

# COSMO Implementation in TURBOMOLE: Extension of an efficient quantum chemical code towards liquid systems†

Ansgar Schäfer,<sup>\*a</sup> Andreas Klamt,<sup>b</sup> Diana Sattel,<sup>a</sup> John C. W. Lohrenz<sup>c</sup> and Frank Eckert<sup>b</sup>

<sup>a</sup> BASF AG, Scientific Computing, ZDP/C-C13, 67056 Ludwigshafen, Germany

<sup>b</sup> COSMOlogic GmbH & Co. KG, Burscheider Strasse 515, 51381 Leverkusen, Germany

<sup>c</sup> Bayer AG, PF-MWF/BT, Pflanzenschutzzentrum Monheim, Geb. 6500, 51368 Leverkusen, Germany

Received 6th January 2000, Accepted 24th March 2000

Published on the Web 2nd May 2000

The most recent algorithmic enhancements of the COSMO solvation model are presented and the implementation in the TURBOMOLE program package is described. Three demonstrative applications covering homogeneous catalysis, tautomeric equilibria, and binary phase diagrams show the efficiency and general applicability of the approach. Especially when combined with the COSMO-RS extension, the method very reliably predicts thermodynamic properties of liquid mixtures.

## 1. Introduction

TURBOMOLE<sup>1</sup> has developed as an efficient and reliable quantum chemical program package which is widely used in academia and in industry. Nevertheless, up to this work TURBOMOLE was restricted to the calculation of molecules *in vacuo*, i.e. to the state of molecules in the ideal gas phase. Due to the overwhelming role of fluid-phase reactions in technical chemistry as well as in biochemistry, it is desirable and in many cases of crucial importance to handle more accurately the special effects caused by a solvent. For this reason the conductor-like screening model (COSMO),<sup>2</sup> which was developed by one of the authors at Bayer AG, has been implemented in TURBOMOLE in a joint initiative of BASF AG and Bayer AG. COSMO is a variant of the dielectric continuum solvation models. For an overview of the different classes of continuum solvation models the interested reader is referred to recent reviews.<sup>3–5</sup>

Since its first implementation in MOPAC in 1993,<sup>2</sup> COSMO has been implemented in several quantum chemical codes, e.g. MNDO,<sup>6</sup> AMPAC,<sup>7</sup> DMol,<sup>8</sup> GAUSSIAN94,<sup>9,10</sup> GAMESS,<sup>11</sup> ADF<sup>12</sup> and AMSOL.<sup>13</sup> The more recent implementations contain slight improvements compared to the originally published method, regarding the cavity construction as well as the algorithmic handling of the model. Meanwhile COSMO turns out to be widely accepted as a mathematically simple and nevertheless very reliable and stable continuum solvation model. This can be seen from the fact that, even in competing approaches like the polarizable continuum model (PCM),<sup>14</sup> the basic ideas of COSMO have been adopted recently.<sup>10</sup> Therefore the implementation of COSMO appears to be one of the most efficient ways towards a more realistic handling of fluid-phase chemistry in quantum chemical codes like TURBOMOLE.

In this paper a compact presentation of the theory and the actual implementation of COSMO is made in Section 2. Section 3 presents some demonstrative applications of TURBOMOLE/COSMO. In Section 3.2, some aspects of the solvent effects in the Heck reaction catalysed by palladacycles

are considered. Section 3.3 demonstrates the calculation of thermo-physical equilibrium data of dihydrotriazine isomers in different solvents, and Section 3.4 shows the calculated phase diagram of the binary mixture of propanol and water, using the extension of COSMO beyond the dielectric limits, COSMO-RS (COSMO for real solvents).<sup>15,16</sup>

## 2. Theory and implementation

The basic idea of COSMO as compared to other dielectric continuum models is to approximate the dielectric continuum by a scaled conductor. This approximation allows for the use of the relatively simple boundary condition of vanishing total potential valid in an infinite conductor instead of the much more complicated ones of a dielectric continuum. The deviations between COSMO and the rigorous dielectric results are negligible in strong dielectrics ( $\epsilon > 20$ ) and can be shown to be less than 10% even in the case of non-polar solvents, which generally have a dielectric constant of about  $\epsilon = 2$ . Since 10% surely is well within the inaccuracy introduced by the general approximation of a dielectric continuum representation of a solvent, it can be concluded that the additional deviations caused by the COSMO approximation are fair compared to the corresponding gain in mathematical simplicity.

In each dielectric continuum solvation model the solute is considered to be located in a cavity of a dielectric continuum. According to the common findings of many workers in this area,<sup>3,5</sup> for a reliable treatment of arbitrarily shaped molecules a molecular-shaped cavity definition is required, using some generalized solvent-accessible surface (SAS) based on van der Waals radii increased by approximately 20%. Such cavities no longer allow for an analytic calculation of the solvation energy like the originally used spherical and ellipsoidal cavities did. Therefore, a numerical treatment of the polarization of the continuum is required, which starts with a discretization of the cavity surface into finite elements, called segments further on. For a mathematical and algorithmic description of the interaction of the solute with the continuum it is most efficient to represent the polarization of the continuum by the corresponding screening charges on the cavity boundary. On each segment  $i$  of the  $m$  segments of the cavity a constant screening charge density  $\sigma_i$  is assumed, corresponding to a

† Dedicated to Prof. Reinhart Ahlrichs on the occasion of his 60th birthday.

