
24

DRUG SOLUBILITY AND REACTION THERMODYNAMICS

KARIN WICHMANN

COSMOlogic GmbH & Co. KG, Leverkusen, Germany

ANDREAS KLAMT

COSMOlogic GmbH & Co. KG, Leverkusen, Germany, and Institute of Physical and Theoretical Chemistry, University of Regensburg, Regensburg, Germany

24.1 INTRODUCTION

24.1.1 Methods for Compounds in Solution

There is a variety of computational methods for the treatment of compounds in solution. The scope of this chapter is not to give an overview of them, but to concentrate on applications of COSMO-RS, a young and very efficient method for the *a priori* prediction of thermophysical data of liquids. COSMO-RS combines unimolecular quantum chemical calculations that provide the necessary information for the evaluation of molecular interaction in the fluid phase, with a very fast and accurate statistical thermodynamic procedure. It has established itself as an alternative to structure-based group contribution methods (GCMs) on the one hand and to force field-based simulation methods on the other hand. Because of its special approach, COSMO-RS is a generally applicable method for compounds in solution. It has been applied successfully in such diverse areas as solvent screening, partitioning behavior, liquid–liquid and vapor–liquid equilibria, and ADME property prediction, and for such diverse compound types as drugs, pesticides, common organic compounds, halocarbons, and ionic liquids. COSMO-RS is used in chemical, pharmaceutical, agrochemical and petrochemical industry.

In this contribution, two application fields important in drug development and drug production will be considered: solubility prediction and prediction of free energy of reaction in solution. Solubility prediction methods are important during the drug design and development process, because in the early drug design phase compounds are often only

virtually considered by computational drug design methods, or the synthesized amount of substance is insufficient for experiments. In both cases the only tools for the selection of promising drug candidates with adequate solubility are computational methods that predict the solubility with sufficient accuracy just from the chemical structure of the compound. A method requiring experimental data for solubility prediction is unfeasible in this situation, since such data will not be available.

Prediction of thermodynamic properties of compounds in solution is also important in industrial process development. Here, specifically reaction energies and equilibrium constants of reaction in solution are of particular interest when a new process is developed or alternative pathways for existing processes are explored. The Gibbs free energy of reaction varies with the choice of the solvent or solvent mixture, and hence the chosen solvent system can strongly influence the process in solution. Generally, experimental data for the equilibrium constant or the free energy of reaction in solution are rare, but are available relatively straightforward from a computational approach. Prediction of thermochemical data like heat of reaction and heat of vaporization furthermore helps designing a chemical process such that process hazards can be prevented.

24.1.2 COSMO

In conventional quantum chemistry, molecules are treated as isolated particles at a temperature of 0K. In physical reality

however, the major part of reactions takes place in solution and at higher temperatures. Since direct treatment of a large number of molecules is computationally very demanding, solvent effects are often treated indirectly by continuum solvation models, where the solute is embedded in a dielectric continuum and the solvent is represented by a mean interaction with a surrounding dielectric medium. The interaction of the solute with such a dielectric solvent is taken into account in the quantum chemical calculation by polarization charges that arise from the dielectric boundary condition.

The “COnductor-like Screening MOdel” (COSMO) is an efficient variant of dielectric continuum solvation methods [1]. In quantum chemical COSMO calculations the solute molecules are calculated in a virtual scaled conductor environment, that is, the scaled boundary condition of a conducting medium is used, where the molecule is ideally screened, and not the exact dielectric boundary condition. In such a conducting environment the solute molecule induces a polarization charge density σ on the interface between the molecule and the conductor, that is, on the molecular surface. These charges act back on the solute and generate a more polarized electron density than in vacuum. During the quantum chemical self-consistency algorithm, the solute molecule is thus converged to its energetically optimal state in a conductor with respect to electron density. Due to the analytic gradients available for the COSMO energy contributions, the molecular geometry can be optimized using conventional methods for calculations in vacuum. The quantum chemical calculation has to be performed once for each molecule of interest.

24.1.3 COSMO-RS

As discussed in more detail elsewhere the simple dielectric continuum models suffer from a number of insufficiencies [2,3]. The polarization charge density σ resulting from unscaled COSMO calculations (also called screening charge density σ), which is a good local descriptor of the molecular surface polarity, is used to extent the model toward “Real Solvents” (COSMO-RS).

In COSMO-RS, a liquid is considered to be an ensemble of closely packed ideally screened molecules, as shown in Figure 24.1. In this figure, each piece of surface has one direct contact partner, but is still separated from its partner by a thin film of conductor. Since the conducting medium that was assumed to surround the molecules in the COSMO calculation is not existent in reality, the energy difference between the pairwise contacts and the ideally screened situation has to be defined as a local electrostatic interaction energy that results from the removal of the conductor film between the molecules. Considering a contact on a region of molecular surface of area a_{eff} (effective contact area), and considering that the two contacting pieces of molecular surface have average ideal screening charge densities σ and σ' in the conductor, it is possible to calculate this interaction energy as the energy that is necessary to remove the residual screening charge density $\sigma + \sigma'$ from the contact. In the special case of $\sigma = -\sigma'$, the contact is an “ideal electrostatic contact” and the interaction energy is zero, because the two molecules screen each other as well as the conductor did. In the general case however $\sigma + \sigma'$ does not vanish and the arising electrostatic interaction energy is

$$E_{\text{misfit}}(\sigma, \sigma') = a_{\text{eff}} e_{\text{misfit}}(\sigma, \sigma') = a_{\text{eff}} \frac{\alpha'}{2} (\sigma + \sigma')^2, \quad (24.1)$$

where $e_{\text{misfit}}(\sigma, \sigma')$ is the misfit energy density on the contact surface and α' is a general constant that can be calculated approximately, but in COSMO-RS is fitted to experimental data as fine-tuning. The misfit term (equation 24.1) subsumes the polarization response of the molecules to the electrostatic misfit quite well [4, 5].

Hydrogen bonding interactions are to some extent already covered by the description of electrostatic interactions, but we still have to parameterize the additional hydrogen bonding energy resulting from interpenetration of the atomic electron densities in some reasonable way. This energy should only be relevant if two sufficiently polar pieces of surface of opposite polarity are in contact, and it should be the more important, the more polar both surface pieces are. Taking the screening charge density σ as a local

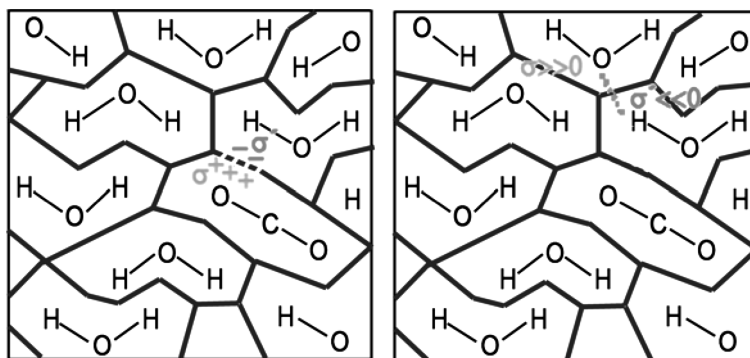


FIGURE 24.1 Schematic illustration of contacting molecular cavities and contact interactions.

measure of polarity, the following function realizes such behavior:

$$E_{\text{hb}}(\sigma, \sigma') = a_{\text{eff}} e_{\text{hb}}(\sigma, \sigma') \\ = a_{\text{eff}} c_{\text{hb}} \min\{0, \min(0, \sigma_{\text{don}} \\ + \sigma_{\text{hb}}) \max(0, \sigma_{\text{acc}} - \sigma_{\text{hb}})\} \quad (24.2)$$

with $\sigma_{\text{don}} = \min(\sigma, \sigma')$ and $\sigma_{\text{acc}} = \max(\sigma, \sigma')$. General parameters are σ_{hb} , the threshold for hydrogen bonding, and c_{hb} , the coefficient for the hydrogen bond strength. Both parameters have to be adjusted to experimental data. With equation 24.2, the hydrogen bond interaction energy is zero, unless the more negative of the two screening charge densities is less than the threshold $-\sigma_{\text{hb}}$, and the more positive exceeds σ_{hb} . Because positive molecular regions have negative screening charge, the negative σ now is the donor part of the hydrogen bond and the positive is the acceptor. In this case, the hydrogen bonding energy is proportional to the product of the excess screening charge densities $(\sigma_{\text{don}} + \sigma_{\text{hb}})(\sigma_{\text{acc}} - \sigma_{\text{hb}})$.

van der Waals interactions are described by element-specific parameters τ in COSMO-RS. The τ parameters have to be fitted to experimental data. Then, the vdW energy gain of a molecule X during the transfer from the gas phase to any solvent is given by

$$E_{\text{vdW}}^X = \sum_{\alpha \in X} a_{\alpha}^X \tau_{\text{vdW}}(e(\alpha)) \quad (24.3)$$

The vdW term is spatially nonspecific. Because E_{vdW} is independent of any neighboring relations, it is not really an interaction energy, but may be considered as an additional contribution to the energy of the reference state in solution. Currently nine of the vdW parameters (for elements H, C, N, O, F, S, Cl, Br, and I) have been optimized. For the majority of the remaining elements reasonable estimates are available. Nonadditive vdW corrections are used for a few element

pairs, but they are of minor importance for the topics of this contribution.

The transition from microscopic surface interaction energies to macroscopic thermodynamic properties of a liquid is possible via a statistical thermodynamics procedure. The exact solution of the thermodynamic problem would require sampling of all different arrangements of all molecules of the systems, weighting the contribution of each arrangement by its Boltzmann factor. This direct approach, which is used in the molecular dynamics and Monte Carlo type methods, is very time-consuming and requires compromises regarding sampling and regarding the accuracy of the energy evaluations. COSMO-RS follows a different concept. The basic approximation is that the ensemble of interacting molecules may be replaced by the corresponding ensemble of independent, pairwise, interacting surface pieces. This approximation implies the neglect of any neighborhood information of surface pieces on the molecular surface and the loss of steric information. The advantage of this approximation is the extreme reduction of the complexity of the problem, which allows for a fast and exact solution. It should be noted that GCMs as UNIFAC are also based on the assumption of independent pairwise interacting surfaces.

Since the screening charge density σ is the only descriptor determining the interaction energy terms in equations 24.1 and 24.2, the ensemble of surface pieces characterizing an ensemble S is sufficiently described by its composition with respect to σ . For this purpose we introduce the molecular σ -profile $p^X(\sigma)$, which is a histogram of the screening charge densities σ on the surface of a molecule X (Figure 24.2). The σ -profile can easily be derived from the COSMO files produced as output of the quantum chemical COSMO calculation for molecule X , applying a local averaging algorithm in order to take into account that only screening charge densities σ averaged over an effective contact area are of physical meaning in COSMO-RS [5].

