Aqueous solubilities of five \(N\)-(diethylaminothiocarbonyl)benzimido derivatives at \(T = 298.15\) K

Bernd Schröder \(^a\),* Mónia A.R. Martins \(^a\), João A.P. Coutinho \(^a\), Simão P. Pinho \(^b\)

\(^a\) CICECO, Departamento de Química, Universidade de Aveiro, 3810-193, Aveiro, Portugal
\(^b\) LSRE-LCM, Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Bragança, 5300-253, Bragança, Portugal

**Highlights**

- Aqueous solubilities of the title compounds were measured at \(T = 298.15\) K.
- Aqueous solubilities of the title compounds were estimated with COSMO-RS.
- Possible environmental partition of the title compounds was qualitatively portrayed.

**Abstract**

\(N\)-(diethylaminothiocarbonyl)benzimido derivatives are polar multifunctional substances. A set of these compounds was synthesised by successive substitution on the enamine side, resulting in similar substances with different polarities, providing a set of model compounds with respect to the study of substituent effects on physico-chemical properties. Experimental aqueous solubility data, at \(T = 298.15\) K, of \(N\)-(diethylaminothiocarbonyl)benzamidine, PhCNH\(_2\)NCSNEt\(_2\) (1), \(N\)-(diethylaminothiocarbonyl)-\(N\)-phenylbenzamidine, PhCNHPhNCSNEt\(_2\) (2), \(N\)-(diethylaminothiocarbonyl)-\(N\)-monoethylbenzamidine, PhCNHEtNCSNEt\(_2\) (3), \(N\)-(diethylaminothiocarbonyl)-\(N\),\(N\)-diethylbenzamidine, PhCNEt\(_2\)NCSNEt\(_2\) (4), and \(N\)-(diethylaminothiocarbonyl)benzimido ethylester, PhCOEtNCSNEt\(_2\) (5) were measured at \(T = 298.15\) K. The obtained data are supplemented by COSMO-RS aqueous solubility predictions as well as other environmentally important partition coefficients. This information is shown in a two-dimensional chemical space diagram, providing indications about the compartment into which the bulk of the compounds is likely to concentrate. The expected quality of COSMO-RS predictions for this type of screening exercise is illustrated on a set of pesticides with established thermophysical property data.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

\(N\)-(diethylaminothiocarbonyl)benzimido derivatives are relatively stable substances belonging to the group of thiourea derivative compounds (Beyer et al., 1984). They are accessible through the reaction of coordinated \(N\)-Acylthiourea with acid chlorides, followed by further derivatisation. Members of this family are known chelating agents (Hartung et al., 1985), with potential applications e.g., radio pharmaceuticals \(^{99m}\text{Tc}\) (Abram et al., 1989). A large number of serine proteinase inhibitors has been developed, starting from benzamidine and its derivatives (Ensinck et al., 1972; Henkel et al., 1983; Jeffcoate and White, 1974). Reports on the thermochemistry of \(N\)-(diethylaminothiocarbonyl)benzimido derivatives (Ribeiro da Silva et al., 2006, 2004b) and some of their Ni-complexes (Ribeiro da Silva et al., 2004a) are available. More recently, they were systematically examined for their crystalline structures (Schröder et al., 2011b, 2011c), with intent to connect to their fusion thermodynamics (Schröder et al., 2011b).

Here, we report on experimental aqueous solubility data, at \(T = 298.15\) K, of the following compounds: \(N\)-(diethylaminothiocarbonyl)benzamidine, PhCNH\(_2\)NCSNEt\(_2\) (1), \(N\)-(diethylaminothiocarbonyl)-\(N\)-phenylbenzamidine, PhCNHPhNCSNEt\(_2\) (2), \(N\)-(diethylaminothiocarbonyl)-\(N\)-monoethylbenzamidine, PhCNHEtNCSNEt\(_2\) (3), \(N\)-(diethylaminothiocarbonyl)-\(N\),\(N\)-diethylbenzamidine, PhCNEt\(_2\)NCSNEt\(_2\) (4), and \(N\)-(diethylaminothiocarbonyl)benzimido ethylester, PhCOEtNCSNEt\(_2\) (5).
These compounds have all the same chemical core structure, only differentiated by successive substitution on the enamine side (1–4) or its complete replacement by an ester group (5), making them an interesting study subject. They represent a class of rather simple, but nevertheless multifunctional compounds, whose thermodynamic properties, due to the combination of their functional groups, are not easy to predict. Quite often, accurate experimental data concerning equilibrium partition constants of such substances are not available. Furthermore, given the large number of chemicals in use, accompanied by frequently incomplete data sets, reliable methods for rapid screening of substances for their persistence in the environment are required (Strempel et al., 2012). A first screening procedure based on physical-chemical equilibrium partitioning data between air, water and octanol can assist in identifying environmental compartments for which degradation half-lives might be needed (Gouin et al., 2000).

