Solubility Enhancement of Hydrophobic Substances in Water/Cycrene Mixtures: A Computational Study

Dinis O. Abranches, Jordana Benfica, Seishi Shimizu, and Joao A.P. Coutinho

Ind. Eng. Chem. Res., Just Accepted Manuscript • DOI: 10.1021/acs.iecr.0c03155 • Publication Date (Web): 17 Sep 2020

Downloaded from pubs.acs.org on September 21, 2020
Solubility Enhancement of Hydrophobic Substances in Water/Cyrene Mixtures: A Computational Study

Dinis O. Abranches¹, Jordana Benfica¹, Seishi Shimizu² and João A. P. Coutinho¹,*

¹ CICECO – Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, 3810-193 Aveiro, Portugal
² York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5DD, United Kingdom
* Corresponding Author Email: jcoutinho@ua.pt

Abstract

Cyrene, a recently introduced bio-based green solvent, behaves as a hydrotrope, greatly increasing the aqueous solubility of hydrophobic substances in water. Moreover, water, when present in small quantities, has also been shown to increase the solubility of hydrophobic substances in Cyrene. Albeit of large practical importance, the understanding of the solvation ability of water/Cyrene mixtures is hampered by a lack of fundamental knowledge of their solvation mechanisms, which is further complicated by a chemical equilibrium established between the ketone form of Cyrene and water, yielding the diol form of Cyrene. In this work, COSMO-RS is used to rationalize the co-solvency effect of water on Cyrene.

It is shown here that COSMO-RS can accurately predict the solubility of hydrophobic substances in water/Cyrene mixtures, in the entire composition range from pure water to pure Cyrene. The predictions of the model show that the ketone form of Cyrene controls its hydrotropic ability, supporting the concept of cooperative hydrotropy. Using COSMO-RS to assess molecular contacts probability, it is also shown that water, when added in small quantities to Cyrene, acts as a co-solvent by preferentially interacting with the solute as a hydrogen bond acceptor, effectively doubling the hydrogen bond donning ability of the solute and, thus, maximizing the interactions with Cyrene. In turn, this leads to a solubility increase of hydrophobic substances in wet Cyrene.

Keywords: Dihydrolevoglucosenone; Cooperativity; Green Chemistry; Solid-Liquid Equilibrium; Solubility; Hydrophobic Compounds
1. Introduction

In 2014, Sherwood et al.\textsuperscript{1} proposed dihydrolevoglucosenone, commercially known as Cyrene, as a bio-based, highly biodegradable solvent, derived directly from cellulose. Cyrene is a polar aprotic solvent, whose properties render it an attractive substitute to industrially relevant solvents such as N-methyl-2pyrrolidine (NMP). Owing to its sustainable character, Cyrene has sparked a great deal of interest in the scientific community.\textsuperscript{2–11} In fact, and rather surprisingly considering its novelty, a demonstration plant able to produce 50 tons per year of Cyrene was opened by the Circa group in early 2019,\textsuperscript{12} and plans to build a plant to produce 1000 tons of Cyrene per year in France were announced in early 2020.\textsuperscript{13}

Mixing with water may greatly expand the potential of Cyrene as a green solvent. De bruyn et al.\textsuperscript{11} showed that Cyrene behaves as a hydrotrope when added to water, being able to greatly increase the aqueous solubility of hydrophobic solutes. Surprisingly, the reverse is also true: water, when added in small quantities to Cyrene, also increases the solubility of hydrophobic solutes, such as salicylic acid, aspirin, ibuprofen, and caffeine, that are of pharmacological relevance. From an application perspective, using water/Cyrene mixtures instead of pure Cyrene may be advantageous when relying on vapor-liquid equilibrium-based separation processes due to i) its decomposition at a temperature below its atmospheric boiling temperature, and ii) its relatively high enthalpy of vaporization at 25 °C (57.65 kJ/mol).\textsuperscript{14}

