ and was described in detail by Chen and co-authors.³¹,³² In this approach, the activity coefficient of solute i is calculated as the sum of the combinatorial (γ_C)i and the residual (γ_R)i contributions

\[
\ln(\gamma_i) = \ln(\gamma_C^i) + \ln(\gamma_R^i)
\]

In summary, the model characterizes the solute and solvent molecules by four conceptual segments representing different surface interactions: hydrophobicity (X), hydrophobicity (Z), polar-attractive (Y'), and polar-repulsive (Y''). These parameters are available in the literature for many solvents,³¹,³²,⁴⁵,⁵⁰ including those addressed in this work. Therefore, only the solute descriptors are required to estimate the solubilities with the NRTL-SAC, which will be regressed using a small set of the experimental solubility data.

The solubility of a solute (i) in a liquid mixture can be described by

\[
\ln(x_i) = \frac{\Delta_m H_i}{R}\left(\frac{1}{T_{m,i}} - \frac{1}{T}\right) + \frac{\Delta C_{p,i}}{R}\left(\frac{T_{m,i}}{T} - \ln\left(\frac{T_{m,i}}{T}\right) - 1\right) - \ln(\gamma_i)
\]

where \(x_i\) is the mole fraction solubility of solute i at temperature \(T\), \(\Delta_m H_i\) and \(T_{m,i}\) are its melting enthalpy and temperature, respectively, \(\Delta C_{p,i}\) is its heat capacity change upon melting, \(\gamma_i\) is its activity coefficient, and \(R\) is the ideal gas constant. The contribution of the \(\Delta C_{p,i}\) term on the calculation of \(x_i\) is often negligible,³¹ and eq 3 can be simplified to

\[
\ln(x_i) = \frac{\Delta_m H_i}{R}\left(\frac{1}{T_{m,i}} - \frac{1}{T}\right) - \ln(\gamma_i)
\]

Equation 4 is frequently used to describe the solid–liquid equilibrium of a pure solid substance in a liquid mixture when no solid–solid transition between \(T\) and \(T_{m,i}\) occur. Whenever such transitions are present, the modified form of eq 4 is recommended³²,³³

\[
\ln(x_i) = \frac{\Delta_m H_i}{R}\left(\frac{1}{T_{m,i}} - \frac{1}{T}\right) + \frac{\Delta S_S H_i}{R}\left(\frac{1}{T_{m,i}} - \frac{1}{T}\right) - \ln(\gamma_i)
\]

where \(T_{m,i}\) is the transition temperature and \(\Delta S_S H_i\) is its enthalpy.

In this work, eq 5 was applied to describe the solubilities of (−)-borneol and (1R)-(+)-camphor with the NRTL-SAC or COSMO-RS models due to the presence of solid–solid transitions between 298.2 K and melting.³⁸,³⁹ For L-(−)-menthol and thymol, eq 4 was employed. All of the phase transition properties of the solutes considered in the solubility estimations were collected from the open literature²⁵,³⁷–³⁹ and are listed in Table 1.

3.3. COSMO-RS Model. COSMO-RS estimates activity coefficients in liquid mixtures, which can be easily coupled with eqs 4 and 5 to estimate solubilities, by computing pair-wise interaction energies between the molecular surface segments of all mixture components. The screened charges of these segments, which are encoded in the so-called sigma surfaces and profiles, are computed using DFT with the COSMO solvation model.

COSMO-RS was employed in this work through its implementation in the COSMOTHERM software package (version 21.0),³⁵ with the BP-TZVPD_FINE_21.cgd parametrization. This parametrization requires sigma surfaces optimized using the de2-TZVPD basis set, the BP-86 DFT functional, and the COSMO solvation model with infinite permittivity. When available, these were collected from the COSMOTHERM Tzvdp-Fine database. Those sigma profiles not available in the database (monoterpenoids and some solvents) were computed using the TmolX (version 4.5) software package,³⁶ coupled with COSMOconfX 2021 for initial conformer generation (BP-TZVPD-FINE-COSMO + GAS_18 template). All conformers obtained were considered to perform the calculations in COSMOTHERM. More details are available in the COSMOTHERM reference manual.³⁶