A computational tool under attention is COSMO-RS, from which a variety of environmentally important partition coefficients was demonstrated on a variety of occasions, with respect to a diversity of compounds as well as properties (Arp et al., 2006; Klamt et al., 2002; Nakajoh et al., 2009; Niederer and Goss, 2008; Oleszek-Kudlak et al., 2005; Schröder et al., 2011a, 2010a, 2010b; Wang et al., 2013; Wittekindt and Goss, 2009; Zhang et al., 2010).

Hence our principal goals are twofold: to provide a) reliable aqueous solubility data and b) check on the predicting performance of COSMO-RS regarding these compounds, with respect to the obtained experimental data. Organizing the results in the form of partition coefficients in a two-dimensional chemical space diagram provides information about the partition behaviour of the title compounds once released in the environment. For the sake of comparison, structurally related, mostly urea-based pesticides, with established thermophysical property data sets were selected and also plotted in the space diagram.

### 2. Methodology

#### 2.1. Compounds

The synthesis was performed as described in detail elsewhere.

<table>
<thead>
<tr>
<th>Cpd. No.</th>
<th>Name</th>
<th>Structural formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N-**(diethylaminothiocarbonyl)**benzamidine</td>
<td><img src="image1" alt="Structure 1" /></td>
</tr>
<tr>
<td>2</td>
<td>N**(diethylaminothiocarbonyl)**-N’-phenylbenzamidine</td>
<td><img src="image2" alt="Structure 2" /></td>
</tr>
<tr>
<td>3</td>
<td>N**(diethylaminothiocarbonyl)**-N’-monoethylbenzamidine</td>
<td><img src="image3" alt="Structure 3" /></td>
</tr>
<tr>
<td>4</td>
<td>N**(diethylaminothiocarbonyl)**-N’-N’-diethylbenzamidine</td>
<td><img src="image4" alt="Structure 4" /></td>
</tr>
<tr>
<td>5</td>
<td>N**(diethylaminothiocarbonyl)**benzimido ethylester</td>
<td><img src="image5" alt="Structure 5" /></td>
</tr>
</tbody>
</table>

Fig. 1. Chemical structures of the title compounds.
(Beyer et al., 1984; Schröder et al., 2011b). A resume as well as details of characterization are given as Supporting Material SM1.

2.2. Aqueous solubility measurements

The experimental solubilities of the \(N\)-(diethylaminothiocarbonyl)benzimido derivatives in water were determined at \(T = 298.15\) K and at atmospheric pressure. The experimental setup, adapted from Andersson and Schrädter (1999); Etzweiler et al. (1995); and Schrädter and Andersson (2001), is presented in Fig. 2.

Initially, test tubes (A) with solutions with an excess of solid were prepared and a dialysis tubing cellulose membrane (B) (DS277 from SIGMA), filled with ultra-pure water. The water used was double distilled, passed by a reverse osmosis system and further treated with a MilliQ plus 185 water purification apparatus (resistivity: 18.2 MΩ cm; TOC < 5 µg dm\(^{-3}\), free of particles > 0.22 µm). The dialysis tubing of around 10 cm length were previously humidified for at least 3 h and cleaned according to the instructions given by the manufacturer. As shown by Fig. 2, one end of the dialysis tubing (B) was closed with a tight knot and the other was fixed to a glass tube (C). This glass tube allows the sampling through a rubber cup (D).