The rationalization of the solvation ability of water/Cyrene mixtures is difficult because of its reaction with water: the ketone (pure) form of Cyrene establishes a chemical equilibrium with water to form its geminal diol (Figure 1). Using the concept of cooperative hydrotropy, where the solubility increase of hydrophobic solutes occurs due to water-mediated aggregation of hydrotrope molecules around solute molecules,\textsuperscript{15,16} Abranches et al.\textsuperscript{2} showed that the geminal diol form of Cyrene acted as poor hydrotrope at low Cyrene concentration, and that the ketone form of Cyrene acted as an efficient hydrotrope at high Cyrene concentrations. Despite this, the molecular mechanism behind the solubility maximum at high Cyrene concentrations (or, in other words, the ability of water to increase the solubility of hydrophobic solutes in Cyrene) is still unexplained.
Figure 1. Chemical structures of the ketone (left) and diol (right) forms of Cyrene.

In this work, COSMO-RS, a quantum chemistry-based excess Gibbs energy model, is used to evaluate the impact of each individual form of Cyrene on the solubility of hydrophobic solutes in ketone/water, diol/water, and ketone/diol/water systems, and to probe the specific intermolecular interactions responsible for the water induced enhanced solubility of hydrophobic solutes in Cyrene, through the assessment of the contacts probability of the various compounds.
2. Computational Details

2.1 Solid-Liquid Equilibrium

The solubility of a compound in a liquid solvent is described by the following solid-liquid equilibrium expression: 17,18

\[
\ln (x_S \gamma_S) = \frac{\Delta_m h}{R \left( \frac{T_m}{T_m} - 1 \right)} + \frac{\Delta_m C_P}{R \left( \frac{T_m}{T} - \ln \frac{T_m}{T} - 1 \right)} \tag{1}
\]

where \(x_S\) is the saturation mole fraction (solubility) of the solute in the system, \(\gamma_S\) is its activity coefficient, \(T_m\) is its melting temperature, \(\Delta_m h\) is its enthalpy of fusion (at temperature \(T_m\)), \(\Delta_m C_P\) is its heat capacity change upon melting, \(R\) is the ideal gas constant and \(T\) is the absolute temperature of the system. Since the influence of the \(\Delta_m C_P\) term in Equation 1 is relatively small, and considering that its value is not easily obtained experimentally and is seldom available in the literature, Equation 1 is often simplified as: 17–19

\[
\ln (x_S \gamma_S) = \frac{\Delta_m h}{R \left( \frac{1}{T_m} - \frac{1}{T} \right)} \tag{2}
\]

In the case where the melting properties of the solute are known \((T_m \text{ and } \Delta_m h)\), Equation 2 can be used to predict solubilities if coupled with a Gibbs excess free energy model capable of predicting the activity coefficient of the solute in the liquid mixture. In this work, COSMO-RS is used to predict \(\gamma_S\), and to estimate the solubility of hydrophobic solutes in water, Cyrene, and Cyrene-water mixtures, through Equation 2. Note that \(\gamma_S\) depends on the composition of the final system and its absolute temperature. Temperature is a known variable; however, the final composition of the system depends on the solubility of the solute \((x_S)\) which, in turn, depends on its activity coefficient. Thus, \(x_S\) is iteratively calculated using Equation 2, where in each iteration COSMO-RS is used to calculate \(\gamma_S\). The melting properties of the solutes studied in this work are listed in Table 1.
Table 1. Melting temperature and corresponding enthalpy of fusion for the solutes studied in this work.