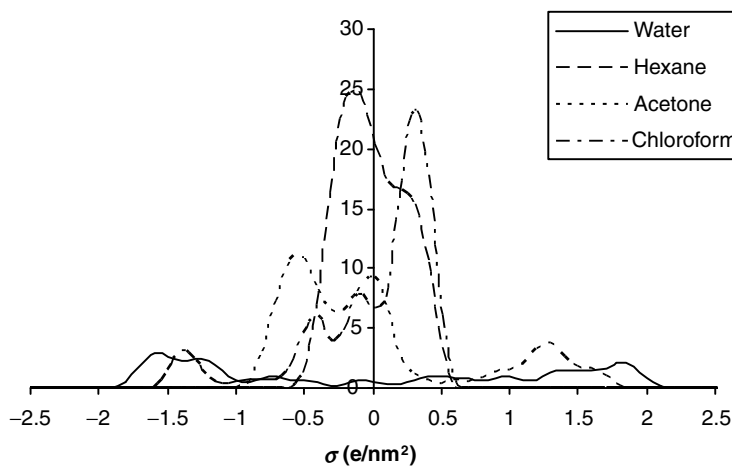


FIGURE 24.2 σ -profiles of common solvents.

The σ -profile for the entire solvent of interest S , which might be a mixture of several compounds, $p_S(\sigma)$, is given by the weighted, surface area normalized sum of the σ -profiles of the components X_i :

$$p'_S(\sigma) = p_S(\sigma)/A_S = \sum_{i \in S} x^i p^{X_i}(\sigma) / \sum_{i \in S} x^i A^{X_i} \quad (24.4)$$

where A^{X_i} is the COSMO surface of a compound X_i in the system.

Under the condition that there is no free surface in the bulk of the liquid, that is, each piece of molecular surface has a direct contact partner, the statistical thermodynamics of the system can be solved using the exact equation:

$$\mu_S(\sigma) = -RT \ln \left\{ \int p'_S(\sigma') \exp \left(\frac{\mu_S(\sigma') - a_{\text{eff}} e(\sigma, \sigma')}{RT} \right) d\sigma' \right\} \quad (24.5)$$

In this equation, $\mu_S(\sigma)$ is the chemical potential of an average molecular contact segment of size a_{eff} in the ensemble S at temperature T , and $e(\sigma, \sigma')$ is the interaction energy functional $e(\sigma, \sigma') = e_{\text{misfit}}(\sigma, \sigma') + e_{\text{hb}}(\sigma, \sigma')$. Since $\mu_S(\sigma)$ appears on both sides of equation 24.5, it must be solved by iteration, starting with $\mu_S(\sigma') = 0$ on the right-hand side. Fortunately, the solution converges rapidly and $\mu_S(\sigma)$ can be computed up to numerical precision within milliseconds on a personal computer. For a formal derivation of equation 24.5, we refer to Ref. 4.

Now it is straightforward to define the chemical potential of a solute X in the ensemble S by

$$\mu_S^X = \mu_{\text{res},S}^X + \mu_{\text{comb},S}^X = a_{\text{eff}}^{-1} \int p^X(\sigma) \mu_S(\sigma) d\sigma + \mu_{\text{comb},S}^X \quad (24.6)$$

where the residual part, that is, the part resulting from the interactions of the surfaces in the liquid, is given by the surface integral of function $\mu_S(\sigma)$ over the solute surface, which is expressed using the σ -profile of the solute in equation 24.6. The second part is the combinatorial contribution, which arises from the different shapes and sizes of the solute and solvent molecules. Expressions based on the surface areas and volume ratios of solvents and solutes, similar to standard chemical engineering expressions as Staverman-Guggenheim, are used in the context of COSMO-RS [6]. COSMO surface areas and volumes are used for the evaluation of the combinatorial term. Hence, equation 24.6 can be completely evaluated based on the information resulting from the COSMO calculations of the individual compounds.

The chemical potential of equation 24.6 is a pseudochemical potential [7], that is, the standard chemical potential without the concentration term $RT \ln x_i$. We will shortly use the term chemical potential for the pseudochemical potential from equation 24.6 throughout this contribution. Providing the chemical potential of an arbitrary compound X in almost arbitrary solvents and mixtures as a function of temperature

and concentration, equation 24.6 allows for the prediction of almost all thermodynamic properties of compounds or mixtures, such as activity coefficients, partition coefficients, or solubility, as shown in the flowchart for a COSMO-RS property prediction in Figure 24.3.

As mentioned above, the COSMO-RS method depends on a small number of adjustable parameters. Some of the parameters are predetermined from physics, while others are determined from selected properties of mixtures. The parameters are not specific to functional groups or types of molecule. As a result, COSMO-RS is the least parameterized of all quantitative methods for the prediction of chemical properties in the liquid phase [8].

24.1.4 Treatment of Conformers in COSMO-RS

Many molecules can adopt more than one conformation, and different conformers of one molecule can have different σ -profiles. The chemical potentials of the individual conformers and hence the conformer distribution as well as the chemical potential of the compound represented by an ensemble of conformers depend on the composition of the system and the temperature. Thus, it is essential for property prediction with COSMO-RS to take conformers with different σ -profiles into account, each described by individual quantum chemical COSMO calculations. The relative contributions of the conformers are determined by an iterative procedure using the Boltzmann-weight of the free energies of the conformers in the liquid phase. This results in a thermodynamically fully consistent treatment of multiple molecular conformations.

24.2 SOLUBILITY PREDICTION WITH COSMO-RS

For the calculation of the solubility S_S^X of a liquid compound X in a solvent S we require the chemical potentials of X in S and in its pure liquid state, μ_S^X and μ_X^X . If S_S^X is sufficiently small, so that the solvent behavior of the X -saturated solvent S is not significantly influenced by the solute X , then the decadic logarithm of the solubility is given by

$$\log S_S^X = \log \left(\frac{\text{MW}^X \rho_S}{\text{MW}_S} \right) - \frac{\ln(10)}{kT} \Delta_S^X \quad (24.7)$$

with the molecular weight MW , the solvent density ρ , and $\Delta_S^X = \mu_S^X - \mu_X^X$. In the case of high solubility (typically for solubility greater than 10 wt%), equation 24.7 becomes approximate and the true solubility would have to be derived from a detailed search for a thermodynamic equilibrium of a solvent-rich and a solute-rich phase. But, in general, at least for the purpose of estimating drug solubility, equation 24.7 is sufficiently accurate.

If the zeroth order $S_S^{X_0}$ as initially provided by equation 24.7, using infinite dilution of X in S , is resubstituted

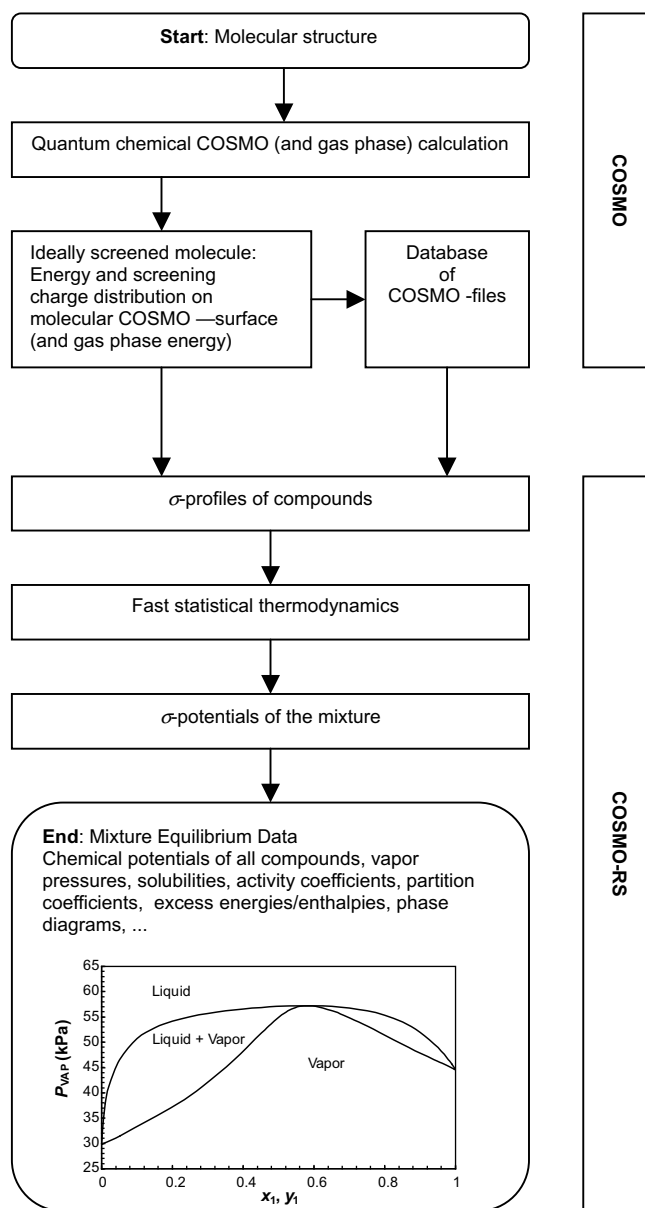


FIGURE 24.3 Flowchart of a property prediction procedure with COSMO and COSMO-RS.

into the solubility calculation via $\Delta_S^{X_1} = \mu_{S(X_0)}^X - \mu_X^X$ a better approximation for S_S^X is achieved. In other words, the solute chemical potential μ_S^X is computed for the solvent–solute mixture with the finite mole fraction of X in S that was predicted by the zeroth order S_S^X . Then, using equation 24.7 with the new $\mu_{S(X_0)}^X$ and the resulting values, an improved solubility $S_S^{X_1}$ is computed. Iterating this process to convergence an iterative solubility can be achieved, which is also implemented in our COSMO-RS program *COSMOtherm* and allows for the accurate prediction of solubility values even for cases of high solubility (solubility up to 50 wt%). Thus, except for rare cases of very high solubility, a complicated

search for a multiphase thermodynamic equilibrium of a solvent-rich and a solute-rich phase can be avoided, but instead equation 24.7 and its iterative refinement can be used.

Drugs are mostly solid at room temperature. Because the solid state of a compound X is related to its liquid state by the free energy difference ΔG_{fus}^X , which is negative in the case of solids, a more general expression for solubility reads

$$\log S_S^X = \log \left(\frac{MW^X \rho_S}{MW_S} \right) + \frac{\ln(10)}{kT} [-\Delta_S^X + \min(0, \Delta G_{\text{fus}}^X)] \quad (24.8)$$

For liquid compounds, ΔG_{fus}^X is positive and equation 24.8 reduces to equation 24.7.

If melting point temperature T_{melt} and heat of fusion ΔH_{fus} or entropy of fusion ΔS_{fus} are known experimentally for a solid compound, the free energy of fusion ΔG_{fus} in equation 24.8 can be estimated from

$$\Delta G_{\text{fus}}(T) = -\Delta H_{\text{fus}} \left(1 - \frac{T}{T_{\text{melt}}} \right) \quad (24.9)$$

or

$$\Delta G_{\text{fus}}(T) = -\Delta S_{\text{fus}}(T_{\text{melt}} - T) \quad (24.10)$$

Equations 24.9 and 24.10 can be complemented by an additional temperature-dependent term using the heat capacity of fusion $\Delta C p_{\text{fus}}$ in order to obtain good absolute predictions, but data for $\Delta C p_{\text{fus}}$ are rarely available from experiment.

