ADF COSMO-RS Manual

ADF Program System
Release 2010.01

Table of Contents

Introduction

The COSMO-RS (COnductor like Screening MOdel for Realistic Solvents) was developed by Klamt and coworkers [1-3]. There are different implementations of COSMO-RS or derivatives, and different parametrizations. The method used in ADF is the one developed by Klamt et al., which is described in detail in Ref. [2], and is called here COSMO-RS. The implementation of COSMO-RS in ADF is described in Ref. [4]. In ADF it is possible to use a thermodynamically consistent combinatorial contribution to the chemical potential as is used in Ref. [3], and a temperature dependent hydrogen bond interaction, also described in Ref. [3]. The parameters in the paper [2] were reparametrized for ADF, see Ref. [4] for details. The ADF COSMO-RS command line program is called *crs*. The main authors of this program are Cory Pye (Saint Mary's University, Halifax NS Canada) and Jaap Louwen (Albemarle Corporation). The COSMO-RS GUI *ADFcrs* contains an input builder for COSMO-RS and can visualize results, see the COSMO-RS GUI tutorial and the COSMO-RS GUI reference manual.

COSMO-RS uses the intermediate results from quantum mechanical (QM) calculations on individual molecules to predict thermodynamic properties of mixtures of these molecules, for example, solubility. There are a fair number of reports of accurate prediction by COSMO-RS of thermodynamic properties in general in the literature. Many of these have been written by Klamt and co-workers, see Ref. [3] and references therein.

There are also empirical methods like UNIFAC that can predict thermodynamic properties (including the activity coefficients). These methods contain group specific parameters and are parametrized against a large data base. They will often do better than COSMO-RS methods (especially, of course, if the system of interest was part of the data base used for parameter estimation). However, these methods cannot handle every type of molecule. In particular when unusual combinations of functional groups occur (such as in drug molecules), no parametrization is available. COSMO-RS methods, on the other hand, only feature general parameters not specific to chemical groups or functionalities. All that is required is that a quantum mechanical calculation can be done on the molecule. Therefore, COSMO-RS can be a valuable tool for the prediction of chemical engineering thermodynamical properties, like, for example, partial vapor pressures, solubilities, and partition coefficients. An additional advantage of COSMO-RS over empirical methods is that the molecules dissolved may in fact be transition states of a chemical reaction. This follows from the fact that all that is required is that one can do a QM calculation on the solute and QM on a transition state has become standard in the last two decades. This affords a unique opportunity to predict the thermodynamics of a reaction including, for instance, the balance between kinetically and thermodynamically favored reaction pathways as a function of the solvent used.

COSMO-RS 2010.01

The major changes of COSMO-RS 2010.01 in comparison to COSMO-RS 2009.01 are described here.

- Fast approximation introduced for COSMO-RS calculations
  - 1-dimensional treatment of charge densities σ_v and $\sigma_v^{\perp}$ instead of 2-dimensional
- COSMO-RS database ADFCRS-2010
  - 1892 precalculated compounds
  - mostly solvents and small molecules
- The recommended settings to produce ADF COSMO result files are now:
  - TZP small core basis set (for all $Z \leq 36$)
  - TZ2P small core basis set basis set for Iodine (for all $Z \geq 37$)
  - the Becke Perdew exchange-correlation functional
  - the relativistic scalar ZORA Hamiltonian
  - integration accuracy 6
- COSMO cavity construction numerically made more stable:
  - merge close lying COSMO surface points

- ◦ reduced dependence results on coincidental close lying integration points with COSMO surface points
- ADF-GUI: easier to use recommended settings to produce COSMO result files:
    - ◦ gas phase molecule: ADFinput → Main Options → Preset → Gas Phase CRS
    - ◦ solvated molecule: ADFinput → Main Options → Preset → Solvent CRS
- COSMO-RS GUI:
    - ◦ improved for easier usage with many compounds
    - ◦ show molecule with ADFview: ADFcrs → View → Show Selected Compound
    - ◦ .coskf and .t21 files can be view with ADFview
   - ◦ more examples

COSMO-RS

Theory

Below some of the COSMO-RS theory is explained, but a more complete description can be found in Refs.[2] and [3].

Although in principle all of chemistry can be predicted by appropriate solutions of the Schrödinger Equation, in practice due to the extreme mathematical complexity of doing so only the smallest systems can be computed at an accuracy rivaling that of the most accurate experiments. However, with suitable approximations, for isolated molecules of up to a few hundred atoms these days quite reasonable results can be obtained. Of course, this means that direct computation of thermodynamic properties is out of reach. Thermodynamic properties can only be computed as an average over a large number of configurations of a large number of molecules. To address this, people have typically resorted to so-called Molecular Dynamics (MD) or Monte Carlo (MC) methods where configurations are generated either by numerically simulating the atomic motions over discrete time steps or by random generation, in either case using empirical molecular models parametrized against quantum mechanical calculations and experimental data to compute energies. However, even these approaches often fall short in generating sufficiently large ensembles, and there is little chance of that situation improving dramatically in the near future.

Around 1995, Andreas Klamt, then working for Bayer, hit upon an approach that made it possible to compute the details of molecules quantum mechanically and subsequently use these details in an approximate statistical mechanics procedure [1]. This approach is called COSMO-RS (COnductor like Screening MOdel for Realistic Solvents) and has proven to be quite powerful. It may currently the best link between the world of chemical quantum mechanics and engineering thermodynamics.