screening charge  $q_i = s_i \sigma_i$ , where  $s_i$  denotes the area of segment  $i$ . In the following, a matrix denotation for the screening charges is used, introducing the  $m$ -dimensional vector of screening charges  $\mathbf{q} = (q_1, \dots, q_m)$ . The Coulomb interactions of the screening charges can be written as a symmetric matrix  $\mathbf{A}$ , which takes into account the self-interaction of the screening charge distribution on each segment by non-zero diagonal elements. Let  $\Phi^X$  be the corresponding vector of the electrostatic potential of the solute X on each segment. Then the conductor boundary equation of vanishing total potential reads

$$0 = \Phi^{\text{tot}} = \Phi^X + \mathbf{A}\mathbf{q}^* \quad (1)$$

with  $\mathbf{q}^*$  denoting the screening charges in a conductor, *i.e.* at a dielectric of strength  $\epsilon = \infty$ . The COSMO approximation now implies that at finite  $\epsilon$  the screening charges can be written as

$$\mathbf{q} = f(\epsilon) \mathbf{q}^* \quad (2)$$

with

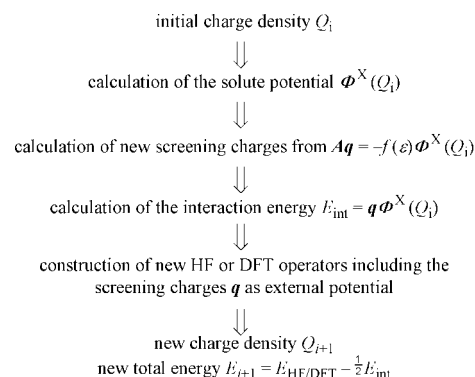
$$f(\epsilon) = \frac{\epsilon - 1}{\epsilon + x} \quad (3)$$

with a small value of  $x$  in the range  $0 \leq x \leq 1$ . As discussed elsewhere,<sup>2,17</sup> the optimal value of  $x$  is 0.5. This value will be used throughout this paper and it is generally used in the TURBOMOLE/COSMO implementation. Hence, one obtains the equation

$$\mathbf{A}\mathbf{q} = -f(\epsilon) \Phi^X \quad (4)$$

which directly relates the screening charges to the solute potential, and hence to the charge density of the solute. In contrast to the original implementations of COSMO,<sup>2</sup> in which this equation was solved by inversion of the matrix  $\mathbf{A}$ , a Cholesky factorization<sup>18</sup> is used nowadays because it is faster and demands less memory. Since the matrix  $\mathbf{A}$  only depends on the molecular geometry, the time-consuming  $O(m^3)$  step of the Cholesky factorization, which is about a factor of 6 faster than the corresponding matrix inversion, has only to be done once for a given molecular structure. The resulting factor matrices can then be used to solve eqn. (4) for  $\mathbf{q}$  in a  $O(m^2)$  step. It should be noted that a specially developed factorization routine is used, which operates on the lower triangle of  $\mathbf{A}$  and hence saves half of the memory compared to most routines available in math libraries.

Regarding the implementation of COSMO into the self-consistency part of Hartree–Fock (HF) or density functional theory (DFT), there are two different alternatives: Alternative 1, which was used in the first COSMO implementation,<sup>2</sup> splits the solute potential and the resulting screening charges into nuclear and electronic contributions and adds the resulting COSMO expressions as additional terms to the nuclear interaction energy, the nuclei–electron interaction operator (one-electron matrix), and the electron–electron interaction (two-electron matrix), respectively. Alternative 2 adds the entire electrostatic potential of the iteratively updated screening charges as an external potential to the one-electron part and takes into account the polarization energy, which is required to generate the screening charges by subsequent subtraction of half of the solute–continuum interaction energy.<sup>17</sup> Both approaches are equivalent and converge to the same total energies and solute densities. In the present implementation, alternative 2 was chosen. Using  $Q$  as a general notation for the solute density, the flow chart of an SCF cycle in HF or DFT calculations is shown in Scheme 1. Thus, the SCF procedure converges directly to the self-consistent state and energy of the molecule embedded in the dielectric continuum. It is not necessary to start with a gas-phase density. In many cases, the number of SCF cycles required is even slightly lower



**Scheme 1** Flow chart of an SCF cycle in HF or DFT calculations with COSMO.

while using COSMO with large values of  $\epsilon$  than it is in the gas phase. This is due to the fact that screening continuum density fluctuations in one part of the molecule are electrostatically decoupled from those in other parts and hence the different parts can converge almost separately.

After convergence of the SCF calculation, the total energy is corrected for the outlying charge error. This error results from the fact that, at reasonable cavity sizes, inevitably a small but significant tail of the solute electron density is located outside the cavity, *i.e.* within the dielectric continuum. This problem has been discussed in detail by Klamt and Jonas<sup>19</sup> showing that COSMO is less sensitive by about a factor of 10 to outlying charge effects on the total energy than methods using the original dielectric boundary conditions. The stability with respect to outlying charge errors is the most important reason why COSMO is starting to replace some of the direct dielectric models. Anyway, the residual outlying charge error even in COSMO often would be in the order of  $2 \text{ kJ mol}^{-1}$  for neutral compounds and up to  $10 \text{ kJ mol}^{-1}$  for ions. Therefore in the TURBOMOLE/COSMO implementation, a stable and local correction algorithm for this error is applied, which is based on an auxiliary cavity located  $0.85 \text{ \AA}$  outside the SAS.<sup>19</sup> The residual electrostatic potential on this cavity arising from the solute and from the original screening charges is calculated. This potential, which should be zero if there would be no outlying charge, originates from the outlying charge density. The respective screening charges on the outer cavity are calculated and projected onto the original cavity. Thus we get a local correction for the outlying charge error, which reduces the energetic error by at least one order of magnitude, *i.e.* to about  $0.2 \text{ kJ mol}^{-1}$  for neutral molecules.

In order to allow for efficient optimization of a solute structure in the presence of the dielectric continuum, analytical derivatives of the COSMO energy have to be considered. For the implementation of energy gradients the mathematical simplicity of the COSMO model is very beneficial. Due to the fully variational treatment of the continuum polarization in HF and DFT calculations, it is only necessary to calculate the derivatives of the COSMO operators, *i.e.* of the matrix  $\mathbf{A}$  and of the “one-electron” integrals for the solute potential on the segments. The formal expression for the gradient simply reads

$$\nabla E_{\text{COSMO}} = \frac{1}{2f(\epsilon)} \mathbf{q} \nabla \mathbf{A} \mathbf{q} + \mathbf{q} \nabla \Phi^X \quad (5)$$

The gradient of the outlying charge correction apparently results in only a small contribution to the total value of the gradient and is thus neglected.