The free energy of fusion of new compounds is often not known, because experimental measurements can be cumbersome and substance may be scarce. Computational prediction of ΔG_{fus}^X requires evaluation of the free energy of a molecule of compound X in its crystal, that is, the crystal structure needs to be known or predicted. In general however, crystal structure prediction for drugs has to be considered as an unsolved problem. Thus, there is no viable way to a fundamental model. As an alternative, a QSPR approximation for ΔG_{fus} can be used in *COSMOtherm*, which is based on a few rather obvious factors that should influence crystallization. Larger molecules should have larger ΔG_{fus} than smaller ones, compounds with more polarity and hydrogen bonding ability should have larger ΔG_{fus} than less polar ones, and also rigidity should give rise to larger ΔG_{fus} . We found that a good regression equation for ΔG_{fus}^X can be achieved by a combination of the descriptors V^X , the cavity volume from the COSMO calculation as size descriptor, $N_{\text{ring atom}}^X$, the number of ring atoms in X as a descriptor of the compounds rigidity, and μ_{W}^X , the compounds chemical potential in water as a combined measure of polarity and hydrogen bonding: [8]

$$\Delta G_{\text{fus}}^X = c_0 + c_1 \mu_{\text{W}}^X + c_2 N_{\text{ring atom}}^X + c_3 V^X \quad (24.11)$$

24.2.1 Relative Solubility and Solubility Screening

The computational prediction of the relative solubility of a drug candidate in a variety of solvents with COSMO-RS is straightforward and can be done without wasting any of the substance in this step. The required DFT/COSMO calculations can be done even before the compound comes to the development laboratory, and a COSMO-RS solubility screening can be already completed when the work in the development department starts.

Experimental data for melting point and free energy of fusion can often be obtained through differential scanning calorimetry. If melting point temperature and heat of fusion or entropy of fusion are known for the compound in question, the free energy of fusion ΔG_{fus} can be calculated according to equations 24.9 or 24.10 and a solubility screening for absolute solubilities can be done. If data for ΔG_{fus} are not known, an estimated ΔG_{fus} may be used, either from the QSPR model implemented in *COSMOtherm* or from an external model.

EXAMPLE 24.1 RELATIVE AND ABSOLUTE SOLUBILITY OF ACETAMINOPHEN IN PURE SOLVENTS AT 30°C

Experimental data for the solubility of acetaminophen in pure solvents were reported by Granberg and Rasmuson [9]. We use this data set to validate the calculated acetaminophen solubilities. Furthermore, a melting point temperature of $T_{\text{melt}} = 441.2\text{K}$, a heat of fusion of $\Delta H_{\text{fus}} = 26.0\text{kJ/mol}$ and an entropy of fusion of $\Delta S_{\text{fus}} = 59.0\text{J/mol}$ have been reported for acetaminophen [10]. These data will be used to compute absolute solubility predictions for acetaminophen in the solvent data set.

With the case study of acetaminophen solubility we want to show first the prediction of relative solubility. For relative solubility calculations, we do not make use of any experimental data like melting point temperature or enthalpy of

fusion, as that kind of data are usually not available in the early drug design phase.

Solute and solvents for this example were calculated on the BP86/TZVP level of theory, which corresponds to the higher quantum chemical level for which the *COSMOtherm* program is parameterized. Different conformations of the compounds were taken into account where the conformers showed different σ -profiles and COSMO energies. All compounds of the data set, including conformers, are available from the *COSMObase*, a collection of validated COSMO files for common compounds and solvents (Figure 24.4).

For the calculation of relative solubility in pure solvents, we do not use the iterative refinement procedure, since the assumed value of $\Delta G_{\text{fus}} = 0\text{kcal/mol}$ will influence the computed zeroth-order solubility and hence the iterative refinement. Therefore, the relative solubilities are calculated in infinite dilution in the respective solvent at 30°C (Figure 24.5).

Calculation results can be read from the output and table files (Figure 24.6). The relative solubility of a compound can be calculated from the chemical potentials of the compound in the solvent $\mu_i^{(j)}$ and in its pure state $\mu_i^{(P)}$ as

$$\log S_{\text{rel}}(x) = (\mu_i^{(P)} - \mu_i^{(j)}) / (RT \ln(10)) \quad (24.12)$$

Since a logarithmic solubility value larger than 0 indicates only that the two compounds are miscible, there is a cutoff at 0 in the *COSMOtherm* results for the logarithmic solubility in Figure 24.6. However, in order to provide insight in the whole range of the solvent data set independent of potential misestimates of ΔG_{fus} , positive values for $\log S_{\text{rel}}(x)$ are allowed here for both relative and absolute solubility predictions. Relative solubility data in Table 24.1 were calculated directly from the chemical potential differences as described in equation 24.12.

The relative solubility predictions correlate well with the experimental data, revealing an overall shift of 1.8 log units, which arises mainly from the neglect of ΔG_{fus} and is therefore irrelevant for real solubility considerations. The predicted

Set	No	COSMO-Name	CAS-Number	MW	Formula	use conf	Search
<input type="checkbox"/>	726	4-hydroxyacetanilide0	000103-90-2	151.16	C8H9NO2	<input checked="" type="checkbox"/>	2
<input type="checkbox"/>	0	h2o	007732-18-5	18.02	H2O	<input type="checkbox"/>	0
<input type="checkbox"/>	1	1-octanol0	000111-87-5	130.23	C8H18O	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	2	formaldehyde	000050-00-0	30.03	CH2O	<input type="checkbox"/>	0
<input type="checkbox"/>	3	dexamethasone0	000050-02-2	392.46	C22H29FO5	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	4	phenobarbital0	000050-06-6	232.24	C12H12N2O3	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	5	lacticacid0	000050-21-5	90.08	C3H6O3	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	6	corticosterone	000050-22-6	332.44	C21H28O4	<input type="checkbox"/>	0
<input type="checkbox"/>	7	hydrocortisone0	000050-23-7	362.46	C21H30O5	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	8	estradiol0	000050-28-2	272.38	C18H24O2	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	9	ddt	000050-29-3	354.49	C14H9Cl5	<input type="checkbox"/>	0
<input type="checkbox"/>	10	benzo(a)pyrene	000050-32-8	252.31	C20H12	<input type="checkbox"/>	0
<input type="checkbox"/>	11	cocaine0	000050-36-2	303.35	C17H21NO4	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	12	desipramine0	000050-47-5	266.38	C18H22N2	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	13	imipramine0	000050-49-7	280.41	C19H24N2	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	14	thioridazine0	000050-52-2	370.58	C21H26N2S2	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	15	chlorpromazine0	000050-53-3	318.87	C17H19ClN2S	<input checked="" type="checkbox"/>	0
<input type="checkbox"/>	16	sorbitol0	000050-70-4	182.17	C6H14O6	<input checked="" type="checkbox"/>	0

FIGURE 24.4 Database view in *COSMOthermX*. Databases can be searched and columns are sortable.

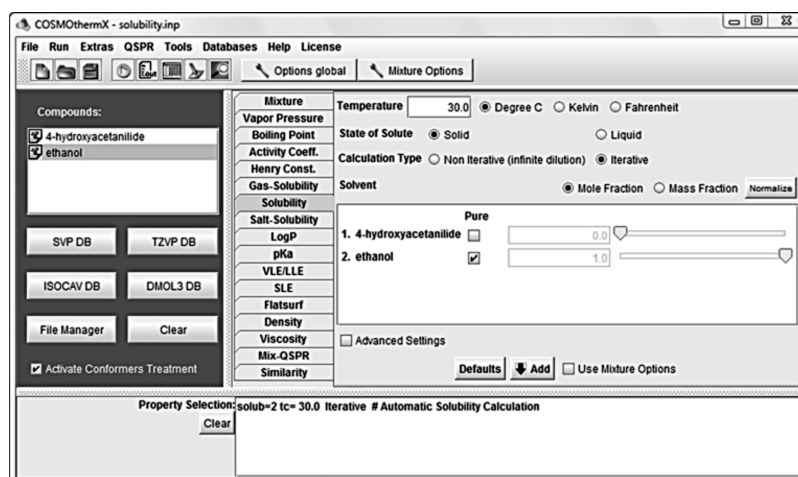


FIGURE 24.5 Overview of COSMOthermX with compound list, solubility panel and input section. When the solute state is set to liquid, $\Delta G_{\text{fus}} = 0$ is used, while with solid solute state, given or estimated values for ΔG_{fus} are used. The iterative refinement can also be set in the solubility panel. In the solvent frame, the solvent composition is set to pure for the respective solvent. Pictured here are settings for absolute solubility using the iterative refinement procedure.

relative solubility data apparently fall into two groups, as can be seen from Figure 24.7. Solubility data in alcoholic solvents are grouped together on a rather straight line with a slope of 0.7 and a relative shift of 1.4 log units compared to the experimental data, while the second, more scattered group has a larger shift (≥ 2 log units), but a similar slope. There is one severe outlier in the data set, carbon tetrachloride, where the acetaminophen solubility is, in contrast to the trend observed in the other solvents, underestimated by 1.5 log units. Since the experimental value for the solubility

in carbon tetrachloride comes from a single measurement, and the other solvents appear to be described reasonably by the model prediction apart from a general overestimation due to the missing free energy of fusion term, we tend to consider this experimental value as questionable.

Using the experimental data for T_{melt} and ΔH_{fus} , absolute solubilities of acetaminophen in the solvent data set are also computed. The absolute predictions are in good quantitative agreement with experimental data, as shown in Figure 24.8. Of the 26 solvents, 4 are predicted with a positive log $S(x)$:

```

Solubility at T = 303.15 K in compound 2 (h2o) - energies are in
kcal/mol volume is in A^3 - Solvent Density = 995.363 [g/l]

Nr Compound      log10(x_solub)    mu(self)    mu(solv)    DG_fus
1 4-hydroxyacetanilide -1.68194563 -5.05034328 -2.71690337 0.00000000
2 h2o              0.00000000 -2.90487525 -2.90487525 0.00000000
...
Solubility at T = 303.15 K in compound 25 (chcl3) - energies are in
kcal/mol volume is in A^3 - Solvent Density = 1478.286 [g/l]

Nr Compound      log10(x_solub)    mu(self)    mu(solv)    DG_fus
1 4-hydroxyacetanilide -1.88836005 -5.05034328 -2.43053523 0.00000000
25 chcl3          0.00000000 -5.29052995 -5.29052995 0.00000000
...
Solubility at T = 303.15 K in compound 26 (ccl4) - energies are in
kcal/mol volume is in A^3 - Solvent Density = 1572.230 [g/l]

Nr Compound      log10(x_solub)    mu(self)    mu(solv)    DG_fus
1 4-hydroxyacetanilide -4.51493557 -5.05034328 1.21343282 0.00000000
26 ccl4           0.00000000 -7.50124553 -7.50124553 0.00000000
...
Solubility at T = 303.15 K in compound 27 (toluene) - energies are in
kcal/mol volume is in A^3 - Solvent Density = 860.666 [g/l]

Nr Compound      log10(x_solub)    mu(self)    mu(solv)    DG_fus
1 4-hydroxyacetanilide -3.05977298 -5.05034328 -0.80538062 0.00000000
27 toluene        0.00000000 -4.62709909 -4.62709909 0.00000000

```

FIGURE 24.6 Excerpt from the COSMOtherm table file for the solubility calculation of acetaminophen in pure solvents. log $10(x_{\text{solub}})$ indicates the logarithmic solubility in mole fractions.