Thermodynamic reference states can be chosen arbitrarily. They do not even have to be physically realizable, as long as it is consistently used. We are at liberty to choose as reference state a molecule embedded in a perfect conductor, that is a material with an infinitely large dielectric constant ('the perfectly screened state'). Suppose a molecule A resides in a molecule shaped cavity. Everywhere outside of this cavity is conductor material. Although it would be hard to realize this in practice, it is relatively easy to do quantum mechanical calculations on this hypothetical state. Since the molecule will in general have a charge distribution and therefore possess an electric field, it will polarize the embedding medium. That will result in another electric field, given by a charge distribution on the surface of the molecule shaped cavity. This charge distribution is generated by the quantum mechanical calculations, for example with ADF if one uses COSMO. From now on the surface of the molecule shaped cavity will be called molecular surface, and the volume of the molecule shaped cavity will be called molecular volume.
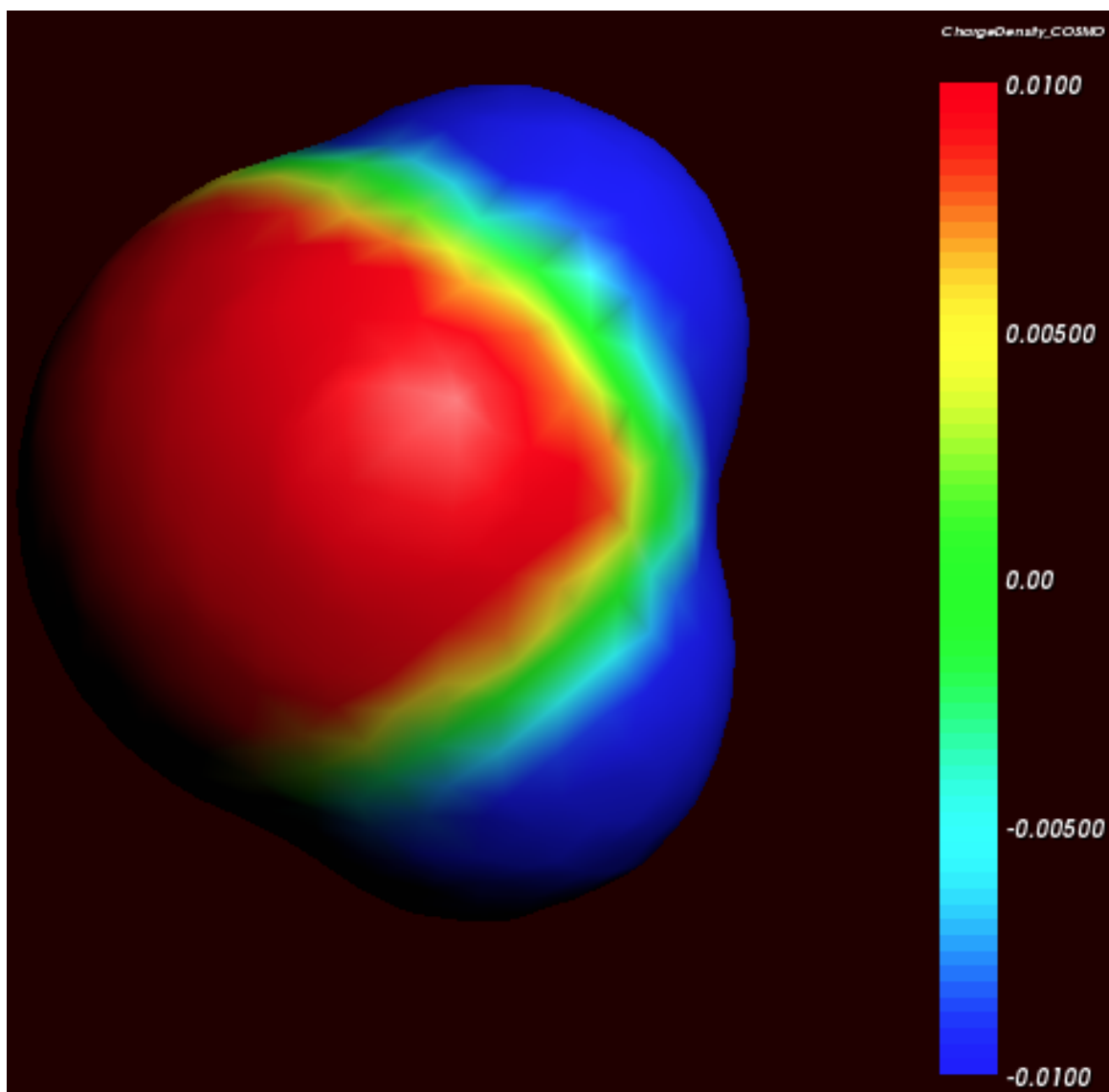
Cosmo charge density on the COSMO surface of water (picture made with ADFview)

Although the actual charge distribution on the molecular surface will be highly detailed, let us for the moment consider the molecular surface as consisting of segments with a constant charge density (i.e. the detailed charge distribution averaged over segments). Now instead of the single molecule A consider, as an arbitrary example, a fluid consisting of three types of molecules: A, B and C. In a fluid not too close too the critical temperature, the molecular surfaces present in the fluid will all be in close contact. That means that the segments of constant density introduced above are in close contact.

We now compare our molecule A in the fluid with our chosen reference state. Any segment of the molecular surface with a charge density of σ_i will be aligned with a segment with charge density σ_j of another molecule. If the two charge densities happen to be opposite (i.e. $\sigma_i + \sigma_j = 0$) the charges required for achieving the perfectly screened state will vanish. However, this will not happen too often and in general an excess charge density is left of $\sigma_i + \sigma_j$ between the two segments. From electrostatic theory it follows that this introduces an energy penalty proportional to the segment size and $(\sigma_i + \sigma_j)^2$. In principle this gives a way to compute the chemical potential of component A, by going over all possible conformations of a large number of molecules A, B and C (in their proper molar fractions) and do computations on the statistical ensemble. However, in practice that would be similar to doing Molecular Dynamics calculations using empirical structure models

and about as computationally prohibitive. Instead, an approximation can be made that is not easily justified a priori and must be judged by the results of subsequent simulations. This assumption is that all segments in the fluid are able to make contact independent of one another. In a way it can be said, that the segments are cut loose from the original (rigid) molecular surfaces.