<table>
<thead>
<tr>
<th>Substance</th>
<th>CAS Number</th>
<th>$T_m/K$</th>
<th>$\Delta_m h/\text{kJ}\cdot\text{mol}^{-1}$</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin</td>
<td>50-78-2</td>
<td>412.7</td>
<td>31.01</td>
<td>20</td>
</tr>
<tr>
<td>Caffeine</td>
<td>58-08-2</td>
<td>507.7</td>
<td>24.8</td>
<td>21</td>
</tr>
<tr>
<td>Ferulic Acid</td>
<td>1135-24-6</td>
<td>444.6</td>
<td>33.338</td>
<td>22</td>
</tr>
<tr>
<td>Ibuprofen (Racemic)</td>
<td>15687-27-1</td>
<td>350.4</td>
<td>39.5</td>
<td>23</td>
</tr>
<tr>
<td>Phthalic Acid</td>
<td>88-99-3</td>
<td>463.5</td>
<td>36.5</td>
<td>24</td>
</tr>
<tr>
<td>Salicylic Acid</td>
<td>69-72-7</td>
<td>434.1</td>
<td>27.1</td>
<td>25</td>
</tr>
<tr>
<td>Syringic Acid</td>
<td>530-57-4</td>
<td>480.3</td>
<td>33.7</td>
<td>26</td>
</tr>
<tr>
<td>Vanillin</td>
<td>121-33-5</td>
<td>355.4</td>
<td>22.4</td>
<td>27</td>
</tr>
<tr>
<td>Benzoic Acid</td>
<td>65-85-0</td>
<td>396.9</td>
<td>16.99</td>
<td>28</td>
</tr>
</tbody>
</table>

2.2 COSMO-RS

COSMO-RS (Conductor-like Screening Model for Real Solvents) is a quantum chemistry-based thermodynamics model able to predict the chemical potential of compounds in liquid mixtures.²⁹⁻³¹ Within the framework of COSMO-RS, the geometry and charge density surface of each molecule needs to be optimized in advance using DFT and the COSMO solvation model. From this, the model constructs a histogram, the so-called $\sigma$-profile, which describes the amount of area segment of charge-density $\sigma$ that a molecule possesses. Finally, the model relies on pair-wise interactions between all area segments of all molecules present in the system in order to calculate their chemical potentials and, thus, their activity coefficients.

Each molecule studied in this work (the solutes listed in Table 1, water, and the ketone and diol forms of Cyrene) was optimized using TURBOMOLE, through its interface TmoleX,³² by adopting the COSMO-BP-TZVP template available in the software package. This template uses a def-TZVP basis set for all atoms, a DFT with the B-P86 functional and the COSMO solvation model (continuum with infinite permittivity). The molecule file obtained was used as input for the software COSMOtherm,³³,³⁴ which implements the model COSMO-RS using the BP_TZVP_19.ctd parametrization. All molecules were also optimized through the COSMO-BP-TZVPD-FINE template of TmoleX,³² which uses a def2-TZVPD basis set for all atoms, a DFT with the B-P86 functional and the COSMO solvation model (continuum with infinite permittivity and a fine grid cavity size). This was carried out to perform a comparison between the BP_TZVP_19.ctd and BP_TZVPD_FINE_19.ctd parametrizations of COSMOtherm.
3. Results and Discussion

3.1 Solubility Prediction

The solubility of the compounds studied here (Table 1) are initially predicted in pure water and in pure Cyrene (ketone form). This was carried out using two different parametrizations of COSMO-RS, BP_TZVP_19.ctd and BP_TZVPD_FINE_19.ctd. The objective is to i) show that the model accurately predicts the solubility of the solutes in pure water and pure Cyrene, and ii) that the more computationally expensive BP_TZVPD_FINE_19.ctd parametrization offers no advantage for these systems over the more commonly used BP_TZVP_19.ctd. These results are reported in Figure 2.