The test tubes solutions were dispersed using an isothermal ultrasonic bath (Branson 250 & 450 Sonifier) during one hour at 60%, and then allowed to equilibrate in a thermostatted Julabo F38 - EH (V2) bath (E) under agitation (F) for at least 24 h, at 298.15 K. Stirring was carried out by Thermo Scientific™Cimarec™ Micro Stirrers. This period of time proved to be enough to guarantee the saturation. The temperature accuracy of the bath was ±0.03 K.

Samples from the inside of the dialysis membrane were collected using plastic syringes maintained at the same temperature of the saturated solution. Solute concentration was obtained by UV–vis spectroscopy, using a SHIMADZU UV-1700 PharmaSpec Spectrometer, at wavelengths of 270, 267, 269, 243, 281 for \(N\)-(diethylaminothiocarbonyl) benzamidine, \(N\)-\(N\)-(diethylaminothiocarbonyl)- \(N\)-diethylbenzamidine, and \(N\)-(diethylaminothiocarbonyl)benzimidethylester and \(N\)-(diethylaminothiocarbonyl)\(N\)-phenylbenzamidine, respectively. These wavelengths were found to be the maximum UV absorption wavelengths for the compounds investigated here. Due to the very low aqueous solubilities, solutions of known solute concentration were prepared in the binary water + methanol mixed solvent containing 65% (mass percentage in solute free basis) of methanol. Diluting this mother solution with the same mixed solvent, calibration curves were built relating absorbance and concentration, in the solute concentration range expected for aqueous solubility.

For sampling, approximately 0.5 g of saturated solution was collected from inside the dialysing tubing, and diluted in methanol, in order to have the same solvent composition as the calibration curve. At least six independent measurements were carried out for each average solubility value reported in Table 1, where the uncertainty is also given. Individual experimental aqueous solubility results are given in Supporting Material SM2.

2.3. Quantum chemical and COSMO-RS calculations

To obtain the necessary COSMO files, quantum-chemical optimizations were performed for all molecules in the gas phase as well as in the COSMO state, using TURBOMOLE TmoleX v.4.0.1 (Steffen et al., 2010; TURBOMOLE, 2007–2014). To generate sets of all relevant conformers, COSMOconX v.3.0 (COSMOconX, 2013) was utilized. The encountered most stable conformers are presented in Supporting Material SM3. The possible existence of tautomeric species in the title compounds was subject to previous studies. All solid structures of 1–3 exclusively show the enamine (a) structure (Braun et al., 1988a, 1988b; Schröder et al., 2011b, 2011c).

At the quantum-chemical level of theory relevant to COSMO-RS, the enamine structures are the energetically more favoured ones. This holds for the gas-phase as well as for the COSMO state, where they are more stable than the respective imino counterparts, in a magnitude of 35–55 kJ mol\(^{-1}\). Although no reports regarding enamine-imino tautomerism of the title compounds in solution exist, experimental work on similar tautomers suggested that the form with the highest dipole moment will predominate in polar solution, e.g. in the case of 3-methylcytosine (Dreyfus et al., 1976) and 1-alkyladenines (Dreyfus et al., 1977).

Supporting Material SM4a gives an overview about COSMO-RS dipole moments. From these findings, the prevalence of the imino form (b) in aqueous solution was assumed for compounds 1 and 2.

While \(pK\) values of a variety of structural analogues were successfully measured by \(pH\) potentiometry before, no data for the title compounds could be obtained (Hartung et al., 1985; Mühl et al., 1986). Hence, at this stage, we assume neutral species being predominant in aqueous solutions at environmentally relevant medium \(pH\) ranges. Furthermore, preliminary \(pK\) data estimations of the title compounds were performed using quantum-chemical calculations at the BP-TZVP level of theory in combination with a linear free energy relationship, as provided by COSMO-RS in its COSMOtherm implementation. The results are given in Supporting Material SM4b. They suggest that the initial assumption of no significant speciation for all assumed species may hold, with exception of the imino form of compound 1, for which an occurrence of...
30% of protonated species at pH = 7 would require a dissociation correction of 0.005 in log S. With respect to pKₐ, no meaningful data could be produced for compound 5.