### 3. Demonstrative applications

#### 3.1. Computational details

The structures of all molecules and transition states have been fully optimized for both the gas phase and the condensed

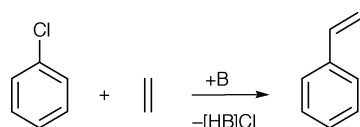
phase using DFT with an approximate treatment of the electronic Coulomb interaction (resolution of identity, RI-DFT).<sup>20</sup> A combination of the exchange functional of Becke<sup>21</sup> with the correlation functional of Perdew,<sup>22</sup> commonly designated as B-P functional, has been employed throughout. The Pd complexes in Section 3.1 have been optimized using split-valence (SV) basis sets<sup>23</sup> with one set of polarization functions for non-H atoms.<sup>24</sup> For the Pd atom, an effective core potential including relativistic corrections<sup>25</sup> for the description of the 28 innermost electrons and a corresponding optimized basis set of SV quality have been applied.<sup>26</sup> After the structure optimization, a single energy calculation with basis sets of triple-zeta valence (TZV) quality<sup>26,27</sup> and polarization functions at all atoms<sup>24</sup> has been performed. For the organic molecules in Sections 3.3 and 3.4, TZVP basis sets have been employed already for the structure optimization.

The calculations in solution have been carried out with the default COSMO parameters of the TURBOMOLE/COSMO implementation, *i.e.* dielectric constant  $\epsilon = \infty$ , number of points per atom in the cavity construction nppa = 1082, number of segments (groups of points) per atom nspa = 92, distance threshold for elements of matrix  $A$  disex = 10.0 Å, distance to outer solvent sphere rsolv = 1.3 Å, and distance of extra solvent sphere for outlying charge correction to outer solvent sphere routf = 0.85 Å (see refs. 2 and 19 for further explanation of the parameters). The optimized atomic COSMO radii ( $r_H = 1.3$  Å,  $r_C = 2.0$  Å,  $r_N = 1.83$  Å,  $r_O = 1.72$  Å,  $r_{Cl} = 2.05$  Å) of ref. 16 have been used, in combination with non-optimized radii of 2.16 Å and 2.28 Å for P and Pd, respectively. Since the latter two elements have four (P) and six (Pd) bond partners in the complexes considered in Section 3.2, they have no contribution to the SAS anyway.

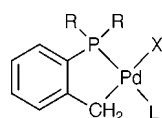
In combination with the COSMO-RS method, TURBOMOLE/COSMO is also capable of calculating all kinds of thermodynamic properties of multi-component mixtures in the liquid phase.<sup>28,29</sup> This is demonstrated by the calculation of Gibbs free energies and Boltzmann weights for a tautomeric equilibrium in Section 3.3, and the phase diagram and the activity coefficients of a binary mixture in Section 3.4. The program COSMOTHERM<sup>29</sup> has been used for the COSMO-RS calculations. It is able to read and process the TURBOMOLE/COSMO output files directly and has been parametrized and optimized for four different density functional/basis set combinations, namely B-P/SV(P), B-P/TZVP, B3-LYP/SV(P) and B3-LYP/TZVP.<sup>28</sup>

### 3.2. Solvent effect on the C–C coupling step in the Heck reaction catalysed by palladacycles

Homogeneous catalysis is one of the most successful fields for the application of density functional theory. However, most of the DFT investigations of reaction mechanisms are performed for the catalyst in the ideal gas phase, thus neglecting the influence of the reaction medium (solvent or reactants). In this subsection we discuss solvent effects on the C–C coupling step in the Heck reaction<sup>30</sup> of chlorobenzene and ethylene,



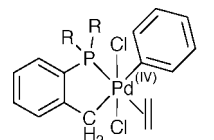
catalysed by palladacycles of the type



with  $R = o\text{-tolyl}$ ,  $X = \text{halide}$  and  $L = \text{solvent, olefin}$ . Recently, these catalysts have emerged as very stable, efficient and selective catalysts for Heck type reactions.<sup>31</sup> Despite the experi-

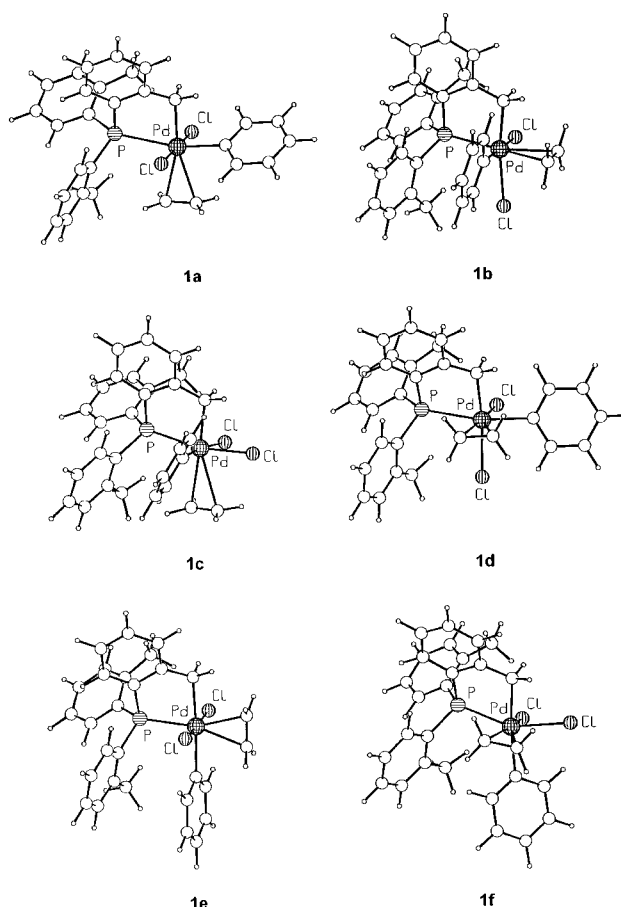
mental success of this class of catalysts, some details about the mechanism of the catalytic reaction are still not understood. There is an ongoing debate in the literature whether the reaction proceeds *via* Pd(II)/Pd(IV) species with the palladacycle staying intact,<sup>32</sup> or whether a reductive opening of the palladacycle takes place before the catalytic process starts, leading to palladium phosphine compounds as in traditional Heck catalysts.<sup>31</sup>