TABLE 24.1 Experimental, Predicted Relative and Predicted Absolute Solubilities of Acetaminophen in Pure Solvents

Solvent	Experimental		Predicted Relative			Predicted Absolute			Error
	c_s (g/kg)	$\log S$ (mg/g)	$\log S(x)$	S (mg/g)	$\log S$ (mg/g)	$\log S(x)$	S (mg/g)	$\log S$ (mg/g)	
Water	17.39	1.24	-1.6819	174.53	2.24	-3.0745	7.07	0.85	-0.39
Methanol	371.61	2.57	0.6155	19465.33	4.29	-1.0119	459.02	2.66	0.09
Ethanol	232.75	2.37	0.4809	9929.62	4.00	-1.0687	280.11	2.45	0.08
1,2-Ethanediol	144.3	2.16	0.1829	3711.29	3.57	-1.2535	135.86	2.13	-0.03
1-Propanol	132.77	2.12	0.2182	4157.36	3.62	-1.2441	143.37	2.16	0.03
2-Propanol	135.01	2.13	0.3858	6115.50	3.79	-1.1268	187.86	2.27	0.14
1-Butanol	93.64	1.97	0.0690	2390.35	3.38	-1.3675	87.50	1.94	-0.03
1-Pentanol	67.82	1.83	-0.0639	1480.10	3.17	-1.4866	55.92	1.75	-0.08
1-Hexanol	49.71	1.70	-0.1756	987.32	2.99	-1.5932	37.75	1.58	-0.12
1-Heptanol	37.43	1.57	-0.2630	709.92	2.85	-1.6779	27.31	1.44	-0.14
1-Octanol	27.47	1.44	-0.3589	507.95	2.71	-1.7716	19.64	1.29	-0.15
Acetone	111.65	2.05	0.9328	22297.15	4.35	-0.8016	411.01	2.61	0.57
2-Butanone	69.99	1.85	0.6508	9380.65	3.97	-0.9153	254.79	2.41	0.56
4-Methyl-2-pentanone	17.81	1.25	0.1126	1956.04	3.29	-1.2844	78.41	1.89	0.64
Tetrahydrofuran	155.37	2.19	1.6842	101306.53	5.01	0.0527	2366.72	3.37	1.18
1,4-Dioxane	17.08	1.23	1.0644	19898.20	4.30	-0.7332	317.14	2.50	1.27
Ethyl acetate	10.73	1.03	0.0872	2097.21	3.32	-1.2415	98.38	1.99	0.96
Acetonitrile	32.83	1.52	0.1079	4720.92	3.67	-1.2689	198.25	2.30	0.78
Diethylamine	1316.9	3.12	3.5201	6845852.30	6.84	1.4457	57683.45	4.76	1.64
<i>N,N</i> -Dimethylformamide	1012.02	3.01	2.2018	276105.51	5.44	0.4567	4965.64	3.70	0.69
Dimethyl sulfoxide	1132.56	3.05	3.3062	3915389.90	6.59	0.2699	3601.53	3.56	0.50
Acetic acid	82.72	1.92	0.3232	5298.62	3.72	-1.1632	172.87	2.24	0.32
Dichloromethane	0.32	-0.49	-1.8354	26.00	1.41	-3.1820	1.17	0.07	0.56
Chloroform	1.54	0.19	-1.8884	16.37	1.21	-3.2647	0.69	-0.16	-0.35
Carbon tetrachloride	0.89	-0.05	-4.5149	0.03	-1.52	-5.9289	0.00	-2.94	-2.89
Toluene	0.34	-0.47	-3.0598	1.43	0.16	-4.4573	0.06	-1.24	-0.77

$\log S(x)$ indicates the logarithmic solubility in mole fractions. Experimental data are taken from Ref. 12.

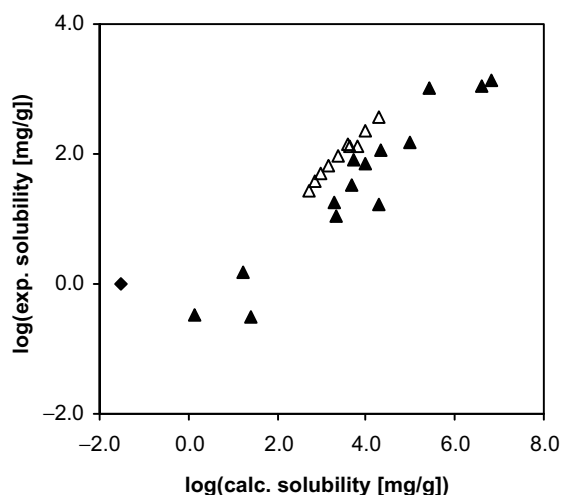


FIGURE 24.7 Predicted relative solubility of acetaminophen versus experimental data in pure solvents at 303.15K. Triangles represent relative solubility data, with empty triangles (Δ) representing alcoholic solvents and water and solid triangles (\blacktriangle) representing the remainder of the solvent data set. One outlier (carbon tetrachloride) is represented by a solid diamond (\blacklozenge).

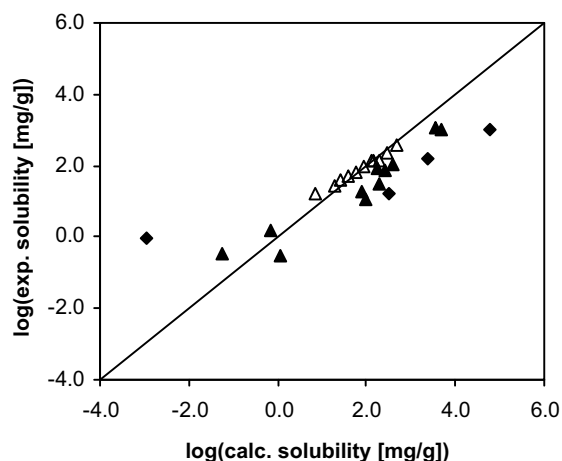


FIGURE 24.8 Predicted absolute solubility of acetaminophen versus experimental data in pure solvents at 303.15 K. Empty triangles (Δ) represent absolute solubility data of alcoholic solvents and water, solid triangles (\blacktriangle) represent the remainder of the solvent data set. Four outliers (carbon tetrachloride, diethylamine, 1,4-dioxane, tetrahydrofuran) represented by solid diamonds (\blacklozenge).

tetrahydrofuran, diethylamine, *N,N*-diethylformamide, and dimethyl sulfoxide. The rmse for all solvents is 0.77 log units. There are, however, four significant outliers in the data set, one of them, carbon tetrachloride, deviates by almost 3 log units from the experimental solubility and has already been discussed above. While the solubility of carbon tetrachloride is severely underestimated by the model prediction, the other outliers are overpredicted solubilities. For diethylamine, the predicted solubility is 1.64 log units too high. We attribute this error to the known systematic error of COSMO and COSMO-RS for secondary and tertiary aliphatic amines [5]. Another outlier is the predicted solubility of acetaminophen in 1,4-dioxane (317.14 mg/g), which deviates by 1.27 log units from the value reported by Granberg and Rasmuson (17.1 mg/g). It is noteworthy that the deviation is much less (0.56 log units) if compared to the experimental solubility reported by Romero et al. (86.9 mg/g) [11]. The fourth outlier in the data set is in tetrahydrofuran, where the predicted solubility deviates by 1.13 log units from the experimental data. Here, we do not have an explanation for the error of the model prediction. However, since tetrahydrofuran and 1,4-dioxane are rather similar solvents, the question arises whether the uncertainty of the experimental data for tetrahydrofuran might be comparable to the case of 1,4-dioxane, where there is a deviation between the published experimental data from the different sources.

With the four outliers removed, the overall rmse reduces to 0.46 log units. While the rmse is very small for alcohols (0.10 log units), the calculated solubilities in polar aprotic solvents like acetone or hydrophobic solvents like toluene are systematically overpredicted. The rmse for the remainder of the data set without the alcoholic solvents is 0.86 log units. It should be noted that the estimate of ΔG_{fus} based on ΔH_{fus} and T_{melt} itself may have an error of the order of 0.5 log units, making these absolute deviations uncertain.

EXAMPLE 24.2 SOLUBILITY OF ACETAMINOPHEN IN BINARY MIXTURES OF WATER-ACETONE AND TOLUENE-ACETONE AT 25°C

Experimental data for acetaminophen solubility in water-acetone and acetone-toluene binary solvents were also reported by Granberg and Rasmuson [12] and are used here for comparison with the model prediction.

As in the previous example, the COSMO files of the compounds were calculated on the BP86/TZVP level of theory. Absolute solubility predictions are calculated using the experimental data for ΔG_{fus} of acetaminophen and employing the iterative refinement procedure for the solubility. The calculations are done for the compositions that were measured by Granberg and Rasmuson (Figure 24.9).

The predicted solubilities of acetaminophen in the binary solvent system can be extracted from the *COSMOtherm* table file shown in Figure 24.10. Table 24.2 lists the predicted solubilities together with the experimental data.

Figures 24.11 and 24.12 show the prediction results for the acetaminophen solubility in water-acetone and acetone-toluene binary solvent mixtures. The solubility of acetaminophen in the water-acetone binary mixture is nonideal, with a solubility peak at ~70% mass fraction of acetone. The nonideal solubility behavior is also captured by the model prediction, with the solubility peak slightly shifted to higher acetone content of the binary solvent. Since the solubility in pure acetone is overpredicted by 0.59 log units, the model prediction for the binary mixture does not show the strong decrease in solubility for very high acetone content of the solvent mixture that is found in the experiment.

The prediction results for the acetone-toluene system (Figure 24.12) are consistent with the trends exhibited by the experimental data. Again, we see effects of the

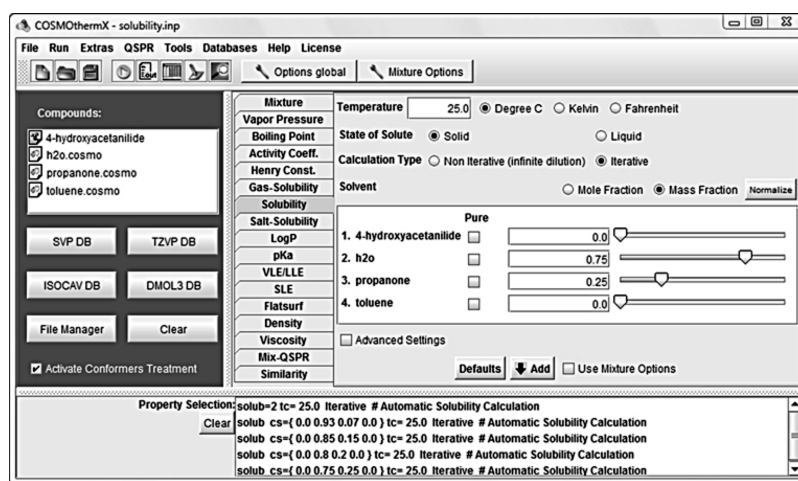


FIGURE 24.9 Input section of *COSMOthermX* with entries for a list of solvent compositions.