3.4. Assessment of the Deviations. The deviations between the experimental and calculated solubilities were assessed by the average relative deviation (ARD)
ARD (%) = \frac{1}{n} \sum \left( \frac{x_{j,\text{exp}} - x_{j,\text{calc}}}{x_{j,\text{exp}}} \right) \times 100 \tag{6}

Table 2. Experimental Solubilities (in Mole Fraction) of the Monoterpenoids in Organic Solvents at 298.2 and 313.2 K$^a,b$

<table>
<thead>
<tr>
<th>solvent</th>
<th>298.2 K</th>
<th>313.2 K</th>
<th>298.2 K</th>
<th>313.2 K</th>
<th>298.2 K</th>
<th>298.2 K</th>
<th>313.2 K</th>
</tr>
</thead>
<tbody>
<tr>
<td>ideal$^c$</td>
<td>0.272</td>
<td>0.333</td>
<td>0.422</td>
<td>0.478</td>
<td>0.740</td>
<td>0.595</td>
<td>0.832</td>
</tr>
<tr>
<td>acetonitrile</td>
<td>0.0451 ± 0.0004</td>
<td>0.0781 ± 0.0002</td>
<td>0.5872 ± 0.0008</td>
<td>0.7212 ± 0.0002</td>
<td>0.4047 ± 0.0019</td>
<td>0.7059 ± 0.0017</td>
<td>0.9001 ± 0.0004</td>
</tr>
<tr>
<td>1-butanol</td>
<td>0.3121 ± 0.0010</td>
<td>0.3480 ± 0.0020</td>
<td>0.4541 ± 0.0020</td>
<td>0.5347 ± 0.0025</td>
<td>0.7103 ± 0.0029</td>
<td>0.6663 ± 0.0045</td>
<td>0.8298 ± 0.0018</td>
</tr>
<tr>
<td>ethanol</td>
<td>0.2977 ± 0.0007</td>
<td>0.3256 ± 0.0010</td>
<td>0.3795 ± 0.0006</td>
<td>0.5215 ± 0.0031</td>
<td>0.6510 ± 0.0022</td>
<td>0.6586 ± 0.0016</td>
<td>0.8435 ± 0.0018</td>
</tr>
<tr>
<td>ethyl acetate</td>
<td>0.2365 ± 0.0013</td>
<td>0.2956 ± 0.0008</td>
<td>0.5686 ± 0.0023</td>
<td>0.7035 ± 0.0037</td>
<td>0.6805 ± 0.0041</td>
<td>0.7090 ± 0.0052</td>
<td>0.8713 ± 0.0033</td>
</tr>
<tr>
<td>hexane</td>
<td>0.0887 ± 0.0012</td>
<td>0.1426 ± 0.0012</td>
<td>0.5237 ± 0.0016</td>
<td>0.5477 ± 0.0057</td>
<td>0.6927 ± 0.0056</td>
<td>0.3380 ± 0.0029</td>
<td>0.8203 ± 0.0044</td>
</tr>
<tr>
<td>R-(+)−limonene</td>
<td>0.1443 ± 0.0012</td>
<td>0.2263 ± 0.0024</td>
<td>0.6278 ± 0.0025</td>
<td>0.6790 ± 0.0016</td>
<td>0.6906 ± 0.0043</td>
<td>0.4797 ± 0.0034</td>
<td>0.8071 ± 0.0027</td>
</tr>
<tr>
<td>1,2-propanediol</td>
<td>0.1236 ± 0.0008</td>
<td>0.1568 ± 0.0014</td>
<td>0.0894 ± 0.0022</td>
<td>0.1163 ± 0.0003</td>
<td>0.6179 ± 0.0012</td>
<td>0.6959 ± 0.0029</td>
<td>0.8456 ± 0.0038</td>
</tr>
</tbody>
</table>

$^a$Temperature and pressure standard uncertainties are $u(T) = 0.10$ K and $u_p = 0.05$, respectively. $^b$Standard deviations are placed after the ± sign. $^c$Calculated for borneol and camphor using the phase transition properties presented in Table 1.