As one would guess, the approximation of independent segments makes the mathematics of computing ensemble properties quite tractable. In fact, computing the chemical potential of component A (or B or C) in the mixture by means of the COSMO-RS and related methods takes in the order of seconds on a normal PC (given the results of quantum mechanical calculations that may have taken days, of course). Note that the molecular surface around the molecule is divided rather arbitrarily in segments and that the assumption was that the segment of one molecule will overlap perfectly with that of another. How can this be true? The answer is that one can split up the molecular surface into segments in an infinite number of ways. However, the molecules in a fluid are always in contact with another. At any given time, molecule A will be in contact with a number of other molecules and share patches of, for example, 7 square Angstroms of its surface with each of the surrounding molecules. At that particular time, the segments will be those patches. A split second later, of course, there will be a different set of segments. That is not a problem. One needs to do statistical mechanics with charged segments for which one needs to know how many 7 square Angstrom segments a particular molecule brings into the fluid and the probability of any segment having an average charge density σ (for all values of σ). Both can be computed from the results of the quantum mechanical calculation on the molecule in the perfect conductor. Just to get a flavor, in the figure below the so-called σ-profile of water is given. These are the statistical distributions of possible segments over charge densities multiplied by the surface area of the molecular volume. The σ-profile relates to the detailed charge distribution on the molecular surface.
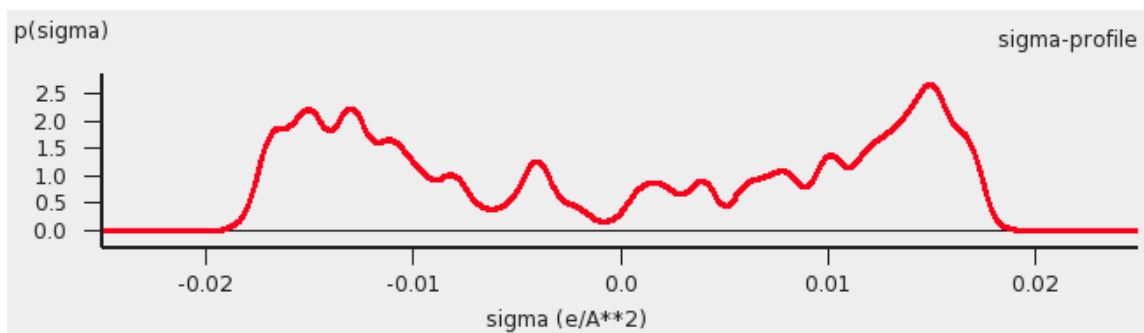
σ-profile of water (picture made with the CRS-GUI), smoothed curve, Delley COSMO surface construction

In principle vapor pressures of pure liquids can be computed directly with COSMO-RS. COSMO-RS calculations yield the chemical potential of a component in a liquid with respect to the perfectly screened reference state. It is easy to compute the energy difference between the reference state and the gas phase by doing an additional quantum mechanical calculation (of the isolated molecule). However, often experimental vapor pressures for the pure liquid are known. Using such experimental data for pure liquids can help in predicting the correct partial vapor pressures in a mixture.

Combinatorial term

In Ref.[2] a thermodynamically inconsistent combinatorial contribution μ_i^{comb} to the chemical potential was used:

$$\mu_i^{comb} = - \lambda RT \ln (q_{av}/Angstrom^2)$$
$$q_{av} = \sum_i x_i q_i$$

In this equation q_i is the surface area of the molecular volume of compound i, x_i is the molar fraction of compound i in the solution, and λ is a COSMO-RS parameter.

The importance of using a thermodynamically consistent combinatorial contribution is discussed in Ref. [3]. In the ADF COSMO-RS program it is possible to use a thermodynamically consistent combinatorial contribution of the form (Equation C.4 of Ref.[3], with $\lambda_0=\lambda_1=\lambda_2=\lambda$):

$$\mu_i^{comb} = \lambda RT \ (1 - r_i/r_{av} + \ln(r_i/r_{av}) + 1 - q_i/q_{av} - \ln \ (q_{av}/Angstrom^2))$$
$$r_{av} \quad = \sum_i x_i \ r_i$$

In this equation r_i is the molecular volume of compound i. In the ADF COSMO-RS program this combinatorial term is used by default, see also Ref. [4].

Temperature dependent hydrogen bond interaction

In Ref.[3] a temperature dependent hydrogen bond interaction is suggested, which is used by default in the ADF COSMO-RS program. The temperature dependence (Equation 6.2 of Ref.[3]) is of the form:

$$term \ (T) = T \ \ln[1+\exp(20 \ kJ/mol/RT)/200]$$
$$f_{hb} \ (T) \quad = term(T)/term(298.15 \ K)$$

In this equation R is the gas constant and T the temperature (in Kelvin). In the ADF COSMO-RS program the hydrogen bond interaction of Ref.[2] is multiplied by this factor f_{hb} (T) to make the hydrogen bond interaction temperature dependent.