Solubility at T = 298.15 K at given concentration CS={ 0.0 85 15 0 } - energies are in kcal/mol volume is in A^3					
Nr	Compound	log10(x_solub)	mu(self)	mu(solv)	DG_fus
1	4-hydroxyacetanilide	-2.39941309	-5.22146393	-3.98117592	2.03362348
2	h2o	0.00000000	-2.98262140	-2.98262140	0.00000000
3	propanone	0.00000000	-1.72077455	-1.72077628	0.00000000
4	toluene	-3.08354307	-4.65191005	-0.44452822	0.00000000

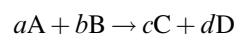
FIGURE 24.10 Excerpt from the COSMO $_{therm}$ table file for the solubility calculation of acetaminophen in a binary solvent system.

overpredicted solubility in pure acetone and the lower predicted solubility in pure toluene, but the relatively ideal solubility behavior of acetaminophen in the acetone–toluene binary mixture is found also by the model prediction.

24.3 CHEMICAL REACTIONS IN SOLUTION

Calculation of reaction energies in the gas phase is a standard application in quantum chemistry. The computational pre-

diction of free energies of reaction in solution is more involved, but still a rather straightforward procedure. Generally, the free energy of reaction is the difference of the total free energies of the reactants and the products of the reaction. For a reaction



where A and B are the reactants with stoichiometric coefficients a and b , and C and D are the reaction products with

TABLE 24.2 Experimental and Predicted Solubilities of Acetaminophen in Water–Acetone and Acetone–Toluene Binary Solvent Mixtures

% Mass Fraction			Experimental			Predicted		Error
Water	Acetone	Toluene	c_s (g/kg)	$\log S$ (mg/g)	$\log S$ (x)	S (mg/g)	$\log S$ (mg/g)	
100	0	0	14.90	1.17	-3.1508	5.93	0.77	-0.40
93	7	0	28.18	1.45	-2.7658	13.69	1.14	-0.31
85	15	0	53.0	1.72	-2.3994	29.99	1.48	-0.25
80	20	0			-2.1978	45.87	1.66	1.66
75	25	0			-2.0125	67.47	1.83	1.83
70	30	0	150.0	2.18	-1.8424	95.66	1.98	-0.20
65	35	0			-1.6879	130.58	2.12	2.12
60	40	0			-1.5500	171.24	2.23	2.23
55	45	0			-1.4290	215.47	2.33	2.33
50	50	0	327.0	2.51	-1.3242	260.58	2.42	-0.10
30	70	0	454.6	2.66	-1.0260	408.68	2.61	-0.05
15	85	0	420.3	2.62	-0.8851	452.18	2.66	0.03
7	93	0	302.2	2.48	-0.8364	438.41	2.64	0.16
3	97	0	197.1	2.29	-0.8247	415.72	2.62	0.32
0	100	0	99.8	2.00	-0.8285	386.32	2.59	0.59
0	95	5	91.7	1.96	-0.8517	359.46	2.56	0.59
0	90	10	82.4	1.92	-0.8774	332.39	2.52	0.61
0	85	15	75.7	1.88	-0.9062	305.14	2.48	0.61
0	80	20	66.4	1.82	-0.9386	277.63	2.44	0.62
0	70	30	52.8	1.72	-1.0165	222.79	2.35	0.63
0	60	40	37.08	1.57	-1.1186	168.79	2.23	0.66
0	50	50	26.56	1.42	-1.2563	117.58	2.07	0.65
0	30	70	8.55	0.93	-1.7149	37.19	1.57	0.64
0	20	80	3.39	0.53	-2.1018	14.50	1.16	0.63
0	15	85	2.12	0.33	-2.3663	7.68	0.89	0.56
0	7	93	0.78	-0.11	-2.9938	1.73	0.24	0.35
0	0	100	0.37	-0.43	-4.6734	0.03	-1.46	-1.03

$\log S(x)$ indicates the logarithmic solubility in mole fractions. Experimental data are taken from Ref. 15.

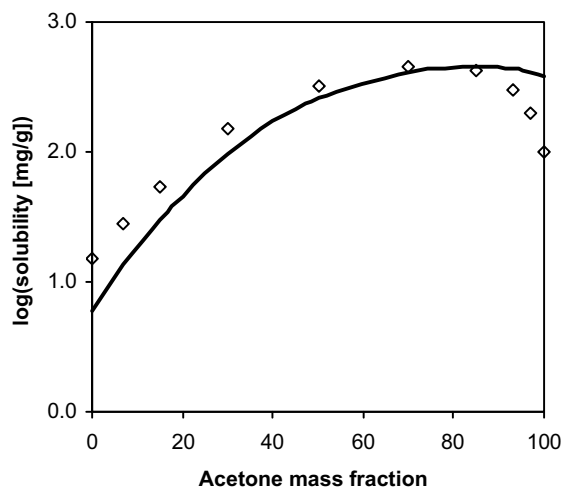


FIGURE 24.11 Experimental and predicted solubility of acetaminophen in acetone–water binary mixtures at 298.15K. Diamonds (◇) are experimental data, the solid line is from the model prediction.

stoichiometric coefficients c and d , the free energy of reaction can be calculated from the difference of the sums of free energies on both sides of the reaction:

$$\Delta G_r = [c \cdot G(C) + d \cdot G(D)] - [a \cdot G(A) + b \cdot G(B)] \quad (24.13)$$

The free energies of the reactants and products, and thus the free energy of reaction, depends on the conditions under which the reaction takes place. The free energy of reaction in the gas phase differs from the free energy of reaction in solution, and it is different in each specific solvent.

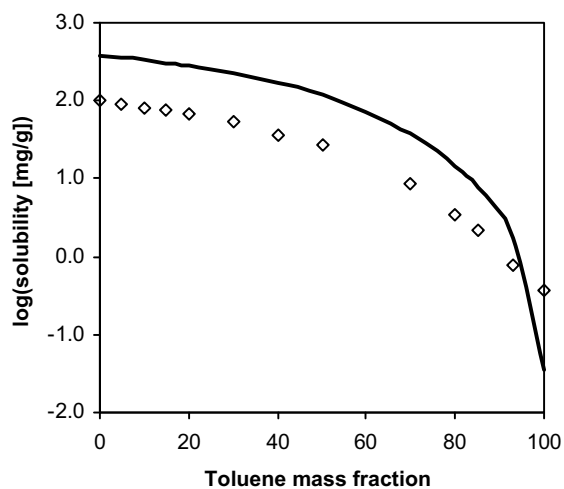


FIGURE 24.12 Experimental and predicted solubility of acetaminophen in toluene–acetone binary mixtures at 298.15K. Diamonds (◇) are experimental data, the solid line is from the model prediction.

In solution, the Gibbs free energy of a species is

$$G(i) = E_{\text{gas}}(i) + \Delta G_{\text{solv}}(i) \quad (24.14)$$

$E_{\text{gas}}(i)$ is the gas-phase energy of the compound, and for computational predictions of the reaction free energy it should be taken from an adequate quantum chemical (DFT or post-Hartree Fock) level. $\Delta G_{\text{solv}}(i)$, the free energy of solvation, describes the change of the free energy that occurs when the compound is dissolved from the gas phase into the liquid phase. This contribution to the total free energy of a compound can be computed using COSMO-RS.

Using the gas-phase energies of the compounds and the free energies of solvation of the compounds, the free energy of reaction in solution can be calculated according to a thermodynamic cycle as depicted in Figure 24.13.

In order to compute the lower horizontal leg of the cycle, corresponding to the reaction in solution, we have to take the appropriate sums and differences of the upper horizontal leg, that is, the gas-phase reaction, and the vertical legs, that is, the solvation energies of the compounds:

$$\Delta G_r(\text{sol}) = \Delta G_r(\text{gas}) + [c\Delta G_{\text{solv}}(C) + d\Delta G_{\text{solv}}(D)] - [a\Delta G_{\text{solv}}(A) + b\Delta G_{\text{solv}}(B)] \quad (24.15)$$

$\Delta G_r(\text{gas})$ can be calculated from the chemical potential of the compounds in the gas phase. As already mentioned, the quantum chemical gas-phase energy E_{gas} of a compound is computed in vacuum at absolute zero. Furthermore, E_{gas} does not account for vibrational motion that is present even at $T=0\text{K}$. The so-called zero-point energy or zero-point vibrational energy (ZPE) can be computed quantum chemically from the vibrational frequencies of the compound and is a standard correction to E_{gas} . Using the ZPE, the free energy of a compound can be calculated as

$$G(i) = E_{\text{gas}}(i) + \text{ZPE}(i) + \Delta G_{\text{solv}}(i) \quad (24.16)$$

As a further refinement for the gas-phase energies of the compounds and the resulting reaction energy, the temperature dependent thermal contributions to the free energy μ_{vib} of the molecule can be calculated. From vibrational frequencies the molecular translational, rotational, and vibrational partition functions, q_{trans} , q_{rot} and q_{vib} , can be calculated,

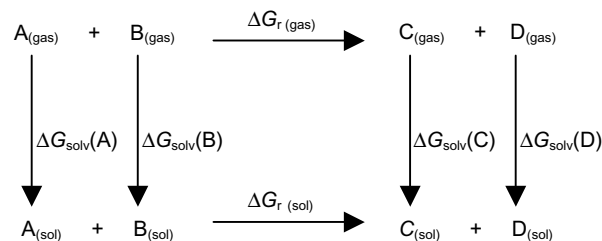


FIGURE 24.13 Cycle for computation of a free energy change in solution.

thus enabling prediction of thermodynamic functions at temperatures other than 0K and finite pressure.

$$\begin{aligned} G(i) &= E_{\text{gas}}(i) + \text{ZPE}(i) - RT \ln(q_{\text{trans}} \cdot q_{\text{rot}} \cdot q_{\text{vib}}) \\ &\quad + \Delta G_{\text{solv}}(i) \\ &= E_{\text{gas}}(i) + \mu_{\text{vib}} + \Delta G_{\text{solv}} \end{aligned} \quad (24.17)$$

The other terms required for equation 24.15, the free energies of solvation ΔG_{solv} , can be obtained from a COSMOtherm prediction of the reverse process, that is, from a vapor pressure prediction. The partial vapor pressure $P(i)$ that is calculated by COSMOtherm corresponds to the pure compound vapor pressure times the activity coefficient and is related to ΔG_{solv} by

$$\Delta G_{\text{solv}}(i) = RT \ln(10) [\log_{10}(P(i)) - \log_{10}(P)] \quad (24.18)$$

with P being the reference pressure at which the reaction takes place.

Note that if the reactant or product compounds are present in the mixture at a finite concentration with a mole fraction $x(i)$ (e.g., if the reaction takes place in bulk reactant liquid), an entropic contribution $RT \ln(x(i))$ of the compound has to be added to the compounds free energy $G(i)$.

24.3.1 Heat of Reaction

The heat of reaction or reaction enthalpy in solution can be calculated by a procedure similar to the free energy of reaction. Instead of the free energy of solvation of the compounds we make use of the heat of vaporization ΔH_{vap} . Since ΔH_{vap} is the enthalpy that is needed to transfer the compound from the liquid phase to the gas phase, it has to be subtracted from the gas-phase energy to obtain the enthalpy of the compound in solution:

$$H(i) = E_{\text{gas}}(i) - \Delta H_{\text{vap}}(i) \quad (24.19)$$

Zero-point vibrational energy corrections or thermal correction terms for the enthalpy in the gas phase can also be used as corrections to the gas-phase energies of the compounds for the calculation of the heat of reaction.