![Predicted vs. experimental mole fraction solubility of the studied solutes in pure water (blue) or pure Cyrene (green), predicted using either the BP_TZVP_19.ctd (diamond) or the BP_TZVPD_FINE_19.ctd (circle) parametrization of COSMO-RS. Experimental data taken from De bruyn et al.\textsuperscript{11} and from Abranches et al.\textsuperscript{2}](image)

Since we have established that COSMO-RS can predict the solubility of the hydrophobic solutes in pure water and pure Cyrene, the model is now applied to predict their solubility curves in the water/Cyrene system. The solubility of each solute was predicted in a three-fold manner, taking into account (i) only the presence of the ketone form of Cyrene, (ii) only the presence of its diol form, and (iii) the presence of both in the composition previously reported by De bruyn et al.\textsuperscript{11} (see section S1 of the Supporting Information). Even though the DFT optimization of the structure of the ketone form of Cyrene is straightforward, in the sense that there are no energetically relevant different conformers, that is not the case for the diol form of Cyrene. In this work, three conformers of the diol were optimized, differing in the number of
intramolecular hydrogen bonds formed between the hydroxyl and ether groups (two, one, or no intramolecular hydrogen bonds). The σ-surface of these conformers are depicted in Figure 3. The σ-surface of the ketone form of Cyrene and of the solutes here studied are depicted in Figure S1 of the Supporting Information. To better demonstrate the polarity difference between the three conformers of the diol form of Cyrene, their σ-profiles are depicted in Figure S2.

![Figure 3. Optimized geometry (σ-surface) of the three relevant conformers of the diol form of Cyrene. Their main difference lies on the extent of intramolecular hydrogen bonding: a) two intramolecular hydrogen bonds, b) one intramolecular hydrogen bond, and c) no intramolecular hydrogen bonds.](image)

Figure 4 depicts the solubility curves of aspirin, ibuprofen, salicylic acid, and vanillin in water-ketone (black dashed line), water-diol (red dashed line), and water-ketone-diol (full line) mixtures. These were the solutes whose solubility was accurately predicted by COSMO-RS in a quantitative manner. The curves of the remaining solutes are shown in Figure S3 of the Supporting Information. The curves depicted in Figure 4 were calculated considering the diol as possessing just one intramolecular hydrogen bond. These calculations were repeated for the other two conformers of the diol, and the resulting solubility curves are reported in Figures S4 and S5 of the Supporting Information. Note that there is no significant difference in the resultant solubility curves between the predictions based on one or two intramolecular hydrogen bonds. However, the predictions fared much worse quantitatively when no intramolecular hydrogen bonding was taken into account.
Figure 4. Predicted solubility of aspirin, ibuprofen, salicylic acid, and vanillin in water-ketone mixtures (black dashed line), water-diol mixtures (red dashed line), and water-ketone-diol mixtures (full line), along with the corresponding experimental data (◆), taken from De bruyn et al. and from Abranches et al., considering the diol conformer with one intramolecular hydrogen bond. The y-axis ($x_S$) is the mole fraction solubility of the solute and the x-axis ($x_H$) is the total mole fraction of Cyrene.

The predictions of COSMO-RS, depicted in Figure 4, are remarkable; without the need for any experimental data other than the melting properties of the solute, the solubility of aspirin, ibuprofen, salicylic acid, and vanillin in water-Cyrene mixtures were predicted successfully. The same methodology was able to capture the general trend satisfactorily for the remaining solutes (Figure S3). Note that COSMO-RS does not predict any synergistic effect between the ketone and diol forms of Cyrene, with the solute solubility in water-ketone-diol mixtures being an intermediate value between the solubility in water-ketone mixtures (always higher) and the solubility in water-diol mixtures (always lower). As such, the solubility in water-ketone-diol mixtures is equal to that of water-diol mixtures for low Cyrene concentration and is equal to that of water-ketone mixtures for high Cyrene concentration.

The fact that COSMO-RS systematically predicts the higher solubility in the binary water-ketone system than on the water-diol system is consistent with our previous work, where the
cooperative hydrotropy model suggested that the ketone form of Cyrene, rather than its diol form, played the dominant role of hydrotrope. Even more interesting, according to the predictions of COSMO-RS, the maximum in solubility of these systems occurs due to the presence of the ketone form. That is, the maximum is correctly predicted by COSMO-RS in the water-ketone system, without the need for considering the presence of the diol form of Cyrene. Patently, there is a change in the solubilization mechanism at the maximum. This can be interpreted as water behaving as a cosolvent when added to Cyrene, and is important from an application perspective, since a small amount of water can increase the solubility of hydrophobic substances in Cyrene. The mechanism behind this phenomenon is rationalized in the next section.