All optimizations were performed at the BP-TZVP as well as the BP-TZVPD-FINE level of theory. In order to assure to have reached a true minimum by excluding the appearance of imaginary frequencies, the encountered global gas phase minimum was subjected to vibrational frequency calculations with AOFORCE, at the respective BP-TZVP level of theory. Resulting COSMO energies of the most stable conformers are given in Supporting Material SM3, as well. Finally, the physico-chemical properties were estimated, using the parameter file BP_TZVPD_C30_1401.cdt (for COSMO files created at the BP-TZVP level of theory) and BP_TZVPD_FINE_C30_1401.cdt (for COSMO files created at the quantum chemical level BP-TZVPD-FINE, with a novel hydrogen bond interaction term and a novel van der Waals dispersion term) (Eckert and Klamt, 2014). In our work, all COSMO-RS calculations were performed with its COSMOtherm implementation (Eckert and Klamt, 2014). In the following, a few more details are given concerning the use of COSMO-RS to calculate aqueous solubility and partition coefficients.

2.3.1. Aqueous solubility

The mole fractions of the solute are refined using the automatic solubility calculation option of COSMOtherm:

\[
\log_{10}(x_2) = \left(\mu_2^{(p)} - \mu_2^{(1)} - \max(0, \Delta_{\text{ fus}} G)\right) / (RT \ln(10))
\] (1)

where \( R \) is the ideal gas constant, \( T \) is the absolute temperature, \( x_2 \) is the mole fraction of solid dissolved in the solvent phase 1 at saturation, \( \mu_2^{(p)} \) is the chemical potential of pure compound 2, \( \mu_2^{(1)} \) is the chemical potential of compound 2 at infinite dilution in the solvent, compound 1. The program possesses experimental free energy of fusion data, \( \Delta_{\text{ fus}} G \).

2.3.2. Air-water partition coefficient

Common representations of the Henry's law constant are depicted in Equations (2) and (3):

\[
H = \left(\frac{Pa \ m^3 \ mol^{-1}}{c_2}\right) = \frac{P_2}{c_2}
\] (2)

with \( c_2 \) as the molar concentration of the solute, and, furthermore, the air-water partition coefficient \( K_{\text{AW}} \) is defined accordingly to

\[
\log K_{\text{AW}} / (\text{dimensionless}) = \log \left(\frac{H}{RT}\right)
\] (3)

2.3.3. Octanol-water partition coefficient

In COSMO-RS, the octanol-water partition coefficient, \( K_{\text{OW}} \), is predicted via computation of the chemical potentials of the solute at infinite dilution in each of the solvents:

\[
\log K_{\text{OW}} = \log_{10} \left\{ \exp \left[ \left( \frac{\mu_1^{xW}}{RT} - \frac{\mu_1^{xO}}{RT} \right) \frac{V_{W,m}}{V_{O,m}} \right] \right\}
\] (4)

with \( \mu_1^{xW} \) as the chemical potential of compound \( x \) in water, and \( \mu_1^{xO} \) as the chemical potential of the same compound in 1-octanol. Two different \( K_{\text{OW}} \) coefficients, the so-called wet and dry, can be calculated, being the last useful to calculate the octanol-air partition coefficient \( K_{\text{OA}} \). A ratio of molar volumes \( V_{W,m}/V_{O,m} \) of 0.1505 (wet) and 0.1141 (dry) was used. Besides the density difference of the solvents, their mutual solubility is taken into account in the case of \( K_{\text{OW}} \) (wet), which corresponds to 0.274 mol fraction of water in the octanol-rich phase (Dallos and Liszi, 1995; Eckert, 2013) at 298.15 K.

2.3.4. Octanol-air and soil sorption coefficients

The octanol-air coefficient is frequently used to describe partitioning of organic substances between air and organic phases in soils, plants and atmospheric aerosols. It has been obtained according Equation (5):

\[
\log K_{\text{OA}} = \log K_{\text{OW}}(\text{dry}) - \log K_{\text{AW}}
\] (5)

The data for \( K_{\text{OW}} \) (dry) can be generated in COSMO-RS using the pure quantum chemical approach. Soil-sorption partition coefficients \( K_{\text{OC}} \) are directly accessible from the “Environmental Property” section of the current version of the program.