Similarly to equation 24.13 for the free energy of reaction ΔG_{r} , the heat of reaction ΔH_{r} can be calculated from the enthalpy of the compounds:

$$\Delta H_{\text{r}} = [c \cdot H(\text{C}) + d \cdot H(\text{D})] - [a \cdot H(\text{A}) + b \cdot H(\text{B})] \quad (24.20)$$

24.3.2 Equilibrium Constants

The equilibrium constant K of a reaction is related to the free energy of reaction by

$$\ln K = -\frac{\Delta G^{\circ}}{RT} \quad (24.21)$$

For a reaction in an ideal solution, that is, in infinite dilution, the equilibrium constant can be calculated using the reaction free energy in solution according to equation 24.15. The free energies of the individual compounds can be computed using different quantum chemical correction terms as described above.

24.3.3 Accuracy

In the described procedure the free energies of the compounds are calculated from two main contributions, the quantum chemical gas-phase energy and the free energy of solvation. The accuracy of the resulting reaction energy is determined mainly by the accuracy of the underlying quantum chemical method. With DFT methods like the BP86 functional, errors of the absolute reaction energy can be in the range of 10 kcal/mol or more [13, 14]. However, for relative reaction energies of one reaction in different solvents in a solvent screening application, that is, if we are looking at the variation of the solvation energy only, calculated as the COSMOtherm contribution in the liquid phase, the accuracy is much higher. For such relative reaction energy predictions considering ΔG_{solv} or ΔH_{vap} from the COSMOtherm vapor pressure prediction only, an accuracy of 0.5 kcal/mol can be expected.

For a higher accuracy of absolute predictions of the reaction energy or enthalpy, it follows that a more accurate quantum chemical method should be applied for the calculation of the gas-phase energy, for example, the MP2 or Coupled-Cluster methods combined with adequate basis sets. Quantitative improvement of the total free energy of a compound $G(i)$ can also be achieved by the ZPE and thermal corrections to the gas-phase energy of the compound.

24.3.4 Calculation of the Free Energy of Reaction and Heat of Reaction

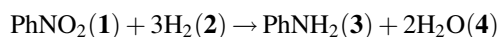
A procedure for the computational prediction of the free energy of reaction ΔG_{r} or heat of reaction ΔH_{r} using quantum chemical gas-phase energies and free energies of solvation or heats of vaporization from COSMO-RS is described as follows:

- Compute the reactant and product molecules using the DFT methods for which COSMOtherm is parameterized. Both COSMO and gas-phase quantum chemical calculations are required. Then use COSMOtherm to obtain the ΔG_{solv} and/or ΔH_{vap} values.
- Compute gas-phase energies E_{gas} of the reactant and product molecules with a high-level *ab initio* method, for example, the Coupled-Cluster method.
- Compute vibrational frequencies for reactant and product molecules to obtain the ZPE correction or the

thermal corrections to the gas-phase energies of the compounds. Vibrational frequency calculations at a DFT level of theory are usually sufficiently accurate.

- Combine E_{gas} , ZPE or thermal corrections and ΔG_{solv} to compute the compounds $G(i)$ according to equations 24.16 or 24.17, or E_{gas} , ZPE or thermal corrections and ΔH_{vap} to compute the compounds $H(i)$.
- Calculate the free energy of reaction ΔG_r or the heat of reaction ΔH_r from the difference of the sums of free energies or enthalpies on both sides of the reaction as described by equations 24.13 and 24.20.

EXAMPLE 24.3 ESTIMATE THE HEAT OF REACTION ΔH_r FOR THE REDUCTION OF NITROBENZENE TO ANILINE IN THE LIQUID PHASE IN THF AT 25°C



Following the procedure described above, we calculate the gas-phase energies and the heats of vaporization. In this example, we calculate the gas-phase energies of the compounds on the MP2/TZVPP quantum chemical level. Furthermore, we employ the ZPE and thermal corrections to the enthalpy in the gas phase to refine the calculated heat of reaction.

For the calculation of the heat of vaporization with *COSMOtherm* we have to provide COSMO files for all compounds involved in the reaction, including the solvent, and the gas-phase energies for the reactants and products. For the compounds involved here, the COSMO files and gas-phase energies are available from the database included in the *COSMOtherm* package. However, to exemplify the procedure, we will give a short overview of the required quantum chemical calculations.

Since we require the gas-phase energy as well as the screening charge surface of the compounds for the calculation of the heat of vaporization, we need to do geometry optimizations both in the gas phase and in the conductor, using the COSMO model with ideal screening. *COSMOtherm* is parameterized for the BP86/TZVP and the BP86/SVP//AM1 quantum chemical levels. Here, we will use the higher one of the two levels, that is, the BP86 functional and the TZVP basis set. The QC calculations are performed using the TURBOMOLE [15] quantum chemical program suite.

For the QC calculation we have to provide 3D structures of the compounds, which can be generated by an external tool or built with the molecular builder of the TURBOMOLE graphical user interface. With the starting geometries, the following QC calculations are performed for each compound

- A gas-phase geometry optimization on the BP86/TZVP quantum chemical level.

- A geometry optimization on the BP86/TZVP/COSMO quantum chemical level.
- A gas-phase geometry optimization on the MP2/TZVPP quantum chemical level.
- A vibrational frequency calculation with the optimized BP86/TZVP molecular gas-phase structure to obtain the zero-point vibrational energy.
- Using the results from the vibrational frequency calculations, we also compute the thermal contributions with the corresponding interactive tool of the TURBOMOLE suite.

These steps are described in more detail in Appendix 24. A. Further information about how the quantum chemical calculations are carried out can also be found in the documentation of the TURBOMOLE program suite and the TmoleX documentation.

The heats of vaporization H_{vap} are calculated with *COSMOtherm* at a temperature of 25°C and a solvent composition of pure tetrahydrofuran. For the COSMO-RS vapor pressure prediction, quantum chemical gas-phase energies from the BP/TZVP level are used to calculate the chemical potential of the compound in the gas phase.

The calculated data for the heats of vaporization of the compounds can be extracted from the *COSMOtherm* output file (Figure 24.14). In Table 24.3, the results for the individual energy contributions of the compounds both from the QC calculations and the *COSMOtherm* calculation are tabulated together with the resulting data for the heat of reaction. The data for the heat of reaction were calculated from $\Delta H_r = [H(3) + 2 \cdot H(4)] - [H(1) + 3 \cdot H(2)]$, taking into account the stoichiometry of the reaction (equation 24.20). For the enthalpy values of the compounds $H(i)$, different correction terms for the quantum chemical gas-phase energies were employed. For comparison, heats of reaction using the quantum chemical gas-phase energies from the MP2/TZVPP level and the BP86/TZVP level are tabulated.

The catalytic reduction of aromatic nitro compounds in the gas phase is known to be a highly exothermic process. For the gas-phase reduction of nitrobenzene, a heat of reaction of $\Delta H_r = -131 \pm 3$ kcal/mol was published [16]. Absolute values for the heat of reaction of the reduction of nitrobenzene to aniline in tetrahydrofuran solution could not be found in the literature, but heats of reaction for the reduction of other R-NO₂ compounds in solution have been found independent of R or solvent to be in the range of -125 to -130 kcal/mol [17].

Using MP2/TZVPP gas-phase energies for the compounds, thermal corrections, and the heat of vaporization from the COSMO-RS prediction, the calculated heat of reaction in solution is $\Delta H_r = -127$ kcal/mol. With ZPE correction only, the calculated heat of reaction in solution

```

Results for mixture 1
-----
Temperature      :      298.150 K

Compound Nr.    :           1           2           3           4           5
Compound        :   nitrobenz   aniline   h2           h2o         thf
Mole Fraction   :      0.0000   0.0000   0.0000   0.0000   1.0000

Compound: 1 (nitrobenzene)
Chemical potential of the compound in the mixture :      -3.54480 kcal/mol
Log10(partial pressure [mbar])                  :      -0.32317
Free energy of molecule in mix (E_COSMO+dE+Mu)   : -274197.04529 kcal/mol
Total mean interaction energy in the mix (H_int)  :      -5.23132 kcal/mol
Misfit interaction energy in the mix (H_MF)       :       2.46378 kcal/mol
H-Bond interaction energy in the mix (H_HB)      :      -0.27976 kcal/mol
VdW interaction energy in the mix (H_vdW)        :      -7.41534 kcal/mol
Ring correction                                  :      -1.14821 kcal/mol
Vapor pressure of compound over the mixture     :           0.00000 mbar
Chemical potential of compound in the gas phase :       0.98955 kcal/mol
Heat of vaporization                             :       12.02087 kcal/mol

```

FIGURE 24.14 Excerpt from the mixture output section of the COSMOtherm output file for a vapor pressure calculation of compounds in pure tetrahydrofuran at 25°C.

is $\Delta H_r = -125$ kcal/mol. Both predicted values are well inside the range of the experimental data for comparable reactions. In contrast, when ZPE and thermal corrections are ignored, the heat of reaction is overestimated by several kcal/mol with a value of $\Delta H_r = -141$ kcal/mol. Table 24.3 also shows heats of reaction calculated from the BP86/TZVP gas-phase energies. The heat of reaction without ZPE or thermal corrections is $\Delta H_r = -125$ kcal/mol, which agrees well with experimental data, but when the correction terms, which should in general lead to a better prediction, are included, the heat of reaction in solution is significantly underestimated with predicted values of $\Delta H_r = -109$ kcal/mol and $\Delta H_r = -111$ kcal/mol, respectively. However, it should be noted that absolute errors in the range of 10–20 kcal/mol are not unusual for pure DFT functionals like BP86 [13,14].

EXAMPLE 24.4 ESTIMATE REACTION FREE ENERGY, EQUILIBRIUM CONSTANT, AND EQUILIBRIUM COMPOSITION FOR THE REACTION OF 1-METHOXY-2-PROPANONE 1 AND ISOPROPYLAMINE 2

First, the gas-phase energies of the reactants and products of the transamination reaction (Figure 24.15) are calculated quantum chemically. In this example, we also employ the MP2 level of theory and the TZVPP basis set. As starting structures for the gas-phase geometry optimizations, the 3D structures from the BP86/TZVP level are used, which are available from the COSMObase, a database of validated COSMO files and gas-phase structures. With these structures we perform gas-phase geometry optimizations using the MP2/TZVPP method and basis set combination.

TABLE 24.3 Gas-Phase Energies, Heat of Vaporization, Zero-Point Vibrational Energies, Thermal Corrections, Total Enthalpies of Compounds, and Heat of Reaction for the Hydrogenation of Nitrobenzene

	1	2	3	4	ΔH_r	
					MP2/TZVPP	(BP86/TZVP)
H_{vap}	12.02	1.35	14.54	10.24		
E_{gas} (BP86/TZVP) (Hartree)	-436.944122	-1.177446	-287.715215	-76.465165		
E_{gas} (MP2/TZVPP) (Hartree)	-435.958485	-1.164647	-287.000283	-76.323461		
ZPE (BP86/TZVP) (Hartree)	0.099295	0.009850	0.113236	0.020640		
ΔH_T (thermal corrections)	67.36	8.26	75.45	15.32		
$H = E_{\text{gas}} - H_{\text{vap}}$	-273580.11	-732.17	-180109.95	-47903.93	-141.19	(-125.06)
$H = E_{\text{gas}} + \text{ZPE} - H_{\text{vap}}$	-273517.80	-725.99	-180038.89	-47890.98	-125.08	(-108.95)
$H = E_{\text{gas}} + \Delta H_T - H_{\text{vap}}$	-273512.75	-723.92	-180034.50	-47888.61	-127.22	(-111.09)

Enthalpy terms are in kcal/mol, quantum chemical gas-phase energies and ZPE are in Hartree.