Figure 1. Comparison of the experimental solubility data of thymol obtained in this work and available in the literature, in different organic solvents: (a) acetonitrile; (b) 1-butanol; (c) ethanol; (d) R-(+)−limonene. Experimental data: (green circle solid) this work; (Δ) Zhu et al.; (□) Bermejo et al.\textsuperscript{36}

4. RESULTS AND DISCUSSION

4.1. Solubility Measurements. The solubilities of (−)-borneol, (1R)-(+)-camphor, l-(−)-menthol, and thymol in the selected organic solvents (acetonitrile, 1-butanol, ethanol, ethyl acetate, hexane, R-(+)−limonene, and 1,2-propanediol) are presented in Table 2 along with the ideal solubility values. The aqueous solubilities of the four monoterpenoids, and the experimental activity coefficients, calculated from the mole fraction solubilities listed in Table 2 using eq 5, are presented in Tables S1 and S2 in the SI, respectively. At this point, it is critical to observe the impact of solid−solid transitions on the ideal solubility and, therefore, on the activity coefficients. While for (1R)-(+)-camphor the impact is almost null, which is consistent with the very low transition enthalpy, for (−)-borneol, at 298.2 K, the ideal solubility would be 0.327, 20.2% higher than that presented in Table 2 amplifying, generally, the deviations from ideality.

The good consistency of the solubility values is demonstrated by the low coefficient of variation, lower than 2.5% for organic systems and 3.7% for the (−)-borneol/water system. Excepting (1R)-(+)-camphor/water system, where the observed solubility at 313.2 K is 3 times lower than the value obtained at 298.2 K,\textsuperscript{24} in all other systems studied at both temperatures, the solubility increases with the temperature. In general, the solubilities of the monoterpenoids in organic solvents are much higher than in water, with
differences varying from 2 orders of magnitudes for (−)-borneol to 4 orders of magnitude for L-(−)-menthol. The ideal mole fraction solubility values listed in Table 2 (higher than 0.27) are much closer to the solubilities in organic solvents, showing better affinity with the solutes. Indeed, due to the hydrophobic nature of monoterpenoids, the activity coefficients of all solutes in water were much higher than 1, ranging from $2.7 \times 10^8$ for (−)-borneol at 298.2 K to $1.7 \times 10^5$ of L-(−)-menthol at 298.2 K.

Concerning the monoterpane alcohols, L-(−)-menthol is more soluble in all of the organic solvents than (−)-borneol, at 298.2 K. This behavior is in line with the ideal solubility trend ($\chi_{\text{ideal}}$). Moreover, they exhibit positive deviations from ideality for most studied binary mixtures, except for (−)-borneol in ethanol (at 298.2 K) or 1-butanol. Among the organic solvents, the highest and the lowest solubilities of both solutes were observed for 1-butanol and acetonitrile, respectively.

Excepting for the solubilities of (1R)-(+) -camphor in 1,2-propanediol and ethanol (at 298.2 K), the solubilities are higher than the ideal value, and consequently, the activity coefficients are lower than 1. The highest solubilities occur in R-(+)-limonene, ethyl acetate, and acetonitrile, whereas 1,2-propanediol is the organic solvent with the poorest affinity to this monoterpane ketone. Taking into account the Hansen solubility parameters, it can be said that solvents of intermediate polarity and small hydrogen-bond character favor camphor solubilization. Different from the other studied solutes, which present one HBA and one HBD site, (1R)-(+) -camphor has only one HB acceptor site (from the carbonyl group). Therefore, solvents with high Hansen HB parameters, such as 1,2-propanediol, might not be as appropriate for solubilizing this monoterpane ketone.