3. Results and discussion

3.1. Experimental aqueous solubilities

Experimental solubility data measured in this work is presented in Table 1, including the expanded uncertainty for a 95% confidence interval. As expected, uncertainty increases with decreasing solubility. The results are, however, very satisfactory since the uncertainty divided by the solubility value presents a maximum around 12.1% for substance 2, the least soluble, while that statistical parameter is lower than 5% for solubility values higher than 1.00E-04 g/gH₂O.

The experimental procedure followed here was used in our laboratory for the first time. In order to validate the method, the experimental determination of the solubility of naphthalene in water at 298.15 K was also carried out, which has a solubility value of the same order of magnitude of the studied substances. Exactly the same steps were followed, and the average value of six independent measurements was 3.219E-05 g/gH₂O, presenting and extended uncertainty of 2.730E-06 g/gH₂O.

In the compilation book by Yalkowsky et al. (2010), 28 solubility values are available at 298.15 K. Eliminating four evident outliers, the solubility ranges from 2.957E-05 g/gH₂O to 3.450E-05 g/gH₂O, presenting an average of 3.189E-05 g/gH₂O. Therefore, the results found in this work show the good reliability of the method.

3.2. COSMO-RS calculations

3.2.1. Qualitative considerations

COSMO sigma profiles (of the energetically most stable conformers) and sigma potentials of the title compounds are shown in Figs. 3 and 4, respectively. The inspection of the sigma profiles highlights differences in the local polarization-charge densities, which ultimately define the differentiation in the interaction energies of the surfaces, and hence, the magnitude of all related properties to be predicted. In the histogram, the range beyond \( \sigma = \pm 0.01 \text{ e A}^{-2} \) is considered as being strongly polar and potentially hydrogen-bonding, while the remaining part is weakly to

<table>
<thead>
<tr>
<th>Substance</th>
<th>Experimental (g g⁻¹H₂O)</th>
<th>U (g g⁻¹H₂O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.280E-04</td>
<td>3.597E-06</td>
</tr>
<tr>
<td>2</td>
<td>8.871E-06</td>
<td>1.071E-06</td>
</tr>
<tr>
<td>3</td>
<td>1.744E-04</td>
<td>9.575E-06</td>
</tr>
<tr>
<td>4</td>
<td>3.168E-04</td>
<td>1.082E-05</td>
</tr>
<tr>
<td>5</td>
<td>3.738E-05</td>
<td>4.199E-06</td>
</tr>
</tbody>
</table>
non-polar. All molecules display a distinct peak at 0.011 e Å⁻², with exception of 3 and 4, which are slightly shifted towards 0.014 e Å⁻². The peak familiar to all molecules arises mainly from the negatively polar thiocarbonyl sulphur. The largest peak belongs to 5; here, the ester oxygen contribution is adding up. On the other side of the histogram, compounds 1 to 3 show potential to form intermolecular hydrogen bonds. This is further depicted in the hydrogen-bond moments, referring to the BP-TZVP level of theory, as Supporting Material SM4a. Compounds 1 to 3 are identified being a hydrogen-bond donor. Dipole moments decrease in the order 4 > 3a > 5 > 1b > 2b. The sigma potential is a measure for the affinity of the system S to a surface of polarity σ. Non-polar molecules with purely dielectric behaviour exhibit a simple parabola centred at σ = 0. Compounds 1 to 3 indicate their hydrogen bonding donor capacity, while the other compounds show nearly parabolic behaviour in the positive region. The sigma potentials of the compounds in the negative region indicate increasingly unfavourable interaction of compounds with themselves in the order 2 > 5 > 3 > 4 > 1. Thermophysical data obtained in a crystal-liquid equilibrium study (Schröder et al., 2011b) of all compounds show (Supporting Material SM5) that 4 exhibits the lowest enthalpy of fusion, followed by 1. Additionally, 4 shows a remarkably low fusion temperature when compared with the other N-
(diethylaminothiocarbonyl)benzimido derivatives. These results were in agreement with the existing crystal packing constraints due to the non-existing intermolecular NH-S hydrogen-bond interactions in 4 and 1 (Braun et al., 1988a). From all of these qualitative considerations, one could preliminarily expect 4 and 1 possessing the highest aqueous solubilities among the compounds studied here.