We have investigated the C–C bond formation step for the Pd(II)/Pd(IV) route, *i.e.* we have calculated structural isomers and activation barriers for complexes of the type



with  $R = o\text{-tolyl}$ . Fig. 1 shows the six calculated isomers of **1**. They all have a *cis* configuration of the phenyl and ethylene ligands, which is necessary for the C–C coupling reaction to take place. Table 1 contains the relative energies of these isomers in both the gas and liquid phases. From the gas-phase calculations, isomer **1a** clearly should be the dominant species. The isomers next higher in energy, **1b** and **1c**, are already 24–25 kJ mol<sup>−1</sup> above **1a**. Hence, in the gas phase, the *trans* configuration of the Cl ligands appears to be most favourable. Isomers **1e** and **1f** are much higher in energy due to the unfavourable *trans* position of the two Pd–C bonds.

The situation changes significantly for the solution phase (assuming a high dielectric constant), which can be deduced from the COSMO results in Table 1. Here, isomers **1b** and **1c** are most stable, and **1a** is 7 kJ mol<sup>−1</sup> higher in energy. Also structure **1d** is strongly stabilized and has only half of the



**Fig. 1** Calculated structures of six isomers of the palladacycle complex (see text), ordered according to the calculated gas-phase energy.

**Table 1** Absolute energies ( $E_h$ ) and energy differences to the most stable structure (in  $\text{kJ mol}^{-1}$ ) for the isomers of the paladacycle complex **1** (see text) as obtained from the RI-DFT(B-P)/TZVP//RI-DFT(B-P)/SV(P) calculations for the gas phase and the solution. For isomers **1a–1d**, calculated activation barriers (in  $\text{kJ mol}^{-1}$ ) for the C–C bond formation are also given

	Absolute energies/ $E_h$		Relative energies/ $\text{kJ mol}^{-1}$		Activation barriers/ $\text{kJ mol}^{-1}$	
	$E_{\text{Gas}}$	$E_{\text{COSMO}}$	$E_{\text{Gas}}$	$E_{\text{COSMO}}$	$E_{\text{Gas}}$	$E_{\text{COSMO}}$
<b>1a</b>	–2512.89568	–2512.91752	0.0	7.3	75.6	75.9
<b>1b</b>	–2512.88670	–2512.92031	23.6	0.0	37.8	34.9
<b>1c</b>	–2512.88596	–2512.92016	25.5	0.4	70.4	75.5
<b>1d</b>	–2512.87981	–2512.91221	41.7	21.3	19.7	11.8
<b>1e</b>	–2512.87395	–2512.89410	57.1	68.8		
<b>1f</b>	–2512.85264	–2512.88431	113.0	94.5		

energy difference to the most stable isomer in the solution phase as compared to the gas phase (21 *vs.* 42  $\text{kJ mol}^{-1}$ ). Apparently, the dielectric continuum stabilizes most the structures with *cis*-coordinated Cl ligands, *i.e.* isomers **1b**, **1c**, **1d** and **1f**, since their higher dipole moments cause stronger polarization. Thus, depending on the solvent, one can expect different species to dominate the catalytic reaction, which then can have a significant influence on the kinetics of the reaction.

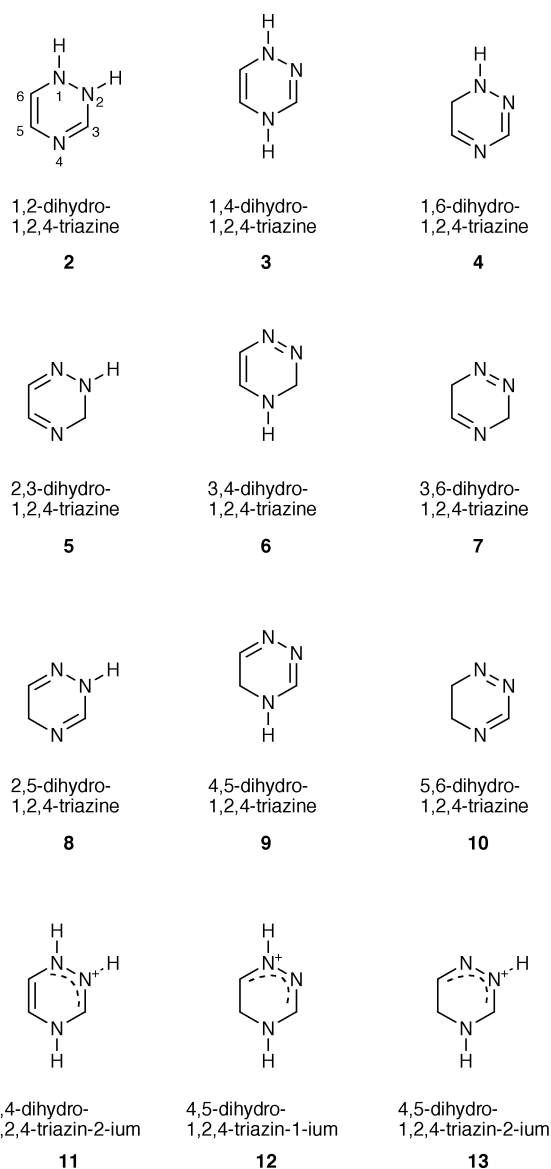
To investigate the latter aspect, we have also determined the activation barriers for the C–C coupling for the four most stable isomers, **1a–1d**, in gas phase and solution, respectively, which are also given in Table 1. The activation energies change by at most 8  $\text{kJ mol}^{-1}$  when going from the gas phase to the solution. In both phases, the barriers are rather high (70  $\text{kJ mol}^{-1}$  or more) for the isomers that have ethylene bound *trans* to the Pd–C bond (**1a** and **1c**). Isomers **1b** and **1d** show much lower barriers (38 and 20  $\text{kJ mol}^{-1}$ , resp., in the gas phase), so that, for the gas phase, the transition state formed from **1b** is on an absolute energy scale 14  $\text{kJ mol}^{-1}$  lower than the one formed from **1a**, although **1b** is 24  $\text{kJ mol}^{-1}$  above **1a**. In solution, isomer **1b** is lowest in energy, and the activation energy for **1b** is further reduced compared to the gas phase. Therefore, it can be concluded that, in polar solvents, isomer **1b** should be the relevant species for the catalytic mechanism of the Pd(II)/Pd(IV) type. Furthermore, polar solvents should significantly increase the reaction rates, since then the energetically most stable isomer also has a very low activation barrier.