Regarding the phenolic thymol, the highest solubility values are observed in acetonitrile, ethyl acetate, and 1,2-propanediol. In contrast, the lowest values among the organics are registered in the nonpolar hexane and R-(+)-limonene. Similar to L-(−)-menthol, the solubilities of thymol in the organic solvents are quite high, probably related to their low melting temperature and moderate melting enthalpies, and the favorable interactions in the solution. However, although thymol and menthol are structurally very similar, the electron coefficients are lower than 1. The highest solubilities occur in acetonitrile, ethyl acetate, and 1,2-propanediol (at 298.2 K). This behavior is in line with the ideal solubility trend ($\chi_{\text{ideal}}$).

For borneol and camphor, a literature search showed that temperature-dependent solubility curves were only available for their racemic mixtures in ethanol. Other data are available, though the stereoisomer of the solute was not specified. As discussed by Coquereel, the solubility behavior of a specific enantiomer or a racemic mixture in a given solvent system might be quite distinct, and a direct comparison of the available solubility data obtained in this work and those reported by Chen et al. is not straightforward. Nonetheless, in those cases, the mole fraction water solubility obtained in this work for (−)-borneol (9.9 $\times 10^{-5}$) is consistent with the literature average and the values presented in Table S3 (8.33 $\times 10^{-5}$). Similarly, moderate relative deviations are observed between the solubilities of camphor measured in this work and those reported by Lin and Nash (at 298.2 K) in ethyl acetate (deviation = 7.5%), hexane (deviation = 18.2%), and 1,2-propanediol (deviation = 11.5%).

### 4.3. Abraham Solvation Model

For all of the studied solutes, the model descriptors ($E$, $S$, $A$, $B$, and $V$) have been previously reported and are presented in Table S4 in the SI. Nevertheless, poor descriptions of the experimental data obtained in this work were found using the descriptors available in the literature, with ARD often superior to 40% (Table S4). The isomeric form of the solute was not specified, and the available descriptors were probably regressed from partition coefficient data only since most solubility data available in the organic solvents were reported recently.

Despite the recommendations in terms of the solute concentration range adequate for this model, the descriptors were regressed in this work for the solutes using a similar strategy applied in a previous work of our group.

Since the Abraham solvation model was developed based on molar solubilities, the experimental mole fraction solubilities ($\chi^{0}$) were converted to the molar basis ($S^{0}$) using the expression

$$
S^{0} = \frac{\chi^{0}}{x^{0}}
$$

### Table 3. Estimated Abraham Descriptors for the Monoterpenoids, Outlier Solvent, and ARD (%) Obtained Using Solubilities in Water and Five Pure Organic Solvents in the Correlation Set

<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>(−)-borneol</td>
<td>0.688</td>
<td>0.589</td>
<td>0.104</td>
<td>0.672</td>
<td>1.359</td>
<td>1,2-propanediol</td>
<td>12</td>
</tr>
<tr>
<td>(1R)-(+) -camphor</td>
<td>0.570</td>
<td>0.654</td>
<td>0</td>
<td>0.583</td>
<td>1.316</td>
<td>acetonitrile</td>
<td>26</td>
</tr>
<tr>
<td>L-(−)-menthol</td>
<td>0.366</td>
<td>0.474</td>
<td>0</td>
<td>0.620</td>
<td>1.468</td>
<td>1,2-propanediol</td>
<td>36</td>
</tr>
<tr>
<td>thymol</td>
<td>0.824</td>
<td>0.647</td>
<td>0.031</td>
<td>0.600</td>
<td>1.339</td>
<td>1,2-propanediol</td>
<td>33</td>
</tr>
</tbody>
</table>

The solubilities of thymol in acetonitrile, 1-butanol, ethanol, and limonene are compared in Figure 1. To the best of our knowledge, the solubilities of L-(−)-menthol in the studied organic solvents are reported here for the first time. As can be seen in Figure 1, the solubilities of thymol in acetonitrile, 1-butanol, and ethanol from this work are in very good agreement with the data reported by Zhu et al. For these systems, an average relative deviation (calculated as the absolute difference between the solubility data obtained in this work and the literature value, divided by the literature solubility value) of 4.4% is found. For the thymol/ethanol mixture, the solubility curve reported by Bermejo et al. is considerably different from the remaining data. Regarding the thymol/R-(+)-limonene, the value obtained in this work at 313.2 K is 14% higher than the value found by Bermejo and co-authors.