3.2.2. Aqueous solubility calculations

Aqueous solubilities were calculated at $T = 298.15 \text{ K}$, both at the BP-TZVP and BP-TZVPD-FINE level of theory, considering the crystalline state of the title compounds, based on the thermophysical fusion data given in Supporting Material SM5. The results are given in Table 2 and compared with the obtained experimental data. Additionally, prediction results with a simpler model as provided by EPI suite WSKOWwin v 1.42 are presented, as well (US EPA EPI Suite 4.1 (2012)). WSKOWwin predicts the water solubility of an organic compound using the compounds log octanol-water partition coefficient as provided by the estimation engine of the KOWwin program, as well as the respective melting point, with the results being subject to certain structure-dependent corrections. In all cases, COSMO-RS tends to underestimate aqueous solubilities, more pronounced with the BP-TZVPD-FINE approach. Overall performance of WSKOWwin falls in-between both COSMO-RS levels of theory, underestimating aqueous solubilities for all compounds except 1, and a remarkably high deviation in 4. WSKOWwin predicts a distinct decrease of aqueous solubility with increasing degree of alkylation from 1 to 3 to 4, which is opposite to the experimental finding.

For the title compounds, the BP-TZVP level of theory performs best overall, with a standard error $\sigma$ of 0.66 in log $S$, and a smallest deviation in log $S$ of 0.2 for compound 4, while the highest deviation in log $S$ is 1.0 for compound 5. At the BP-TZVP level of theory, the experimental finding of similar magnitudes of aqueous solubilities of compounds 1, 3 and 4 could be reproduced. For a small set of urea-based pesticides containing similar structural features and for which experimental fusion and aqueous solubility data are available, a standard error of 0.25 in log $S$ was obtained, slightly underestimating the solubilities in all cases except for propchloron, propanil and carbaryl, considering the BP-TZVP level of theory, as given as Supporting Material SM6 and SM7.

To the best of our knowledge, no vapour pressure data of the title compounds are reported, yet, neither for the (subcooled) liquid nor the crystalline state. Calorimetrically obtained standard molar enthalpies of sublimation, $\Delta H_{m}^{0}(298.15 \text{ K})$, of N-(diethylaminothiocarbonyl)benzimido derivatives lit-erature data and COSMO-RS results, both at the BP-TZVP and BP-TZVPD-FINE level of theory.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Experimental</th>
<th>BP-TZVP</th>
<th>BP-TZVPD-FINE</th>
<th>WSKOWwin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>126.0 ± 1.5a</td>
<td>117.7</td>
<td>115.3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>159.4 ± 3.3a</td>
<td>134.8</td>
<td>128.4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>141.2 ± 1.2b</td>
<td>139.7</td>
<td>130.0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>122.2 ± 2.0c</td>
<td>126.0</td>
<td>119.3</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>135.6 ± 2.6c</td>
<td>130.9</td>
<td>122.7</td>
<td></td>
</tr>
</tbody>
</table>

* Ref.: Ribeiro da Silva et al. (2004b).

The deviation for compounds 2 is biggest for both methods and rather distinct from the others. The overall performance is good and slightly better for BP-TZVP (AAD of 5.9%) than BP-TZVPD-FINE (AAD of 9.6%).

3.2.3. Further equilibrium partition coefficients as obtained from COSMO-RS

Today, a variety of partitioning property estimation methods is used in the attempt to prioritize substances according to their potential environmental hazards. Main targets of screening exercises are information about persistence, bioaccumulation potential, toxicity and long-range transport potential of existing and new chemicals. Equilibrium partition coefficients like log $K_{AW}$, log $K_{OW}$, log $K_{OA}$ and log $K_{OC}$ are key parameters in the process. Since experimental data are often unavailable, reliable prediction methods are of utmost importance. Current prediction methods are consistent only to a certain degree: a screening test of 529 substances using four established prediction methods showed a mere consistency for ~70% of the set members (Zhang et al., 2010).

Further physico-chemical data relevant for environmental purposes obtained from COSMO-RS are presented in Supporting Material SM8. Results for the chosen urea-based pesticides are given as Supporting Material SM9. The deviations are given in Supporting Material SM10 and show a good performance of COSMO-RS predictions. Supporting Material SM11 compiles data on hybrid air-water partition coefficient of the title compounds, where the experimental value stems from this work aqueous solubility measurements, while vapour pressures were estimated with COSMO-RS.