Fast approximation for COSMO-RS calculations

In the 1998 COSMO-RS model each segment of the molecular surface has a charge density of σ_v, but also a second charge density $\sigma_v^{\perp}$, which is a descriptor for the correlation between the charge density on the segment with its surrounding. In the original ADF COSMO-RS implementation this was treated as a 2-dimensional problem, in the fast approximation this is effectively reduced to 1-dimension. Starting from COSMO-RS 2010 this fast approximation is now the default. This approximation reduces the computation time, especially in cases of more than 1 compound.

Calculation of properties

The COSMO-RS method allows to calculate the chemical potential of a compound in the liquid phase, as well as in the gas phase, see the the COSMO-RS theory that was discussed before and Ref.[2]. In the ADF COSMO-RS implementation the following equations were used to calculate properties using these chemical potentials.

$$\sum_i x_i \quad = 1$$
$$p_i^{pure} \quad = \exp \ \{(\mu_i^{pure}-\mu_i^{gas})/RT\}$$
$$p_i^{vapor} = x_i \ \exp \ \{(\mu_i^{solv}-\mu_i^{gas})/RT\}$$
$$p^{vapor} \quad = \sum_i p_i^{vapor}$$
$$y_i^{vapor} = p_i^{vapor}/p^{vapor}$$
$$\gamma_i \qquad = \exp \ \{(\mu_i^{solv}-\mu_i^{pure})/RT\}$$
$$a_i \qquad = \gamma_i \ x_i$$
$$G^E \qquad = \sum_i x_i \ (\mu_i^{solv}-\mu_i^{pure})$$
$$H^E \qquad = \sum_i x_i \ (E_i^{HB}+E_i^{misfit}-E_i^{HB \ pure}-E_i^{misfit \ pure})$$
$$k_H \qquad = 1/V_{solvent} \ \exp \ \{(\mu_i^{gas}-\mu_i^{solv})/RT\}$$
$$k_H^{cc} \quad = k_H \ RT$$

$$x_i^{SOL} = 1/\gamma_i \quad (T > T_m)$$

$$x_i^{SOL} = 1/\gamma_i \exp\{\Delta H_{fus}(1/T_m - 1/T)/R - \Delta C_p(\ln(T_m/T) - T_m/T + 1)/R\} \quad (T < T_m)$$

$$\Delta G_{solv}^{liq\text{-}solv} = \mu_i^{solv} - \mu_i^{pure}$$

$$\Delta G_{solv}^{gas\text{-}solv} = \mu_i^{solv} - \mu_i^{gas} + RT \ln(V_{solvent}/V_{gas})$$

$$\log_{10}P_{solv1/solv2} = 1/\ln(10)\,(\mu_i^{solv2} - \mu_i^{solv1})/RT + \log_{10}(V_{solv2}/V_{solv1})$$

The above equations are not always exact, some assume ideal gas behavior, for example.

The molar fraction x_i of each compound i of the solvent should add up to 1.

With the COSMO-RS method it is possible to predict vapor pressures. However, it is also possible to use experimental vapor pressures of pure compounds as input data for the calculation. This may increase the accuracy of the calculated vapor pressures in a mixture, for example.

In the COSMO-RS method the volume of 1 molecule in the liquid phase is approximated with the volume of the molecule shaped cavity, that is used in the COSMO calculations. In this way it is possible to calculate the volume of 1 mole of solvent molecules in the liquid phase. However, for properties that depend on such volumes, one can also use (related) experimental data as input data for the calculation.

The calculation of the boiling temperature of a solvent is performed with an iterative method. The temperature is varied until the calculated vapor pressure is within a certain threshold of the desired pressure.

Also the calculation of solubility of compound i is performed with an iterative method, since the activity coefficient γ_i depends on the molar fraction of this compound. The COSMO-RS method does not predict ΔH_{fus}, ΔC_p, or T_m. These can be given as input data for the calculation of solubility calculations of solid compounds.

Quantity	Meaning
R	Gas constant
T	Temperature
μ_i^{gas}	The chemical potential of the pure compound i in the gas phase
μ_i^{pure}	The pseudochemical potential of the pure compound i in the liquid phase
μ_i^{solv}	The pseudochemical potential of compound i in a liquid solution
x_i	The molar fraction of compound i in a liquid solution
y_i	The molar fraction of compound i in the gas phase
γ_i	Activity coefficient of compound i in a liquid solution
a_i	Activity of compound i in a liquid solution
p_i^{pure}	The vapor pressure of the pure compound i
p_i^{vapor}	The partial vapor pressure of compound i
p^{vapor}	The total vapor pressure
G^E	The excess Gibbs free energy
H^E	The excess enthalpy
$E_i^{HB\ pure}$	The hydrogen bond energy of the pure compound i in the liquid phase, see Ref.[2]
E_i^{HB}	The partial hydrogen bond energy of compound i in a liquid solution
$E_i^{misfit\ pure}$	The misfit energy of the pure compound i in the liquid phase, see Ref.[2]
E_i^{misfit}	The partial misfit energy of compound i in a liquid solution
k_H	Henry's law constant: ratio between the liquid phase concentration of a compound and its partial vapor pressure in the gas phase
k_H^{cc}	dimensionless Henry's law constant: ratio between the liquid phase concentration of a compound and its gas phase concentration

$\Delta G_{solv}^{liq-solv}$	The solvation Gibbs free energy from the pure compound liquid phase to the solvated phase
$\Delta G_{solv}^{gas-solv}$	The solvation Gibbs free energy from the pure compound gas phase to the solvated phase, with a reference state of 1 mol/L in both phases
$V_{solvent}$	Volume of 1 mole of solvent molecules in the liquid phase
V_{gas}	Volume of 1 mole of molecules in the gas phase (at 1 atm, ideal gas)
x_i^{SOL}	Solubility of compound i in a solvent (molar fraction)
ΔH_{fus}	The enthalpy of fusion of compound i
ΔC_p	The Δ heat capacity of fusion of compound i
T_m	The melting temperature of compound i
$\log_{10}P_{solv1/solv2}$	The logarithm of the partition coefficient P, which is the ratio of the concentrations of a compound in two immiscible solvents, solvent 1 and solvent 2

COSMO result files

COSMO-RS needs as input for the calculation so called COSMO result files for each compound, which are results of quantum mechanical calculation using COSMO. In ADF such a COSMO result file is called a TAPE21 (.t21) file or a COSKF file (.coskf). In other programs such a file can be a .cosmo file. For example, at http://www.design.che.vt.edu/VT-Databases.html a database of .cosmo files can be found, which were made with a different program. Note that the optimal COSMO-RS parameters may depend on the program chosen.