### 3.3. Relative stabilities of dihydro-1,2,4-triazines

Dihydro-1,2,4-triazines are well known experimentally and their derivatives have found a broad range of applications in organic chemistry and life sciences.<sup>33–35</sup> Their structures have been studied and characterized by  $^1\text{H}$  NMR<sup>36</sup> and UV spectroscopy<sup>37</sup> as well as *via* preparative considerations.<sup>36–39</sup> However, the experimental work on these compounds left some uncertainty about their detailed geometric and electronic structure. Recently, a theoretical study on the relative stabilities of the dihydro-1,2,4-triazines and the dihydro-1,2,4-triazinium cations based on density functional theory and high level *ab-initio* quantum chemical calculations has been published.<sup>40</sup> So far, all of the theoretical calculations done on these molecules have been restricted to the gas phase, although the experimental results indicate a significant contribution of solvent effects on the relative stability of the dihydro-1,2,4-triazines (refs. 33–35, also compare ref. 40). This work attempts to clarify further the tautomerism of the dihydro-1,2,4-triazines by extending the theoretical approach to the liquid phase.

Fig. 2 shows the nine possible isomers of dihydro-1,2,4-triazine as well as the three possible structures of the dihydro-1,2,4-triazinium cation. In the first preparative articles, the 2,5-dihydro (**8**) and the 4,5-dihydro (**9**) isomers have been reported as reaction products.<sup>41,42</sup> In addition, UV spectroscopic measurements of 1,2,4-triazines with aryl substituents in the 3, 5

and 6 position of the heterocyclic ring indicate a tautomeric equilibrium between **8** and **9** in ethanol, consisting of predominantly 2,5-dihydro isomer **8**.<sup>37,39</sup> If a substituent is present at the N-4 atom, the most stable species is the 4,5-dihydro isomer,<sup>37</sup> whereas the 1,6-dihydro isomer is preferred if a substituent is present at N-1.<sup>39</sup> Thus, for the unsubstituted dihydro-1,2,4-triazine a tautomeric equilibrium between mainly the 2,5-dihydro- (**8**), 4,5-dihydro (**9**) and 1,6-dihydro (**4**) isomers can be expected.<sup>40</sup> Only the unsubstituted dihydro-1,2,4-triazines and a selection of dihydro-1,2,4-triazinium cations are considered in our calculations, in order to allow



**Fig. 2** Structures of nine possible dihydro-1,2,4-triazines and three possible dihydro-1,2,4-triazinium cations.



for an unambiguous comparison with the gas-phase stabilities of these molecules calculated by Nagy *et al.*<sup>40</sup>

Table 2 shows the optimized gas-phase and COSMO energies for the nine dihydro-1,2,4-triazines and three dihydro-1,2,4-triazinium cations. The relative gas-phase energies are in good agreement with the results of Nagy *et al.* that were calculated on a similar level of density functional theory.<sup>40</sup> The 2,5-dihydro isomer **8** was found to be most stable, followed by the 1,6- and the 4,5-dihydro isomers (**4** and **9** resp.), which are about 6.9 and 8.6 kcal mol<sup>−1</sup> higher in energy. The 1,4- and 2,3-dihydro isomers (**3** and **5**) are energetically comparable to **9**, whereas all other isomers are much higher in energy. This situation changes significantly with the application of COSMO: As expected, the 2,5-dihydro isomer **8** is still lowest in energy, and **2**, **7** and **10** are still the least stable structures, with no significant change in energy relative to **8**. However, the order in energy as well as the energy differences of the other isomers have changed. Isomers **3** and **9** are strongly stabilized, whereas the relative energy of the 1,6-dihydro isomer **4** has increased. In addition, isomers **5** and **6** have changed places on the energy scale. From the COSMO calculations, **9** is second lowest in energy, which is consistent with the experimental UV spectroscopic results mentioned above.

Regarding the cations, 4,5-dihydro-1,2,4-triazin-2-ium **13** is found to be the most stable isomer *in vacuo*, in agreement with the results of Nagy *et al.*<sup>40</sup> and experimental findings.<sup>37,39</sup> Upon the application of COSMO, the relative energies do not change much compared to the gas-phase results. The energy difference between 4,5-dihydro-1,2,4-triazin-1-ium **12** and **13** is lowered by 0.7 kcal mol<sup>−1</sup> whereas 1,4-dihydro-1,2,4-triazin-2-ium **11** is destabilized by 1.3 kcal mol<sup>−1</sup>.