The derived partition coefficients may be introduced into a two-dimensional plot describing a hypothetical chemical space as depicted in Fig. 5. Here, the environment is modelled as volumes of air, water, and octanol, where octanol represents the organic fraction appearing in soils and sediments (Gouin et al., 2000). More details on the approach are given as Supporting Material SM12.

The plot allows a first screening of compounds with respect to their probable distribution in the environment once (hypothetically) released. It also gives an idea about the deviation between experimental data and COSMO-RS predictions with respect to the equilibrium partition coefficients and the impact on chemical space distribution.

For the title compounds, experimental solubilities are used in the calculation of log $K_{AW}$, all other properties were calculated with COSMO-RS. Since experimental log $K_{OW}$ (dry) data were not available, log $K_{OW}$ (wet) data were used throughout. The rather small deviations between both, as indicated in Supporting Material SM8 and SM9, justifies this approach in this qualitative screening exercise.

The 1% and 99% lines in Fig. 5 divide the $K_{AW}/K_{OW}$ space into

regions in which partitioning is occurring almost exclusively into one medium. It is likely that degradation processes in that medium are most important, and hence, respective data are to be collected, in order to proceed with a more complete study of the compound’s behaviour once released into the environment. For instance, from the title compounds, 2 is to be expected to mainly partition into soil and sediment, its rather high log $K_{OW}$ value renders the compound prone to be persistent in the environment. The other title compounds will partially be found in the water phase, as well; especially, due to their relative lower content of carbohydrate fragments, compounds 1 and 4.

In a first approximation, the substances of the set are mainly partitioning into soil and the aquatic environment. The quality of used input data may alter the detected principal environmental compartment of their occurrence. In our example, the input only differs in log $K_{OW}$, which leads to vertical shift between the data based on experimental aqueous solubility data and COSMO-RS solubilities, while all other data are obtained with COSMO-RS. The air-water partition coefficient is the compound’s property most intimately related to aqueous solubility and vapour pressure. Simultaneous over (-or under)estimation of a comparable magnitude of aqueous solubilities as well as vapour pressures can largely be cancelled out in the Henry’s law constant itself, as was found e.g., in the case of alkylated naphthalenes (Schröder et al., 2013).

Until experimental vapour pressures or experimentally determined air-water partition coefficients of the title compounds are known for confirmation, the COSMO-RS log $K_{OW}$ data provide a reasonable estimate. The shift resulting from the different input data sets does, in this specific case, not influence the expected environmental compartment in which the bulk of each compound will likely to partition.

In order to give indications about the adequacy of COSMO-RS to build a preliminary chemical space diagram, Fig. 5 also includes a comparison for a set of urea-based pesticides with established consistent thermophysical property data. As can be seen the results are promising as the predictions are globally close to the experimental information. On the other hand, calculated environmental properties of some compounds show rather distinct deviations from experiment, and hence, substantially different placements in Fig. 5. Fenuron and linuron for instance exhibit deviations in their aqueous solubilities falling within the expected range. But rather high deviations in their vapour pressure predictions are encountered, when compared to recommended data. These recommendations are known for their limited quality, e.g. solid vapour pressures of linuron (Mackay et al., 2006). Barban and diphenamid, with acceptable to excellent deviations concerning vapour pressures and aqueous solubilities, show a remarkable, in terms of magnitude, atypical overestimation of their lipophilicity - when compared to experiment. In such cases, COSMO-RS can serve as a screening tool for sorting out problematic experimental data, given the known difficulties in obtaining them.