ADF COSMO calculation

Here it is described briefly how to make COSMO result files consistent with the way they were made for the ADF parametrization of COSMO-RS to ensure full parameter applicability. First a gas phase geometry optimization should be performed with ADF, with a small core TZP basis set, the Becke-Perdew functional, the relativistic scalar ZORA method, and an integration accuracy of 6:

```
BASIS
type TZP
core Small
I ZORA/TZ2P/I.4p
END
XC
GGA Becke Perdew
END
INTEGRATION 6 6 6
relativistic scalar zora
```

For heavy elements like iodine, a small core TZ2P basis set is required. The resulting TAPE21 of the molecule (rename it compound_gasphase.t21) is used as a restart file in the COSMO calculation. The ADF COSMO calculation is performed with the following settings:

```
SYMMETRY NOSYM
SOLVATION
    Surf Delley
    Solvent name=CRS emp=0.0 cav0=0.0 cav1=0.0
    Charged method=CONJ corr
    C-Mat EXACT
    SCF VAR ALL
    RADII
      H  1.30
      C  2.00
      N  1.83
      O  1.72
      F  1.72
      Si 2.48
      P  2.13
      S  2.16
      Cl 2.05
      Br 2.16
      I  2.32
    SubEnd
END
XC
```

```
GGA Becke Perdew
END
INTEGRATION 6 6 6
relativistic scalar zora
RESTART compound_gasphase.t21
```

In this COSMO calculation the non-default Delley type of cavity construction is chosen (See Ref.[5] for details on the Delley surface construction). The name of the solvent is CRS, which sets the dielectric constant to infinite and sets the radius of the probing sphere to determine the solvent excluded part of the surface to 1.3 Angstrom. In the Radii subblock key the Klamt atomic cavity radii are chosen. The parameters `emp`, `cav0`, and `cav1` are zero. The `corr` option to the CHARGED subkey constrains the computed solvent surface charges to add up to the negative of the molecular charge. Specifying `exact` for the C-MAT subkey causes ADF to compute straightforwardly the Coulomb potential due to the charge q in each point of the molecular numerical integration grid and integrate against the electronic charge density. This is, in principle, exact but may have inaccuracies when the numerical integration points are very close to the positions of a charge q. To remedy this, in ADF2010 the electrostatic potential is damped in case of (very) close lying numerical integration points and COSMO surface points. The numerical stability of the results compare to those of ADF2009 was increased as a result of this. Specifying `exact` for the C-MAT subkey also requires that the ADF calculation uses SYMMETRY NOSYM.

The resulting TAPE21 (rename it compound.t21) of the COSMO calculation is a COSMO result file.

In a COSMO-RS calculation only the 'COSMO' part of this file is needed. One can make a kf file compound.coskf, which only consists of the section 'COSMO' if one does:

```
$ADFBIN/cpkf compound.t21 compound.coskf "COSMO"
```

The file compound.coskf should not exist before this command is given. Note that such a .coskf file is not a complete TAPE21 anymore, and can not be viewed with ADFview, for example. It is only useful for COSMO-RS calculations.

Accuracy

Several parameters in the COSMO calculation can influence the accuracy of the result of the quantum mechanical calculation. Some of these parameters will be discussed. Note that if one chooses different parameters in the COSMO calculation one may also have to reparametrize the ADF COSMO-RS parameters. A list of some of the ADF COSMO parameters.

- XC functional
- basis set
- fit set
- atomic cavity radii and radius of the probing sphere
- cavity construction
- geometry

The atomic cavity radii and the radius of the probing sphere are the same as in Ref. [2], which describes the COSMO-RS method developed by Klamt et al., which is implemented in ADF. The Becke Perdew functional is relatively good for weakly bound systems, but may not be so good in other cases. The basis set TZP is a compromise basis set. For heavy elements like iodine, a TZ2P basis set is required, including the relativistic scalar ZORA method. Since the relativistic method hardly cost extra time compared to a non-relativistic method, the scalar relativistic scalar ZORA method is recommended to be used also for light elements. The Delley type of cavity construction in ADF can give a large number of COSMO points. The XC functional, basis set, and cavity construction chosen in the ADF COSMO calculation have a similar accuracy as those

that were used in Ref. [2]. Note that they are not exactly the same as were used in Ref. [2], since in that paper a different quantum mechanical program was used.

In the parametrization for ADF the same geometry was used for the gas phase and the COSMO calculation, which is different than in Ref. [2]. It depends on the actual solvent if reoptimizing the molecule in the COSMO calculation may give better results. Note that the dielectric medium used in the COSMO model has an infinite dielectric constant in the COSMO-RS model. Thus a geometry optimization of the molecule in the COSMO calculation might be more appropriate for a molecule dissolved in water than for a molecule dissolved in n-hexane.