3.2 Water-Solute Specific Interactions

The fact that the addition of small quantities of water to Cyrene increases the solubility of hydrophobic solutes is, at first glance, paradoxical. By their apolar nature, hydrophobic substances do not favorably interact with water, which is the reason behind the low water solubility of the solutes here studied. To rationalize this phenomenon of practical importance in terms of applying Cyrene as a solvent is the objective of this section.

The good prediction of the solubility curves depicted in Figure 4 indicates that COSMO-RS correctly describes the underlying phenomenon and the molecular mechanism behind this enhanced solubility. The results obtained suggest that the solubility maximum is caused by the ketone form of Cyrene rather than the diol form. Hence the interpretation of the problem can be simplified by ignoring the presence of the diol form of Cyrene. This is reasonable also from a physical perspective, considering that the ketone is its prevailing form at high Cyrene concentration.

Within the framework of COSMO-RS, the software COSMOtherm calculates the most probable contacts between molecule pairs in a mixture. Since these contacts may help shed light on the role of water in increasing the solubility of hydrophobic solutes in Cyrene, they were calculated for the systems depicted in Figure 4 at low water concentration and are depicted in Figure 5.
Figure 5. Interaction geometry of the most probable contacts between the solute and water in water/Cyrene mixtures at low water concentration, predicted using COSMO-RS.

Figure 5 reveals an interesting pattern; regardless of the solute, water always behaves as the hydrogen bond acceptor when interacting with the solute (at low water concentration). This is not surprising for the acidic solutes (aspirin, ibuprofen and salicylic acid), since the hydrogen bond between their acidic proton and the oxygen atom of water is quite strong. However, it is surprising for vanillin, whose sole positive proton is hindered by its intramolecular hydrogen bond with the neighbor methoxy group. Thus, the results depicted in Figure 5 suggest that water preferentially interacts with the solute by behaving as a hydrogen bond acceptor. This can be rationalized by considering the structure of the ketone form of Cyrene. As evident from its σ-surface (Figure S1), the ketone form of Cyrene possesses three hydrogen bond acceptor centers but no hydrogen bond donor centers. Thus, Cyrene has a preference to interact with hydrogen bond donors. As such, water may be acting as a cosolvent by establishing a favorable interaction with the solutes, effectively doubling their hydrogen bond donating ability, thereby enhancing their interaction capability with Cyrene and, consequently, increasing their solubility.
To test the hypothesis presented in the previous paragraph, the contact probability ($P$) between water molecules was calculated within the framework of COSMO-RS, using the software COSMOtherm, for the aspirin system. $P$ was calculated for the entire composition range of the water/Cyrene system, starting from solute concentration of zero (to serve as reference) and increasing the concentration of the solute, until saturation. These results are reported in Figure 6. Note that it is advantageous to plot the probability of water-water contacts as a function of its mole fraction (as opposed to plotting it against the mole fraction of Cyrene) since this allows a fair comparison between systems that possess the same amount of water.

![Figure 6](image.png)

**Figure 6.** Probability of water-water contacts in aspirin-water-Cyrene mixtures, considering only the presence of the ketone form of Cyrene, for different solute mole fractions (see inset).