The relative stability of the isomers in different solvents can be rationalized *via* their dielectric energy (defined as half of the

solvent–solute interaction energy,  $E_{\text{int}}/2$ ), which is shown in the third column of Table 3. The 4,5-dihydro isomer **9** has the lowest dielectric energy of the nine dihydro-1,2,4-triazines. This is also reflected in the significant energy gain in the COSMO calculation compared to the gas phase, which is also found for the other molecules with a strongly negative dielectric energy, *i.e.* **3**, **6** and cation **12**. In addition, the dielectric energy allows for a qualitative prediction of the stabilities of the isomers in solvents of different polarity. In general, a strongly negative dielectric energy will stabilize the solute molecule in polar solvents and destabilize it in non-polar solvents. Thus, in a polar solvent, the 4,5- and possibly also the 1,4- and 3,4-dihydro isomers **9**, **6** and **3** should be the predominant species besides **8**, whereas in a non-polar solvent **4** and **5** should be more stable. Isomers **2**, **7** and **10** are far too high in energy to make a significant contribution in the tautomeric equilibrium, independent of the solvent.

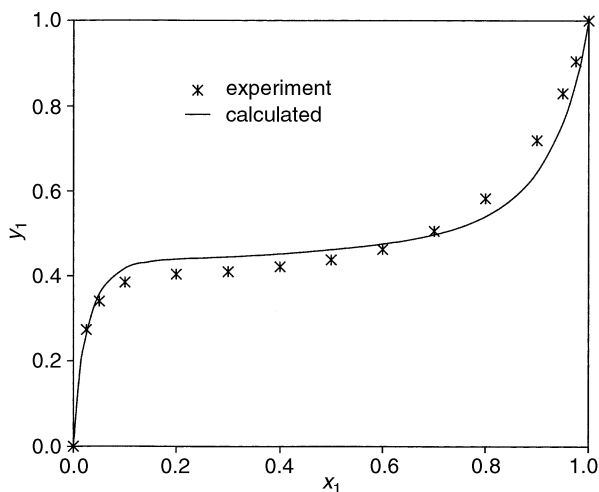
Such qualitative considerations can be turned into quantitative estimates using the COSMO-RS approach, which is a generalization of the COSMO model towards the realistic and physically correct representation of solvation effects by means of a rigorous statistical thermodynamic treatment of the solute–solvent interaction. Using the program COSMO-THERM, the tautomeric equilibria have been determined in a self-consistent way, *i.e.* the relative Boltzmann weights of the tautomers in the mixture have been allowed to relax until self-consistency of their chemical potentials has been reached.<sup>28</sup> Table 3 shows the calculated Gibbs free energies and Boltzmann weight factors for the nine isomers of dihydro-1,2,4-triazine and the three isomers of the dihydro-1,2,4-triazinium ion at infinite dilution in various solvents. The qualitative predictions from the dielectric energies are well met: in the non-polar solvent cyclohexane, 2,5-dihydro-1,2,4-triazine (**8**) is the

**Table 2** Optimized TURBOMOLE RI-DFT (B-P)/TZVP gas-phase and COSMO energies of nine dihydro-1,2,4-triazine isomers and three dihydro-1,2,4-triazinium cations. Values of Nagy *et al.* (CBS4 method including free energy correction) are taken from ref. 40

		Absolute energies/ $E_{\text{h}}$		Relative energies/kcal mol <sup>−1</sup>		
		$E_{\text{Gas}}$	$E_{\text{COSMO}}$	Nagy <i>et al.</i>	$E_{\text{Gas}}$	$E_{\text{COSMO}}$
1,2-dihydro-1,2,4-triazine	<b>2</b>	−281.622966	−281.638251	14.48	15.03	14.96
1,4-dihydro-1,2,4-triazine	<b>3</b>	−281.632861	−281.651203	9.46	8.82	6.83
1,6-dihydro-1,2,4-triazine	<b>4</b>	−281.635927	−281.650621	9.33	6.90	7.20
2,3-dihydro-1,2,4-triazine	<b>5</b>	−281.632704	−281.644334	10.39	8.92	11.14
3,4-dihydro-1,2,4-triazine	<b>6</b>	−281.628723	−281.648088	15.13	11.42	8.79
3,6-dihydro-1,2,4-triazine	<b>7</b>	−281.622376	−281.636781	16.08	15.40	15.88
2,5-dihydro-1,2,4-triazine	<b>8</b>	−281.646918	−281.662090	0	0	0
4,5-dihydro-1,2,4-triazine	<b>9</b>	−281.633158	−281.656743	8.89	8.63	3.35
5,6-dihydro-1,2,4-triazine	<b>10</b>	−281.615725	−281.628804	18.39	19.57	20.89
1,4-dihydro-1,2,4-triazin-2-ium	<b>11</b>	−281.997270	−282.098384	10.71	11.38	12.63
4,5-dihydro-1,2,4-triazin-1-ium	<b>12</b>	−282.004946	−282.109233	8.95	6.56	5.82
4,5-dihydro-1,2,4-triazin-2-ium	<b>13</b>	−282.015399	−282.118512	0	0	0

**Table 3** COSMO-RS dielectric energies, Gibbs free energies and Boltzmann weights of nine dihydro-1,2,4-triazine isomers and three dihydro-1,2,4-triazinium cations in various solvents