For borneol and camphor, a literature search showed that temperature-dependent solubility curves were only available for their racemic mixtures in ethanol. Other data are available, though the stereoisomer of the solute was not specified. As discussed by Coquereel, the solubility behavior of a specific enantiomer or a racemic mixture in a given solvent system might be quite distinct, and a direct comparison of the available solubility data obtained in this work and those reported by Chen et al. is not straightforward. Nonetheless, in those cases, the mole fraction water solubility obtained in this work for (−)-borneol (9.9 $\times 10^{-5}$) is consistent with the literature average and the values presented in Table S3 (8.33 $\times 10^{-5}$). Similarly, moderate relative deviations are observed between the solubilities of camphor measured in this work and those reported by Lin and Nash (at 298.2 K) in ethyl acetate (deviation = 7.5%), hexane (deviation = 18.2%), and 1,2-propanediol (deviation = 11.5%).
Figure 2. Comparison between the experimental and calculated solubility data by the Abraham solvation model, at 298.2 K: (a) correlation; (b) prediction. Filled symbols correspond to: ○, (−)-borneol; □, (1R)-(−)-camphor; △, 1-(−)-menthol; and ◊, thymol.

Table 4. Estimated NRTL-SAC Molecular Parameters, Outlier Solvent, and ARD (%) for the Studied Monoterpenoids

<table>
<thead>
<tr>
<th>compound</th>
<th>$X$</th>
<th>$Y^+$</th>
<th>$Y^−$</th>
<th>$Z$</th>
<th>outlier</th>
<th>ARD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(−)-borneol</td>
<td>0.734</td>
<td>0</td>
<td>0</td>
<td>0.374</td>
<td>acetonitrile</td>
<td>19</td>
</tr>
<tr>
<td>(1R)-(−)-camphor</td>
<td>1.324</td>
<td>2.223</td>
<td>0.189</td>
<td>0.089</td>
<td>water</td>
<td>27</td>
</tr>
<tr>
<td>1-(−)-menthol</td>
<td>0.844</td>
<td>0.216</td>
<td>0</td>
<td>0.054</td>
<td>ethanol</td>
<td>7</td>
</tr>
<tr>
<td>thymol</td>
<td>0.748</td>
<td>0.757</td>
<td>0.832</td>
<td></td>
<td>hexane</td>
<td>8</td>
</tr>
</tbody>
</table>

$$S_i^{\text{exp}} \approx \frac{x_i^{\text{exp}}V_{\text{solute}}}{[x_i^{\text{exp}}V_{\text{solute}} + (1 - x_i^{\text{exp}})V_{\text{solvent}}]} \quad (7)$$

where $V$ is the molar volume (dm$^3$·mol$^{-1}$), calculated from density data collected from the literature.

The descriptors $E$ and $V$ were calculated using the procedure illustrated by Abraham and co-authors, while the other remaining solute parameters ($S$, $A$, and $B$) were obtained by simultaneously solving a reduced set of LFERs (eq 1) using the solubility data measured by us, at 298.2 K, in water and five organic solvents (acetonitrile, ethanol, ethyl acetate, hexane, and 1,2-propanediol). For (1R)-(−)-camphor, the hydrogen-bond acidity parameter ($A$) was set to 0 since no HB-donor site is available in its chemical structure. The estimated Abraham solute descriptors are presented in Table 3, along with the obtained ARDs and the outlier systems.

A global ARD of 27% was obtained in the estimation of the model descriptors, which is comparable to the results found in our previous work. The best description was obtained for (−)-borneol (ARD = 12%), while the biggest deviations are observed for 1-(−)-menthol (ARD = 36%) and thymol (ARD = 33%). The latter solutes generally present higher solubilities in the studied organic solvents than the other two monoterpenoids (with only a few exceptions), while values registered for (−)-borneol are typically the lowest. This confirms that the Abraham model is more suitable for representing moderate to low solubility data values.