The fit set in ADF is not always able to describe the Coulomb potential accurately at each of the COSMO surface points. In regular ADF calculations this problem is not apparent since the numerical errors in the integrals computed in the vicinity of the COSMO surface have little impact. However, in COSMO calculations this may have some effect. This is why the option `C-Mat exact` was selected above, instead of the default `C-Mat pot` option. Another possibility is to add more fit functions, for example, using the ADDDIFFUSEFITFUNCTIONS key in the input for the adf calculation.

Cavity construction

The Esurf type of cavity construction in ADF with default settings does not give a large number of COSMO points. Therefore it is recommended to use the so called Delley type of cavity construction (Ref.[5]), which allows one to construct a surface which has many more points. The Esurf type of cavity construction also allows many more points if one sets the option NFDiv of the subkey DIV of the key SOLVENT to a larger value than the default value of 1. This will not be discussed here further. In ADF2010 the numerical stability of the Delley surface has been improved, by merging close lying COSMO surface points, and removing COSMO surface points with a small surface area. A figure of a COSMO surface with the Esurf type of cavity construction with default settings is given below. In this figure the small spheres represent the COSMO surface points that are used for the construction of the COSMO surface. The red part represents positive COSMO charge density, the blue part negative COSMO charge density (the coloring scheme is chosen to match the one by Klamt):
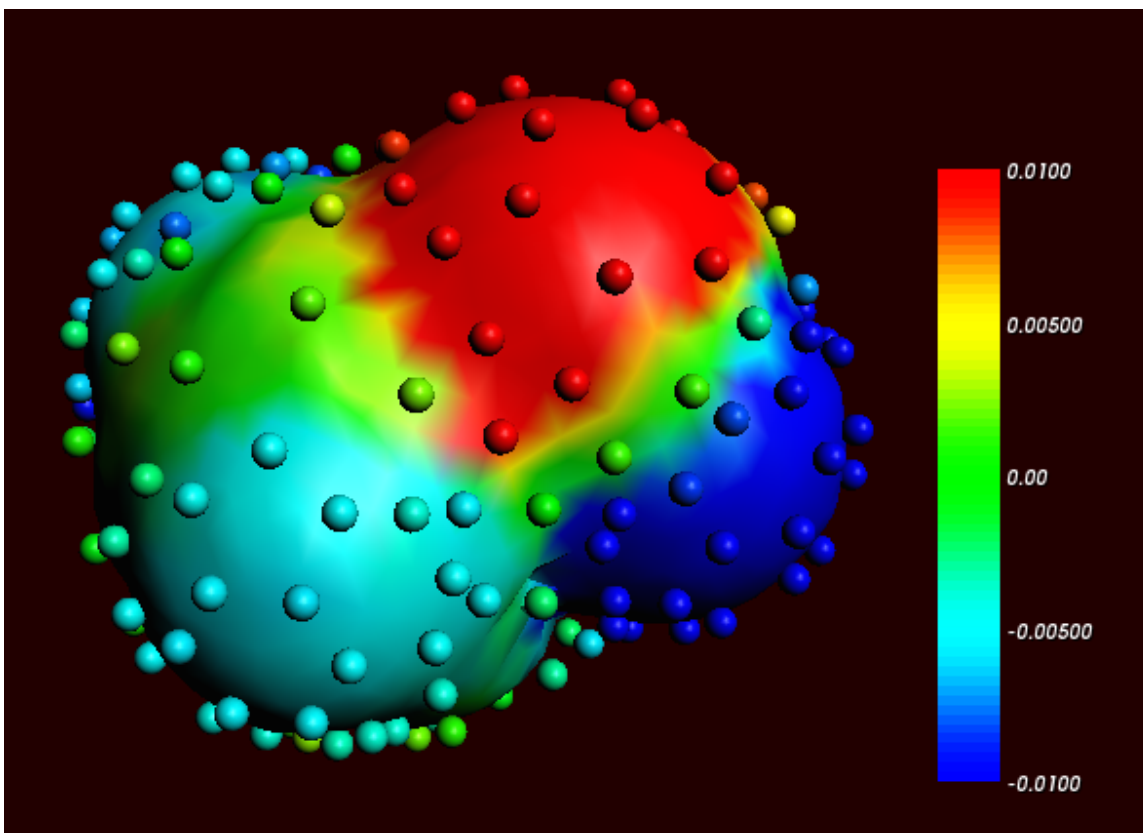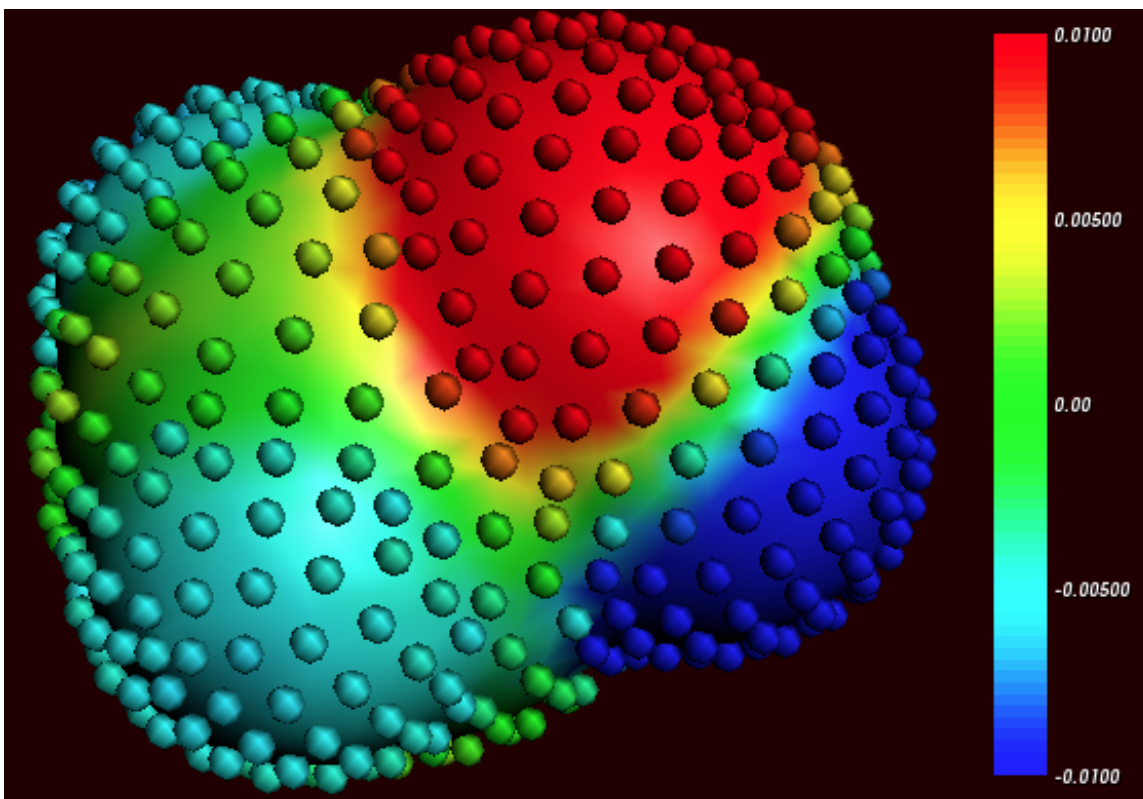
Cosmo charge density on the COSMO surface of methanol, Esurf surface (picture made with ADFview)