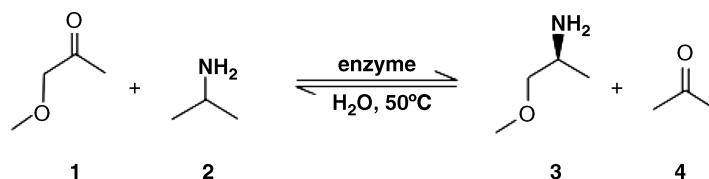


FIGURE 24.15 Catalytic transamination of 1-methoxy-2-propanone and isopropylamine.

As a further refinement for the gas-phase energies, vibrational frequency calculations for the compounds are carried out, and the thermal contributions from the vibrational frequencies to the gas-phase energy are computed using the corresponding tool of the TURBOMOLE suite. These calculations are carried out on the BP86/TZVP level.

In our next step for the calculation of the reaction free energy, the free energy of solvation of each compound is calculated from equation 24.18, using the partial pressures from the vapor pressure prediction of the COSMOtherm program. Reactants, products, and solvent for the reaction are taken from the COSMObase. The conditions for the vapor pressure calculation are set to a temperature of 50°C and the solvent composition is set to pure water. Quantum chemical gas-phase energies of the compounds are used to calculate the chemical potentials of the compounds in the gas phase. The calculated partial pressures can be extracted from the COSMOtherm output file (Figure 24.16).

Energy terms and corrections for the reactants and products from this procedure are tabulated in Table 24.4. Calculated data for the free energy of reaction ΔG_r are also tabulated, using different QC correction terms. ΔG_r is calculated according to the stoichiometry of the reaction as $\Delta G_r = [G(3) + G(4)] - [G(1) + G(2)]$.

With the MP2/TZVPP gas-phase energies and the solvation free energies of the compound, the free energy of reaction in solution is $\Delta G_r = -3.27$ kcal/mol, corresponding to an equilibrium constant of $\ln K = 5.10$. The zero-point

vibrational energy corrections for the reactants and products have very little influence on the overall reaction energy ($\Delta G_r = -3.29$ kcal/mol, $\ln K = 5.12$ kcal/mol), but when thermal correction are included in the free energies $G(i)$ of the compounds, the free energy of reaction decreases to $\Delta G_r = -2.20$ kcal/mol and the equilibrium constant decreases to $\ln K = 3.43$. The equilibrium constant indicates that the equilibrium position of the reaction lies on the right-hand side of the reaction equation. The relative amount of reactants at equilibrium is 0.08 each, and the relative amount of products at equilibrium is 0.42 each. The experimental equilibrium constant for the reaction of 1-methoxy-2-propanone with isopropylamine is $K = 7.8$ ($\ln K = 2.05$) [18]. Thus, the calculated free energy of reaction and equilibrium constant are in excellent agreement with the experimental data.

24.4 CONCLUSION AND OUTLOOK

In this chapter, we presented an overview of two different applications of COSMO-RS in the drug development process. The power and main benefit of the COSMO-RS model is that properties in solution can be obtained from *ab initio* calculation without any experimental input. It does not require external data for modeling and can also be used when empirical models are not parameterized. Complex multifunctional molecules and new chemical functionalities are treated on the same footing as simple organic molecules.

```

Results for mixture 1
-----
Temperature      : 323.150 K

Compound Nr.    :      1      2      3      4      5
Compound        : isopropyl 1-methoxy 1-methoxy propanone h2o
Mole Fraction   : 0.0000 0.0000 0.0000 0.0000 1.0000

Compound: 1 (isopropylamine)
Chemical potential of the compound in the mixture : -1.17237 kcal/mol
Log10(partial pressure [mbar])                  : 4.11202
Free energy of molecule in mix (E_COSMO+dE+Mu) : -109542.55539 kcal/mol
Total mean interaction energy in the mix (H_int) : -7.42349 kcal/mol
Misfit interaction energy in the mix (H_MF)      : 2.15679 kcal/mol
H-Bond interaction energy in the mix (H_HB)      : -4.65492 kcal/mol
VdW interaction energy in the mix (H_vdW)        : -4.92535 kcal/mol
Ring correction                                  : 0.00000 kcal/mol
Vapor pressure of compound over the mixture     : 0.00000 mbar
Chemical potential of compound in the gas phase : -2.81690 kcal/mol
Heat of vaporization                             : 10.89292 kcal/mol

```

FIGURE 24.16 Excerpt from the mixture output section of the COSMOtherm output file for a vapor pressure calculation of compounds in pure water at 50°C.

TABLE 24.4 Energy Terms and Energy Corrections for the Reactants and Products for reaction 2 and Free Energies of Reaction with the Various Correction Terms

	1	2	3	4	ΔG_r	$\ln K$
$\log 10(P(i))$ (mbar)	2.72032	4.11202	3.12339	3.75518		
G_{solv}	-0.41	1.64	0.18	1.12		
E_{gas} (MP2/TZVPP) (Hartree)	-307.100753	-174.113883	-288.440644	-192.779317		
ZPE (BP86/TZVP) (Hartree)	0.112968	0.117222	0.149297	0.080872		
μ_{vib} (thermal corrections)	54.44	227.770	71.54	32.03		
$G = E_{\text{gas}}$ (MP2) + G_{solv}	-192709.05	-109256.47	-180999.06	-120969.74	-3.27	5.10
$G = E_{\text{gas}}$ (MP2) + G_{solv} + ZPE	-192638.16	-109182.91	-180905.38	-120918.99	-3.29	5.12
$G = E_{\text{gas}}$ (MP2) + G_{solv} + μ_{vib}	-192660.99	-109202.03	-180927.52	-120937.70	-2.20	3.43
Relative amount of compound in equilibrium	0.08	0.08	0.42	0.42		

The relative amount of the compounds in equilibrium has been calculated from the equilibrium constant $\ln K = 3.43$. Free energy terms and chemical potential are in kcal/mol, quantum chemical gas-phase energies and ZPE are in Hartree.

Prediction of relative drug solubility with COSMO-RS is based on a consistent thermodynamic modeling of interactions in the solvent and the supercooled state of the drug. Solvent mixtures can be treated in the same way as pure solvents and with similar accuracy. Absolute solubility prediction is limited by the availability of free energy of fusion data.

Although COSMO-RS in its present state cannot be proven to be more accurate than more empirical models with many adjusted parameters, its strength is the essential independency from experimental data. This allows for independent modeling and avoids errors resulting from erroneous experimental data on which empirical models rely. Potential improvement to the current COSMO-RS solubility prediction model include a more accurate fusion term for absolute solubility prediction, and improvement of the COSMO-RS interaction terms themselves, especially for the chemically important group of secondary and tertiary amines, but requires more reliable experimental data as are available at present.

With COSMO-RS, solubility prediction is also possible for salts. This is important as many drugs are formulated as salts. Solvent systems involving ionic liquids can also be treated with very good accuracy [19]. Furthermore, different conformational forms of molecules can be used for solubility screening, and the relative weight of conformers in different solvents can be determined. This feature basically allows for examination of conditions influencing the crystallization process and solvent screening for conformational polymorphism and pseudopolymorphism [20, 21].

Another application of COSMO-RS frequently used in pharmaceutical and agrochemical industry deals with reaction modeling. In principle, reaction equilibrium, mechanism, rate and by-product formation may be solvent dependent. Here, we investigated the influence of solvation on the free energy and heat of reaction and the reaction equilibrium only. A straightforward procedure for the computational

prediction of the free energy of reaction and heat of reaction has been shown, and the effect of the employed quantum chemical level on the absolute heat of reaction has been demonstrated. Elsewhere, it has been shown that although, depending on the quantum chemical level, absolute values for the free energy of reaction may differ substantially from experimental data, general trends are predicted correctly [22].

All applications of COSMO-RS require quantum chemical calculations of compounds, taking into account the various molecular conformations of the compounds. This constitutes the computationally most demanding part of the procedure, but has to be done only once per compound. The COSMO files of the involved compounds can be reused for other projects and all types of properties. Thus, if combined with a database of precalculated COSMO files for common compounds, thermodynamic property calculations with *COSMOtherm* can be carried out quite fast and efficiently.

24.A APPENDIX

24.A.1 Details of COSMO and Gas-Phase Calculations

For later use of the COSMO files in *COSMOtherm*, the details of the quantum chemical COSMO calculation should be consistent with one of the parameterizations of *COSMOtherm*. There are two levels of different quality mainly used in *COSMOtherm*, a lower, computationally less expensive level, and a higher level, which is computationally more time-consuming but better suited for chemical engineering applications. The higher level has also been used throughout in this contribution and requires

- BP86 DFT geometry optimization with a TZVP quality basis set and the RI approximation applied in the gas phase and in the conductor.

- For COSMO calculations only: COSMO applied in the conductor limit ($\epsilon = \infty$) using optimized element-specific COSMO radii for the cavity construction.
- If more than one conformation is considered to be potentially relevant for a compound, QC calculations have to be done for all conformations.

24.A.2 Steps for Calculating the Free Energy or Enthalpy of a Compound

Apart from the heat of vaporization or free energy of solvation that are calculated by COSMO $therm$, the free energies and enthalpies of compounds involved in the reaction examples are composed from several quantum chemical contributions.

24.A.2.1 Gas-Phase Energy The BP86/TZVP gas-phase energy is obtained from a gas-phase geometry optimization of the compound using the settings described in Appendix 24.A.1. The resulting energy value can be found in output files of the TURBOMOLE program suite. The protocol file `job.last` comprises the output of the last complete cycle and information about the settings that were applied in the program run, for example, convergence criteria. The gas-phase energy can be read from “total energy” line in the output section displayed in Figure 24.17. If the graphical user interface `TmoleX` is used, the gas-phase energy can also be taken from the Energy block of the Job-Results panel. This panel also displays information about the run of the job (Figure 24.18).

The MP2/TZVPP gas-phase energy is obtained from an MP2 geometry optimization employing the TZVPP basis set (also called `def-TZVPP`) and the RI approximation (Figure 24.19). The module used for this type of calculation is `ricc2`. In `TmoleX`, the method has to be set to RI-MP2 in the “Level” section of the “Level of Theory” panel. In the MP2 calculations for this contribution, the 1s orbitals of elements C, N, and O were kept frozen. Details about settings for frozen core orbitals can be found in the TURBOMOLE documentation.

24.A.2.2 Correction Terms from Vibrational Frequency Calculations For the zero-point vibrational energy contribution, a vibrational frequency calculation is required. This type of calculation can be done with the `aoforce` module of TURBOMOLE. The ZPE can be taken either directly from the `aoforce.out` file or from the interactive module `freeh`, which also allows for calculation of the molecular partition functions at temperatures other than 0K and finite pressure. Results of the `freeh` module are printed to standard I/O (Figure 24.20). Note that vibrational frequency calculations are based on the assumption of an harmonic oscillator and partition functions are computed within the assumption of an ideal gas and no coupling between degrees of freedom.