		$E_{\text{dielectric}}/\text{kcal mol}^{-1}$	Relative free energy $G_{\text{rel}}$ of the isomers in various solvents/kcal mol <sup>−1</sup>					Boltzmann weight $W_{\text{b}}$ of the isomers in various solvents				
			Cyclohexane	Chloroform	Ethanol	Acetic acid	Water	Cyclohexane	Chloroform	Ethanol	Acetic acid	Water
1,2-dihydro-1,2,4-triazine	<b>2</b>	−12.71	14.76	14.63	13.94	13.91	14.18	0	0	0	0	0
1,4-dihydro-1,2,4-triazine	<b>3</b>	−14.38	8.14	7.65	7.18	7.25	6.93	0	0	0.00001	0.00001	0.00001
1,6-dihydro-1,2,4-triazine	<b>4</b>	−11.72	6.40	6.76	7.45	7.66	7.69	0.00002	0.00001	0	0	0
2,3-dihydro-1,2,4-triazine	<b>5</b>	−10.07	9.86	10.35	10.92	11.15	11.41	0	0	0	0	0
3,4-dihydro-1,2,4-triazine	<b>6</b>	−17.19	10.71	9.63	9.10	8.90	8.50	0	0	0	0	0
3,6-dihydro-1,2,4-triazine	<b>7</b>	−11.48	15.39	15.58	18.17	17.70	17.70	0	0	0	0	0
2,5-dihydro-1,2,4-triazine	<b>8</b>	−11.63	0	0	0	0	0	0.99997	0.99967	0.99884	0.99070	0.98516
4,5-dihydro-1,2,4-triazine	<b>9</b>	−21.13	7.09	4.74	3.97	2.73	2.45	0.00001	0.00031	0.00115	0.00929	0.01483
5,6-dihydro-1,2,4-triazine	<b>10</b>	−9.97	19.59	20.09	22.87	22.62	22.68	0	0	0	0	0
1,4-dihydro-1,2,4-triazin-2-ium	<b>11</b>	−64.52	12.36	12.44	9.79	10.98	10.42	0	0	0	0	0
4,5-dihydro-1,2,4-triazin-1-ium	<b>12</b>	−66.80	6.60	6.33	6.05	5.86	5.69	0.00002	0.00003	0.00004	0.00006	0.00007
4,5-dihydro-1,2,4-triazin-2-ium	<b>13</b>	−65.74	0	0	0	0	0	0.99998	0.99997	0.99996	0.99994	0.99993



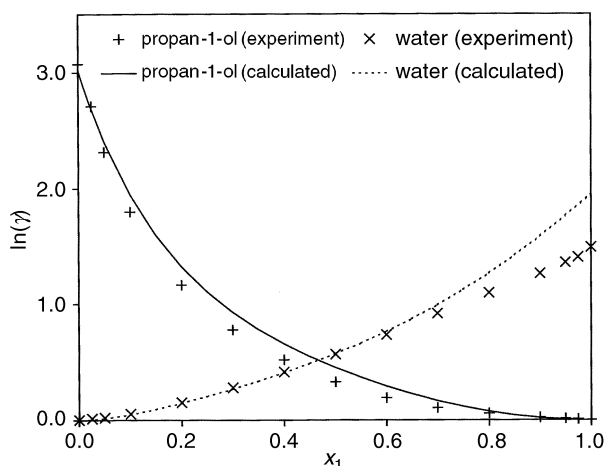
**Fig. 3** Phase diagram of a binary mixture of propan-1-ol (1) and water (2) at  $T = 333.15$  K;  $x_i$  = mole fraction in the liquid phase;  $y_i$  = mole fraction in the gas phase. Experimental values are taken from ref. 44.

predominant isomer, with a small amount of the 1,6-dihydro isomer **4** in the mixture. Proceeding to more polar solvents, the amount of isomer **8** (having a small dielectric energy) in the mixture decreases and isomer **4** completely vanishes. On the other hand, the free energy gain of isomers **9**, **6** and **3** is quite strong for polar solvents, resulting in an increased share (i.e. increased Boltzmann factor  $W_B$ ) in the equilibrium of the tautomers. Based on UV spectroscopic measurements in ethanol and acetic acid, Nagy *et al.*<sup>39,40</sup> have assumed a tautomeric equilibrium between predominantly isomer **8** and a small amount of isomer **9**. This result is reproduced by the present calculations, as is apparent from Table 3. The effect of different solvents on the dihydro-1,2,4-triazinium isomers is smaller but still visible: In polar solvents, the weight of isomer **12** is increased.

### 3.4. Phase diagrams of binary mixtures of liquids

The phase diagram and the activity coefficients of a binary mixture of propan-1-ol with water at  $T = 333.15$  K have been calculated. In the computation of the gas-phase compositions, experimental vapour pressures of the pure compounds (propan-1-ol 192.0 mbar, water 198.7 mbar<sup>43</sup>) have been used.

The results for the binary liquid *vs.* gas-phase composition diagram are presented in Fig. 3. The RI-DFT(B-P)/COSMO/COSMO-RS values are in good agreement with the experi-



**Fig. 4** Activity coefficients of the binary mixture of propan-1-ol (1) and water (2) at  $T = 333.15$  K. Experimental values are calculated from the thermodynamic data of Schreiber *et al.*<sup>44</sup>

mental results of Schreiber *et al.*<sup>44</sup> In particular, the flat region of the phase curve, between  $x_1 = 0.2$  and  $x_1 = 0.7$ , is represented very well, showing only small deviations of less than 10% from the experiment. In addition, the absence of a miscibility gap for the propanol–water system at  $T = 333.15$  K is predicted correctly. Fig. 4 shows the activity coefficients of propan-1-ol and water *vs.* the mole fraction of propan-1-ol at  $T = 333.15$  K. Apparently the correspondence between the predicted values and the experimental data of Schreiber *et al.*<sup>44</sup> is very good.

## 4. Conclusions

It has been shown that the DFT/COSMO method and its extension, the COSMO-RS approach, provide a valuable (in some cases even indispensable) tool for the theoretical rationalization and prediction of chemistry in solution. As demonstrated by the given examples, the TURBOMOLE/COSMO implementation allows a computationally efficient description of solvation effects, which can be assessed quantitatively using the COSMO-RS extension, and additionally provides a vivid picture of the solute–solvent interactions in the liquid phase.

However, the range of applications is not restricted to chemical and engineering thermodynamics. For example, TURBOMOLE/COSMO/COSMOTHERM can also provide valuable descriptors for QSAR applications in the life sciences and related fields of research<sup>28</sup> (see also [http://www.cosmologic.de/ct\\_examples.htm](http://www.cosmologic.de/ct_examples.htm)).