In most cases, the estimated E descriptors are close to those found in the literature, but larger differences can be observed for $S$, $A$, and $B$, particularly for thymol, where differences around 0.2–0.4 units are observed. For the other solutes, the model descriptors obtained in this work usually deviate less than 0.2 units from the values listed in Table S4. It is relevant to mention that the descriptors reported by Abraham and co-authors were estimated using reduced sets of experimental partition coefficients data, which might not entirely cover the solvation behavior of the saturated conditions addressed here. On the other hand, the methodology is highly dependent on the available solubility values in water, which were updated with new information in this work.

The parameters presented in Table 3 were used to estimate the solubility data, at 298.2 K, of the stereoisomers studied in this work, in solvents for which the Abraham descriptors are also available (solubility of all solutes in 1-butanol measured in this work and of thymol in other pure solvents). A complete overview of the correlation and predicted results is presented in Figure 2.

The highest relative deviations in the correlation step were found in 1,2-propanediol for 1-(−)-menthol and thymol, but in acetonitrile, the results are also poor. Nonetheless, the model usually delivers very good solubility estimations in ethanol, ethyl acetate, and hexane. Moreover, as shown in Figure 2b, good prediction results are offered by the model with global ARDs of 8%.

4.4. NRTL-SAC. To the best of our knowledge, the NRTL-SAC conceptual segments for the studied monoterpenoids were not reported yet. The optimization of the solute parameters ($X$, $Y^+$, $Y^−$, $Z$) was carried out using the nonlinear least-squares method. The objective function to minimize was defined as the global ARD (eq 6).

From the previous experience in our group, a reduced set of experimental solubility data obtained in structurally different solvent systems is an adequate correlation database. Thus, the experimental solubilities in five solvents (water, acetonitrile, ethanol, ethyl acetate, and hexane), at 298.2 and 313.2 K, were used. After the predictive capability of the model was also checked by estimating the solubilities in 1-butanol, R-(−)-limonene, and in other solvents found in the literature. Again, the solubility data of the racemic mixtures or of stereoisomers that were not identified, were not included in...
the predicted set due to the solubility dependency on the solute’s phase transition properties (eqs 4 and 5).

The estimated NRTL-SAC parameters, the global ARD, and the outlier solvent for each solute are summarized in Table 4, while a visual description of the correlation and prediction results is given in Figure 3. For a better comparison, the solubilities in water were excluded from Figure 3 and are presented separately in Figure S1 in the SI. The solvent’s molecular parameters were retrieved from the literature.32,45,50

A very good description of the available solubility data for the monoterpenoids is obtained using the NRTL-SAC model, with global ARDs obtained in the correlation (15%) and prediction (14%). Excellent results were obtained for L(-)-menthol and thymol (ARD lower than 8%) during correlation, and even lower for the predictions, where ARD is close to 4% for both solutes. In the case of (-)-borneol, ARDs of 19 and 26% were obtained in the correlation and prediction, respectively.

Among the four monoterpenoids, the overall weakest performance is for (1R)-(+)-camphor, with an ARD of 27 and 21% for correlation and prediction, respectively. Although these deviations are superior to those obtained for all other systems, they are in line with or better than those reported in previous works for other biomolecules.40,46,49,66,67 Particulatly for water, the NRTL-SAC is not capable of predicting the reduction in the solubility of camphor as the temperature rises from 298.2 to 313.2 K (shown in Table 2), providing a considerably lower solubility estimate ($x_{\text{calc}} = 3.9 \times 10^{-5}$) than the experimental value ($x_{\text{exp}} = 1.4 \times 10^{-4}$ at 298.2 K.24 Notably, for such small solubility values, the model delivers very reliable water solubility estimations for all other monoterpenoids, as illustrated in Figure S1, with a global ARD of 11% considering the four solutes.