One can construct a surface which has many more points using a so called Delley surface. For the subkey SURF of the key SOLVENT one can choose `delley`. The subkey DIV of the key SOLVENT has extra options `leb1` (default value 23), `leb2` (default value 29), and `rleb` (default value 1.5 Angstrom). If the cavity radius of an atom is lower than `rleb` use `leb1`, otherwise use `leb2`. These values can be changed: using a higher value for `leb1` and `leb2` gives more surface points (maximal value `leb1`, `leb2` is 29). A value of 23 means 194 surface points in case of a single atom, and 29 means 302 surface points in case of a single atom Typically one could use `leb1` for the surface point of H, and `leb2` for the surface points of other elements.

The next figure is made with the following (default for the Delley surface) settings:

```
SOLVATION
    SURF Delley
    DIV leb1=23 leb2=29 rleb=1.5
    ....
END
```

Cosmo charge density on the COSMO surface of methanol, Delley surface (picture made with ADFview)

The different ways of constructing the cavity has some consequences for the σ-profile of methanol, see the figure below:
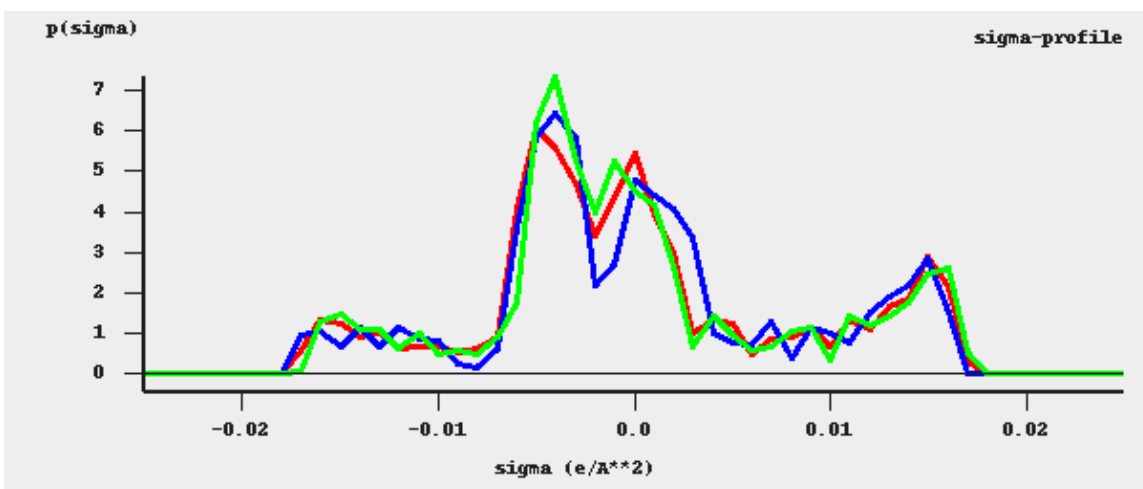
σ-profiles of methanol (picture made with the CRS-GUI)

In this picture the blue line is the σ-profile with the Esurf type of construction, the red line is that with the Delley type of construction with many surface points. For comparison, the green line is the σ-profile of methanol if a large QZ4P basis set is used, again with the Delley type of construction with many surface points.

The COSMO-RS program

The ADF COSMO-RS command line program *crs* is described here, including all input options.

Running the COSMO-RS program

Running the COSMO-RS program involves the following steps:

- Construct an ASCII input file, say in.
- Run the program by typing (under UNIX):
  ```
  $ADFBIN/crs <in >out
  ```
- Move / copy relevant result files (in particular CRSKF) to the directory where you want to save them,
  and give them appropriate names.
- Inspect the standard output file out to verify that all has gone well.

COSMO-RS parameters

The COSMO-RS model has several parameters. The general COSMO-RS parameters can be given a value with the key `CRSPARAMETERS` and the element specific parameters can be given a value with `DISPERSION`.