Furthermore, COSMO-RS results for a set of more modern fluorine-containing pesticides [G. Theodoridis, Fluorine-Containing Agrochemicals: An Overview of Recent Developments, in Fluorine...].
and the Environment, Vol. 2, Chapt. 4, Elsevier, 2006) are included (please refer to Supporting Material SM13 and SM14). These are mostly benzoylphenyl urea derivatives used as commercial insect growth regulators (IGRs). To our best knowledge, no agreed-upon consistent thermophysical property data are available in the literature, at this point of time (except diflubenzuron). Introducing electron-withdrawing substituents often extends a compound’s pectesticidal spectrum, but also affects its environmental behaviour. In the exemplary case of dichlorobenzuron, a 2,6-dichloro benzoyl derivative, and diflubenzuron, its 2,6-difluorinated congener, COSMO-RS calculations suggest that the substitution of chlorine by fluorine augments the vapour pressure more than it increases its aqueous solubility, in relative terms. Next to the therefore increased Henry constant, a less pronounced lipophilic character is illustrated in the lower log $K_{ow}$ value of diflubenzuron, leading to the spatial separation of both compounds in Fig. 5. Furthermore, dichlorobenzuron is known to degrade in soil within six to twelve months, while diflubenzuron has a half-life in soil of about three days (Luteijn and Tipker, 1986). The different sizes of the introduced halogens lead to distinctively different molecular structures and hence, to different metabolic pathways (jeschke, 2004).

To obtain further hints on a possible PBT character of the title compounds, another screening exercise was performed with the online-calculator PBT Profiler, whose estimates are designed for screening-level assessments on persistence (P), bioaccumulation (B), and/or toxicity ($T$) (PBT Profiler, 2015).

The PBT Profiler expresses (reactivity-based) persistence in single medium half-lives, measured in days, in air, water, soil, and sediment. The program first determines the media a chemical is most likely to be found in, using a Level III multi-media model where advective losses are accounted for. Furthermore, the program provides estimates regarding the bioconcentration factor (BCF) as well as the chemical’s relative toxicity in the form of a long-term toxicity value (Fish ChV). The results with regard to the title compounds are given as Supporting Material SM15 and SM16.

They suggest that actually all title compounds might be of interest in terms of potential PBT characteristics. Their estimated half-life in soil, 75 days, exceeds the EPA criteria of ≥ 2 months. Therefore, they are estimated to be persistent in the environment, once released. All compounds are assigned a chronic toxicity value of fish above EPA threshold limits, except 1.

4. Conclusion

The applied aqueous solubility measurement method proved to give reliable and precise experimental data. COSMO-RS tends to underestimate aqueous solubilities of the title compounds, with respect to the experimental data, yielding best results at the BP-TZVP level of theory.

COSMO-RS is a useful tool to predict the distribution of a species between different compartments in preliminary screening exercises with the aim to prioritize compounds with potential impacts in the environment. The reliable prediction of physico-chemical properties of multi-functional substances remains a challenging task, as well as the quantification of numerical hazard estimates with respect to their uncertainties (Arnot and Mackay, 2008; Strempel et al., 2012).

This in turn calls for improvements in the availability of reliable experimental physico-chemical property data. Although measuring all relevant partition parameters of all multi-functional compounds is impossible, given their shear number, the opportunity given by the technological developments in high-accuracy property determinations should be concisely applied. A concerted effort should be focused in measuring sets of polar key compounds to be defined, and re-measuring other important compounds where available data are of dubious quality, even more so as these data are supposed to serve as the future’s foundation of predictive routines.

Acknowledgments

Bernd Schröder acknowledges Fundação para a Ciência e a Tecnologia (FCT) and the European Social Fund (ESF) under the 3rd Community Support Framework (CSF) for the award of a post-doctoral grant (grant number SFRH/BPD/38637/2007) and the award of research project PTDC/AAC-AMB/121161/2010. This work was developed in the scope of the projects CICECO-Aveiro Institute of Materials (Ref. FCT UID/CTM/50011/2013) and LSRE-LCM Associate Laboratory (Project POCI-01-0145-FEDER-006984), financed by national funds through the FCT and by FEDER funds through COMPETE 2020-POCI. M.A.R.M thanks to FCT for financial support for the PhD grant SFRH/BD/87084/2012.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.chemosphere.2016.06.042.

References


COSMOCanOnX v3.0, COSMOlogic GmbH & Co. Leverkusen; 2013.


Henkel, H.J., Labes, D., Hagen, V., 1983. Quantum chemical studies on substituted
benzamidines as inhibitors of the serine proteinases trypsin and thrombin. Pharmazie 38, 342–346.