For high concentration of water (above a water mole fraction of about 0.5), Figure 6 reveals that the presence of the solute does not significantly change the interactions between water molecules. This is consistent with the mechanism of hydrotropy, where water is the effective solvent of the system and the main interactions of the solute are water-mediated aggregation of apolar moieties of the solute and the hydrotrope (Cyrene). However, this clearly changes at low water concentration, where hydrotropy ceases to occur and co-solvency sets in. At that concentration, the probability of water-water contacts changes with solute concentration. That is, water-water contacts are maximized in the absence of the solute and decrease with increasing solute concentration. This is an indicator that water is in fact favorably interacting with the solute, supporting the hypothesis of strong molecular pair interaction between water and the solute as the reason behind the co-solvent behavior of water in Cyrene.

The strong solute-water interactions, suggested by Figures 5 and 6 as the reason behind the co-solvency behavior of water in Cyrene, are supported further by the total hydrogen bonding
energy of water and the ketone form of Cyrene, predicted by COSMO-RS and depicted in Figure 7.

![Graphs showing hydrogen bonding energy of water and ketone form of Cyrene](image)

**Figure 7.** Total hydrogen bonding interaction of water (left) and the ketone form of Cyrene (right) in aspirin-water-Cyrene mixtures, considering only the presence of the ketone form of Cyrene, for different solute mole fractions (see inset).

Figure 7 shows that the presence of the solute severely impacts the total hydrogen bonding energy of water at low water concentration. More specifically, the increase of solute mole fraction leads to much higher (in absolute terms) hydrogen bonding energy, suggesting that water is establishing a strong hydrogen bonding with the solute. This effect is no longer significant for a water mole fraction larger than approximately 0.3, shown in Figure 6. In sharp contrast to water, the total hydrogen bonding interaction energy of Cyrene is not affected by the presence of the solute, as has been shown by Figure 7. This suggests that Cyrene preferably establishes hydrogen bonds with water and, thus, supports the role of water as an interaction bridge between the solute and Cyrene.

**4. Conclusions**

COSMO-RS revealed in this work to be a valuable tool in predicting the solubility of hydrophobic substances in water/Cyrene mixtures, being capable of describing the solubility curves of most systems remarkably well. Furthermore, the model reveals that the ketone form of Cyrene is responsible for the solubility enhancement in the hydro trope regime and in the co-solvency regime. In fact, COSMO-RS correctly predicts the maximum in the solubility curves of the studied systems, attributing its mechanism to the ketone form of Cyrene, with the presence of its diol form not significantly affecting the position of the maximum.

Furthermore, Cyrene has been shown not to behave always as a hydro trope, contrary to previous expectations. In fact, depending on its concentration, Cyrene can act either as a
hydrotrope or as a co-solvent. In the latter case, for high concentration of Cyrene, water tends to act as a hydrogen bond acceptor for the solutes, effectively doubling their hydrogen bond donating ability, which increases their interaction capacity with Cyrene, thus increasing their solubility. Thus, this work shows that hydrotropy and co-solvency may occur for the same system, depending on its concentration, and that the same molecule (Cyrene) can act both as a hydrotrope or as a co-solvent.

Supporting Information

The Supporting Information is available free of charge at: https://pubs.acs.org/doi/10.1021/acs.iecr.8b03221.

SI includes calculation details, polarity and geometry of molecules, and additional solubility results. Figure S1: chemical structures and σ-surface. Figure S2: σ-profiles of the three relevant conformers of the diol form of Cyrene. Figure S3: predicted solubility considering the diol conformer with one intramolecular hydrogen bond. Figure S4: predicted solubility considering the diol conformer with no intramolecular hydrogen bond. Figure S5: predicted solubility considering the diol conformer with two intramolecular hydrogen bonds.

Acknowledgements

This work was developed within the scope of the project CICECO-Aveiro Institute of Materials, UIDB/50011/2020 & UIDP/50011/2020, financed by national funds through the FCT/MEC and when appropriate co-financed by FEDER under the PT2020 Partnership Agreement. FCT is also acknowledged for the financial support through the project SAICTPAC/0040/2015.
5. References


(11) De bruyn, M.; Budarin, V. L.; Misefari, A.; Shimizu, S.; Fish, H.; Cockett, M.; Hunt,


Table of Contents