4.5. COSMO-RS. In Figure 4, the experimental solubilities of the monoterpenoids in organic solvents compiled in this work30,36,37 are compared with the values obtained with COSMO-RS. The COSMO-RS model delivers a very good representation of the solubility data of the monoterpenoids in organic solvents, showing a global ARD of 16%. High-quality predictions were obtained for thymol (ARD = 7%), (-)-borneol (ARD = 8%), and L(-)-menthol (ARD = 13%). For these solutes, worse results were observed for L(-)-menthol + acetonitrile and thymol/(-)-borneol + hexane.

In particular for thymol, a remarkably low ARD of 2.5% is observed for the available solubility data at 313.2 K, showing that COSMO-RS can deliver reliable predictions considering the solubility change with temperature.

Concerning the solubility of (1R)-(+)-camphor in organic solvents, the calculated global ARD is 37%. In Figure 4, the systematic underestimation of (1R)-(+)-camphor solubility for most solvents is notable, which is very significant in organic solvents of different characteristics, namely, acetonitrile, hexane, limonene, and ethyl acetate. Nonetheless, the (1R)-(+)-camphor predicted solubilities in alcohols are satisfactory (ARD < 25%).

The experimental and predicted solubilities in water are depicted in Figure S2 in the SI. For these systems, the best solubility description is achieved for L(-)-menthol, followed by thymol, (-)-borneol, and (1R)-(+)-camphor. The values are now overestimated, and excepting (1R)-(+)-camphor, the model provides water solubilities of the same order of magnitude as the experimental ones. As discussed in our
previous work,\textsuperscript{24} the calculation of monoterpenoids solubility in aqueous systems, using either semiempirical or predictive models, is a challenging task. Nevertheless, the predictive COSMO-RS often provides better performances than other more empirical options.

Since appropriate descriptions of the solubility data at 298.2 and 313.2 K were obtained with both NRTL-SAC and COSMO-RS models, they were tested to represent the temperature solubility curves of thymol using data available in the literature\textsuperscript{30,36,37} and measured in this work. The temperature solubility curves are depicted in Figure S3 in the SI, while a comparison of the experimental and calculated solubilities is presented in Figure 5.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Comparison of the temperature-dependent experimental and calculated solubilities of thymol by the COSMO-RS model (filled symbols) and the NRTL-SAC model (empty symbols). The dashed lines correspond to $x = 0.15$.}
\end{figure}

In general, both models offer a qualitative picture of the solubility curves, with NRTL-SAC (ARD = 9%) performing better than COSMO-RS (ARD = 17%). Two lines representing ARD limits of 15% were included in Figure 5. As can be seen, much of the values are inside that region, being the alcohols 1-propanol, 2-propanol, and 1-butanol the outliers for COSMO-RS, while acetonitrile is the solvent presenting most of the data points out of the delimited region. When comparing both models, the NRTL-SAC depends on the availability of the model parameters for the solute and the solvent, limiting its application for systems where this information is known. On the other hand, COSMO-RS presents a full predictive character, capable of providing solubility curves only from the chemical structures of the individual compounds in a mixture. Therefore, the ability of the COSMO-RS model to deliver solubility estimations suggests its application in the preliminary solvent screening of different solutes, particularly when little experimental information is available.

\section*{5. CONCLUSIONS}

New experimental solubility data of three monoterpenoids ((−)-borneol, (1R)-(−)-camphor, and thymol) in acetonitrile, 1-butanol, ethanol, ethyl acetate, hexane, R-(−)-limonene, and 1,2-propanediol, at 298.2 and 313.2 K, are reported. For (−)-borneol, experimental water solubilities at 298.2 and 312.2 K were also measured. The solubilities of 1-(−)-menthol in the same set of organic solvents were investigated at 298.2 K. The Abraham solvation and NRTL-SAC models were successfully applied to describe the solubility data measured here, and most data were retrieved from the open literature. In the case of the Abraham model, global ARD of 27 and 8% were obtained in the correlation and prediction steps, respectively, while the lowest global deviations were observed for camphor and borneol. Regarding the NRTL-SAC model, the ARD for the solutes varied between 7 and 27% in the correlation and between 4 and 26% in the prediction, considering the available solubility at 298.2 and 313.2 K, being the best performances found for thymol and 1-(−)-menthol (ARDs = 6%). Nevertheless, the use of these models depends on the availability of the solute parameters or a small set of experimental data to estimate them. Moreover, the absence of parameters for some solvents (1,2-propanediol for the NRTL-SAC model, and ethyl lactate and limonene for the Abraham solvation model) precludes covering all of the solubility data.