## References

- 1 R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chem. Phys. Lett.*, 1989, **162**, 165; R. Ahlrichs and M. v. Arnim, in *Methods and Techniques in Computational Chemistry: METECC-95*, ed. E. Clementi and G. Corongiu, STEF, Cagliari, 1995, p. 509ff.
- 2 A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2*, 1993, 799.
- 3 C. J. Cramer and D. G. Truhlar, *Chem. Rev.*, 1999, **99**, 2161.
- 4 J.-L. Rivail and D. Rinaldi, in *Computational Chemistry—Reviews of Current Trends*, ed. J. Leszczynski, World Scientific, Singapore, 1995, vol. 33, pp. 139–174.
- 5 J. Tomasi and M. Persico, *Chem. Rev.*, 1994, **94**, 2027.
- 6 A. Gelessus, PhD Thesis, University of Zürich, 1997.
- 7 A. J. Holder, program package AMPAC 6.0, 1997.
- 8 J. Andzelm, C. Kölmel and A. Klamt, *J. Chem. Phys.*, 1995, **103**, 9312.
- 9 T. N. Truong and E. V. Stefanovich, *Chem. Phys. Lett.*, 1995, **240**, 253.
- 10 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995.
- 11 K. Baldrige and A. Klamt, *J. Chem. Phys.*, 1997, **106**, 6622.
- 12 C. C. Pye and T. Ziegler, *Theor. Chem. Acc.*, 1999, **101**, 396.
- 13 D. M. Dolney, G. D. Hawkins, P. Winget, D. A. Liotard, C. J. Cramer and D. G. Truhlar, *J. Comput. Chem.*, in the press.
- 14 S. Miertus, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117.
- 15 A. Klamt, *J. Phys. Chem.*, 1995, **99**, 2224.
- 16 A. Klamt, V. Jonas, T. Bürger and J. C. W. Lohrenz, *J. Phys. Chem. A*, 1998, **102**, 5074.
- 17 A. Klamt, in *Encyclopedia of Computational Chemistry*, ed. P. v. R. Schleyer, Wiley, New York, 1998, pp. 604–615.
- 18 W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in FORTRAN*, Cambridge University Press, New York, 2nd edn., 1992, pp. 89–91.
- 19 A. Klamt and V. Jonas, *J. Chem. Phys.*, 1996, **105**, 9972.
- 20 K. Eichkorn, O. Treutler, H. Öhm, M. Häser and R. Ahlrichs, *Chem. Phys. Lett.*, 1995, **242**, 652.
- 21 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098.
- 22 (a) J. P. Perdew, *Phys. Rev. B*, 1986, **33**, 8822; (b) J. P. Perdew, *Phys. Rev. B*, 1986, **34**, 7406(E).
- 23 A. Schäfer, H. Horn and R. Ahlrichs, *J. Chem. Phys.*, 1992, **97**, 2571.
- 24 The following exponents of polarization functions have been used: H p = 0.8, C d = 0.8, N d = 1.0, O d = 1.2, P d = 0.45, Cl d = 0.65.
- 25 D. Andrae, U. Haeussermann, M. Dolg, H. Stoll and H. Preuss, *Theor. Chim. Acta*, 1990, **77**, 123.

- 26 K. Eichkorn, F. Weigend, O. Treutler and R. Ahlrichs, *Theor. Chim. Acta*, 1997, **97**, 119.
- 27 A. Schäfer, C. Huber and R. Ahlrichs, *J. Chem. Phys.*, 1994, **100**, 5829.
- 28 A. Klamt and F. Eckert, *Fluid Phase Equilibria*, 2000, in the press.
- 29 A. Klamt and F. Eckert, COSMOTHERM, version C1.1, revision 12/99; COSMOlogic GmbH & Co. KG, Leverkusen, Germany, 1999. See also <http://www.cosmologic.de>
- 30 W. A. Herrmann, in *Applied Homogeneous Catalysis with Organometallic Compounds*, ed. B. Cornils and W. A. Herrmann, VCH, Weinheim, 1996, vol. 2, pp. 712–732.
- 31 W. A. Herrmann, V. P. W. Böhm and C.-P. Reissinger, *J. Organomet. Chem.*, 1999, **576**, 23.
- 32 B. L. Shaw, *New J. Chem.*, 1998, 77.
- 33 V. N. Charushin, S. G. Alexeev, O. N. Chupahkin and H. C. van der Plas, in *Advances in Heterocyclic Chemistry*, ed. R. A. Katritzky, Academic Press, San Diego, 1989, vol. 46, pp. 74–135.
- 34 A. L. Weis, in *Advances in Heterocyclic Chemistry*, ed. R. A. Katritzky, Academic Press, San Diego, 1985, vol. 38, pp. 1–103.
- 35 H. Neunhöfer, in *The Chemistry of Heterocyclic Compounds*, ed. A. Weissberger and E. C. Taylor, Wiley Interscience, New York, 1978, vol. 33, pp. 189–1072.
- 36 S. Konno, S. Ohba, M. Sagi and H. Yamanaka, *Chem. Pharm. Bull.*, 1987, **35**, 1378.
- 37 J. Nagy, A. Horváth, Á. Szöllösy and J. Nyitrai, *Eur. J. Org. Chem.*, 1999, 685.
- 38 S. Konno, M. Sagi, Y. Yuki and H. Yamanaka, *Heterocycles*, 1985, **23**, 2807.
- 39 J. Nagy, R. Rapp, M. Alexovics, D. Döpp and J. Nyitrai, *J. Chem. Soc., Perkin Trans. 1*, 1993, 661.
- 40 J. Nagy, J. Nyitrai, I. Vágó and G. I. Csonka, *J. Org. Chem.*, 1998, **63**, 5824.
- 41 R. Metze, *Chem. Ber.*, 1958, **91**, 1863.
- 42 C. M. Atkinson and H. D. Cossey, *J. Chem. Soc.*, 1962, 1805.
- 43 *Handbook of Chemistry and Physics*, ed. R. C. Weast, CRC Press, Boca Raton, FL, 1976.
- 44 E. Schreiber, E. Schüttau, D. Rant and H. Schuberth, *Z. Phys. Chem. (Leipzig)*, 1976, **247**, 23.