Vibrational frequency calculations can also be started from `TmoleX`. Start the job as a Single-Point calculation with the “Frequency Analysis” radio button ticked in the “Job Selection” section of the Single-Point calculation panel.

The thermal enthalpy contribution ΔH_T can be calculated from the thermal energy contribution printed in the `freeh` output. This is done via $\Delta H_T = \Delta G_T + RT$, where ΔG_T is the value in the “energy” column of the `freeh` output, T is the temperature, and R is the gas constant $R = 8.314472 \text{ J/(K mol)}$.

For the calculation of the thermal contribution to the free energy of a compound, μ_{vib} , the `freeh` value for “chem. pot.” should be used and not the value from the “energy” column.

After converting the contributions from the individual steps to consistent units, the total enthalpy H of a compound can be calculated as the sum of the individual contributions. For compound nitrobenzene on the MP2/TZVPP level of theory, this involves the sum $H = E_{\text{gas}}(\text{MP2}) + \Delta H_T - H_{\text{vap}}$ consisting of the following contributions listed in Table 24.3

- $E_{\text{gas}}(\text{MP2}) = -435.958485 \text{ Hartree} = -273568.09 \text{ kcal/mol}$ from the MP2 gas-phase geometry optimization.
- $\Delta H_T = \Delta G_T + RT = 279.36 \text{ kJ/mol} + 8.314472 \text{ J/(K mol)} \cdot 298.15 \text{ K} = 281.84 \text{ kJ/mol} = 67.36 \text{ kcal/mol}$

```

*****
nitrobenzene
*****

-----
| total energy      =  -436.94412249634 |
-----
: kinetic energy    =   435.14545720492 :
: potential energy  =  -872.08957970126 :
: virial theorem    =    1.99588353476 :
: wavefunction norm =    1.00000000000 :
: .....

```

FIGURE 24.17 Excerpt from the TURBOMOLE output file `job.last` from a gas-phase geometry optimization of nitrobenzene.

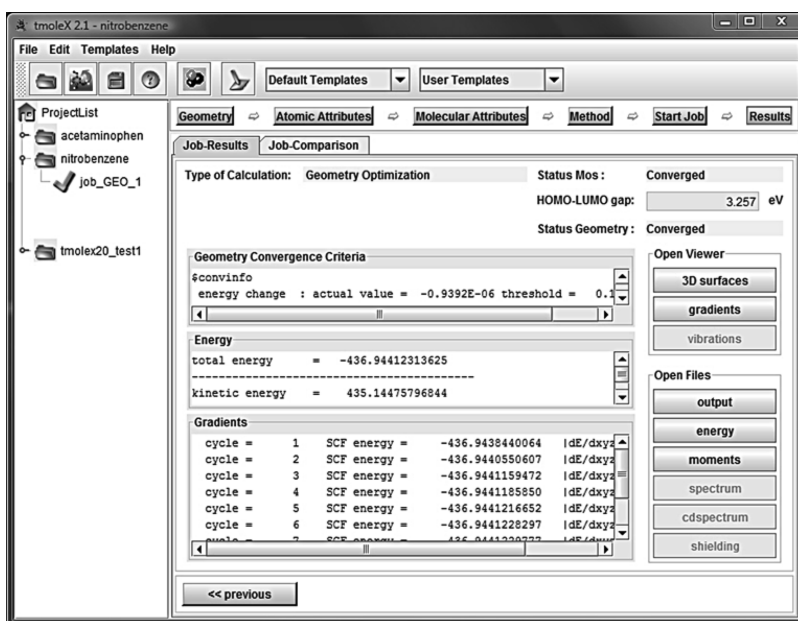


FIGURE 24.18 Job-results panel of the graphical user interface TmoleX of the TURBOMOLE program suite.

from the vibrational frequency calculation and subsequent computation of partition functions with the freeh module.

- $H_{\text{vap}} = 12.02$ kcal/mol from the COSMO-RS vapor pressure prediction.

Note that quantum chemical energies are usually expressed in Hartree, which is the atomic unit of energy. The conversion factor to the kcal/mol unit system is 627.5095 kcal/mol.

The free energy G of a compound can similarly be calculated from

$$G = E_{\text{gas}}(\text{MP2}) + G_{\text{solv}} + \mu_{\text{vib}}$$

For compound isopropylamine, this requires the following energy terms, listed in Table 24.4:

- $E_{\text{gas}}(\text{MP2}) = -174.113883$ Hartree = -109258.16 kcal/mol from the MP2 gas-phase geometry optimization.
- $\mu_{\text{vib}} = 227.77$ kJ/mol, the chemical potential value (“chem. pot.”) from a vibrational frequency calculation and subsequent computation of partition functions with the freeh module.
- $G_{\text{solv}} = 1.64$ kcal/mol, calculated from $\Delta G_{\text{solv}}(i) = RT \ln(10)[\log_{10}(P(i)) - \log_{10}(P)]$ (equation 24.18). The partial pressure of isopropylamine, $\log_{10}(P(i)) = 4.11202$ mbar, was taken from a COSMOtherm

```
*****
*
* <<<<<<<<< GROUND STATE FIRST-ORDER PROPERTIES >>>>>>>> *
*
*****

-----
Method           : MP2
Total Energy      :  -435.9584853137
-----
```

FIGURE 24.19 Excerpt from the TURBOMOLE output file job.last from an MP2 gas-phase geometry optimization of nitrobenzene using the ricc2 module of TURBOMOLE.

```

enter the range of temperatures (K) and pressures (MPa)
at which you want to calculate partition sums and free enthalpies :

tstart=(real) tend=(real) numt=(integer) pstart=(real) pend=(real) nump=(integer)

default values are :
tstart=298.15 tend=298.15 numt=1 pstart=0.1 pend=0.1 nump=1

or enter q or * to quit

-----
your wishes are :
-----

pstart= 0.1000      pend= 0.1000      nump= 1
tstart= 298.1      tend= 298.1      numt= 1

zero point vibrational energy
-----
zpe= 260.7      kJ/mol

T      p      ln(qtrans) ln(qrot) ln(qvib) chem.pot. energy entropy
(K)    (MPa)
298.15 0.1000000    17.82   12.90   3.49   175.88   279.36   0.35539

T      P      Cv      Cp
(K)    (MPa)    (kJ/mol-K) (kJ/mol-K)
298.15 0.1000000    0.1128303 0.1211446

```

FIGURE 24.20 Excerpt from the interactive output of the freeh module of TURBOMOLE.

vapor pressure prediction (Figure 24.16) and the reference pressure P was assumed to be 1 bar.

ΔG_{solv} free energy of solvation
 ΔH_{vap} heat of vaporization
 μ chemical potential
 σ screening charge density

24.6 ABBREVIATIONS

BP86	approximate DFT functional, consisting of Becke's exchange functional [23] and Perdew's correlation functional [24]
COSMO	conductor-like screening model
COSMO-RS	conductor-like screening model for realistic solvation
DFT	density functional theory
GCM	group contribution method
MC	Monte Carlo method
MD	molecular dynamics method
MP2	second-order Møller-Plesset perturbation theory
QC	quantum chemistry/quantum chemical
rmse	root mean square error
TZVP, TZVPP	Ahlrich's triple-zeta valence polarization basis sets [25]
vdW	van der Waals interaction

24.7 SYMBOLS

E_{gas}	gas-phase energy
ΔG_{fus}	free energy of fusion

REFERENCES

1. Klamt A, Schüürmann G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans.* 2 1993;799–805.
2. Klamt A. Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *J. Phys. Chem.* 1995;99:2224–2235.
3. Klamt A. *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, Elsevier, Amsterdam, 2005.
4. Gmehling J. Present status of group-contribution methods for the synthesis and design of chemical processes. *Fluid Phase Equilib.* 1998;144:37–47.
5. Klamt A, Jonas V, Bürger T, Lohrenz JCW. Refinement and parameterization of COSMO-RS. *J. Phys. Chem. A* 1998; 102: 5074–5085.
6. Klamt A, Eckert F. COSMO-RS: a novel and efficient method for the *a priori* prediction of thermophysical data of liquids. *Fluid Phase Equilib.* 2000; 172: 43–72.
7. Ben-Naim A. *Solvation Thermodynamics*, Plenum Press, New York and London, 1987.

8. Klamt A, Eckert F, Hornig M, Beck ME, Bürger T. Prediction of aqueous solubility of drugs and pesticides with COSMO-RS. *J. Comput. Chem.* 2002;23:275–281.
9. Granberg RA, Rasmuson AC. Solubility of paracetamol in pure solvents. *J. Chem. Eng. Data* 1999;44:1391–1395.
10. Manzo RH, Ahumada AA. Effects of solvent medium on solubility. V: Enthalpic and entropic contributions to the free energy changes of di-substituted benzene derivatives in ethanol:water and ethanol:cyclohexane mixtures. *J. Pharm. Sci.* 1990;79:1109–1115.
11. Romero S, Reillo A, Escalera B, Bustamante P. The behavior of paracetamol in mixtures of amphiprotic and amphiprotic–aprotic solvents. relationship of solubility curves to specific and non-specific interactions. *Chem. Pharm. Bull.* 1996;44:1061–1064.
12. Granberg RA, Rasmuson AC. Solubility of paracetamol in binary and ternary mixtures of water + acetone + toluene. *J. Chem. Eng. Data* 2000; 45: 478–483.
13. Koch W, Holthausen MC. *A Chemist's Guide to Density Functional Theory*, Wiley-VCH, Weinheim, 2001.
14. Cramer CJ. *Essentials of Computational Chemistry*, Wiley, Chichester, 2002.
15. TURBOMOLE V6.0 2009, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007. Available from <http://www.turbomole.com>.
16. Macnab JI. The role of thermochemistry in chemical process hazards: catalytic nitro reduction processes. *I. Chem. E. Symp. Ser.* 1981;68(3/S):1–15.
17. am Ende DJ. Private communication, 2009.
18. Matcham G, Bhatia M, Lang W, Lewis C, Nelson R, Wang A, Wu W. Enzyme and reaction engineering in biocatalysis: synthesis of (*S*)-methoxyisopropylamine (= (*S*)-1-methoxypropan-2-amine). *Chimia* 1999;53:584–589.
19. Diedenhofen M, Eckert F, Klamt A. Prediction of infinite dilution activity coefficients in ionic liquids using COSMO-RS. *J. Chem. Eng. Data* 2003;48:475–479.
20. Abramov Y, Mustakis J, am Ende DJ. Private communication, 2009.
21. Geertman R. Private communication and presentation at *COSMO-RS Symposium 2009*, <http://www.cosmologic.de/index.php?cosId=1901&crId=5>.
22. Peters M, Greiner L, Leonhard K. Illustrating computational solvent screening: prediction of standard Gibbs energies of reaction in solution *AICHE J.* 2008;54:2729–2734.
23. Becke AD. Density-functional exchange-energy approximation with correct asymptotic behaviour. *Phys. Rev. A* 1988;38:3098–3100.
24. Perdew JP. Density-functional approximation for the correlation-energy of the inhomogeneous electron gas. *Phys. Rev. B* 1986;33:8822–8824.
25. Schäfer A, Huber C, Ahlrichs R. Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* 1994;100:5829–5835.