On the other hand, COSMO-RS offers a good representation of the solubility data in organic solvents at 298.2 and 313.2 K, presenting ARDs between 7% (thymol) and 37% ((1R)-(−)-camphor), and a global ARD of 16%. The overall performance of COSMO-RS in describing the solubilities of the monoterpenoids was comparable to those registered by other hybrid approaches, which is a remarkable achievement of this fully predictive model that requires only structural information from the components present in the mixture.

\section*{ASSOCIATED CONTENT}

\subsection*{Supporting Information}

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.iecr.2c03991.

Analytical procedure for quantification of experimental solubility data by UV–vis spectroscopy and gas chromatography; experimental water solubilities of the solutes at 298.2 and 313.2 K; experimental activity coefficients for the solutes in water and organic solvents at 298.2 and 313.2 K; overview of the available solubility data for the binary systems under study; list of the Abraham descriptors for the solutes found in the literature; and comparison of the experimental temperature-dependent solubility curves with those predicted by the NRTL-SAC and COSMO-RS models (PDF).

\section*{AUTHOR INFORMATION}

Corresponding Authors

Olga Ferreira – Centro de Investigação de Montanha (CIMO), Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal; Labaratório para a Sustentabilidade e Tecnologia em Regiões de Montanha, Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal; orcid.org/0000-0001-8414-3479; Phone: +351273303087; Email: oferreira@ipb.pt

Simão P. Pinho – Centro de Investigação de Montanha (CIMO), Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal; Labaratório para a Sustentabilidade e Tecnologia em Regiões de Montanha, Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal; orcid.org/0000-0002-9211-857X; Phone: +351273303086; Email: spinho@ipb.pt
Authors
Sérgio M. Vilas-Boas – Centro de Investigação de Montanha (CIMO), Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal; Laboratório para a Sustentabilidade e Tecnologia em Regiões de Montanha, Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal; CICECO – Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, 3810-193 Aveiro, Portugal; orcid.org/0000-0001-8179-935X
Isabella W. Cordova – Centro de Investigação de Montanha (CIMO), Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal; Laboratório para a Sustentabilidade e Tecnologia em Regiões de Montanha, Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal; CICECO – Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, 3810-193 Aveiro, Portugal; orcid.org/0000-0001-8179-935X
Dinis O. Abranches – CICECO – Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, 3810-193 Aveiro, Portugal
João A. P. Coutinho – CICECO – Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, 3810-193 Aveiro, Portugal; orcid.org/0000-0002-3841-743X
Complete contact information is available at: https://pubs.acs.org/10.1021/acs.iecr.2c03991

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS
This work was developed within the scope of the project CIMO-Mountain Research Center, UIDB/00690/2020 and LA/P/0007/2020, and CICECO-Aveiro Institute of Materials, UIDB/50011/2020, UIDP/50011/2020, and LA/P/0006/2020, financed by national funds through the Portuguese Foundation for Science and Technology (FCT)/MCTES. Sérgio M. Vilas-Boas (SRFH/BD/138149/2018 and COVID/BD/152936/2022) and Isabella W. Cordova (2022.12407.BD) thank FCT and the European Social Fund (ESF) for their Ph.D. grants.

REFERENCES
(3) Sharmean, J. B.; Mahomooodally, F. M.; Zengin, G.; Maggi, F. Essential Oils as Natural Sources of Fragrance Compounds for Cosmetics and Cosmeceuticals. Molecules 2021, 26, 666.