General parameters

```
CRSPARAMETERS
   {RAV         rav}
   {APRIME      aprime}
   {FCORR       fcorr}
   {CHB         chb}
   {SIGMAHBOND  sigmahbond}
   {AEFF        aeff}
   {LAMBDA      lambda}
   {OMEGA       omega}
   {ETA         eta}
   {CHORTF      chortf}
   {combi1998 | combi2005}
   {hb_all    | hb_hnof}
   {hb_temp   | hb_notemp}
   {fast      | nofast}
End
```

The ADF default values are optimized parameters for ADF calculations. The Klamt values can be found in Ref. [2]. See also Ref. [2] for the meaning of the parameters.

symbol	ADF Default Ref. [4]	ADF combi1998 Ref. [4]	Klamt Ref. [2]
rav (r_{av})	0.400	0.415	0.5
aprime (a')	1510.0	1515.0	1288.0
fcorr (f_{corr})	2.802	2.812	2.4
chb (c_{hb})	8850.0	8850.0	7400.0
sigmahbond (σ_{hb})	0.00854	0.00849	0.0082

aeff (a_{eff})	6.94	7.62	7.1
lambda (λ)	0.130	0.129	0.14
omega (ω)	-0.212	-0.217	-0.21
eta (η)	-9.65	-9.91	-9.15
chortf ($c^{\perp}$)	0.816	0.816	0.816
combi1998 | combi2005	combi2005	combi1998	combi1998
hb_all | hb_hnof	hb_hnof	hb_hnof	hb_hnof
hb_temp | hb_notemp	hb_temp	hb_notemp	hb_notemp
fast | nofast	fast	fast	fast

`chortf`

See Ref. [2] for the definitions: $\sigma_v^{\perp} = \sigma_v^0 - c^{\perp} \sigma_v$

`combi1998 | combi2005`

If the subkey `combi1998` is included a thermodynamically inconsistent combinatorial contribution to the chemical potential μ_i^{comb} of Ref.[2] is used. If the subkey `combi2005` is included (default) a thermodynamically consistent combinatorial contribution of Ref.[3] is used. See the section on the combinatorial term and Ref.[3].

`hb_all | hb_hnof`

If the subkey `hb_all` is included hydrogen bond interaction can be included between segments that belong to H atoms and all other segments. If the subkey `hb_hbnof` is included (default) hydrogen bond interaction can be included only between segments that belong to H atoms that are bonded to N, O, or, F, and segments that belong to N, O, or F atoms.

`hb_temp | hb_notemp`

If the subkey `hb_notemp` is included the hydrogen bond interaction is not temperature dependent, as in Ref.[2]. If the subkey `hb_temp` is included (default) the hydrogen bond interaction is temperature dependent, as in Ref.[3]. See the section on the temperature dependent hydrogen bond interaction and Ref.[3].

`fast | nofast`

If the subkey `fast` is included the fast approximation is used. This fast approximation is the default. Use `nofast` for the original approach. See the section on the fast approximation for COSMO-RS calculations.

Element specific parameters

```
DISPERSION
   {H   dispH}
   {C   dispC}
   {N   dispN}
   {...}
End
```

The following table gives the element specific dispersion constants. The ADF default values are optimized parameters for ADF calculations. The Klamt values can again be found in Ref. [2]. The constants for F, Si,

P, S, Br, and I in the ADF defaults were only fitted to a small number of experimental values or taken from Ref. [3].

element	ADF Default	ADF combi98	Klamt (Ref.[2])
H	-0.0340	-0.0346	-0.041
C	-0.0356	-0.0356	-0.037
N	-0.0224	-0.0225	-0.027
O	-0.0333	-0.0322	-0.042
Cl	-0.0485	-0.0487	-0.052
F	-0.026		
Si	-0.04		
P	-0.045		
S	-0.052		
Br	-0.055		
I	-0.062		

Note that not for all elements COSMO-RS parameters were fitted.

Compounds

For each compound one has to add the keyword `COMPOUND`

```
COMPOUND filename
{cosmofile}
{drophbond}
{NRING nring}
{FRAC1 frac1}
{FRAC2 frac2}
{PVAP pvap}
{TVAP tvap}
{Antoine A B C}
{MELTINGPOINT meltingpoint}
{HFUSION hfusion}
{CPFUSION cpfusion}
End
```

`filename`

   The filename (can be a full path, otherwise relative path is assumed) should be a COSMO result file. How to make an ADF COSMO result file can be found here.

`cosmofile`

   If the subkey `cosmofile` is included the file should be an ASCII COSMO file (.cosmo). If not specified (default) the file should be a kf file, either an ADF COSMO result file TAPE21 (.t21) or a COSKF file (.coskf).

`drophbond`

   If the subkey `drophbond` is included no hydrogen-bond terms will be included for this compound. If not specified (default) the hydrogen-bond terms are included for this compound.

`nring`

18

The number of ring atoms. This is a COSMO-RS parameter. It should be 6 for benzene, for example. Default value is 0.

`frac1`

The molar fraction of the compound in the solvent. This is solvent 1 in case of the calculation of partition coefficients.

`frac2`

The molar fraction of solvent 2, only used in case of the calculation of partition coefficients.

`pvap, tvap`

Pure compound vapor pressure `pvap` (bar) at temperature `tvap` (Kelvin). Used only if both `pvap` and `tvap` are specified, and then will have an effect on the calculated vapor pressures or boiling points. Alternative is to set the Antoine coefficients. If both are not specified the pure compound vapor pressure is approximated using the COSMO-RS method.

`A, B, C`

`A`, `B`, and `C` are the pure compound Antoine coefficients, such that:
`log P = A - B/(T+C)`
This Antoine equation is a 3-parameter fit to experimental pure compound vapor pressures P (bar) over a restricted temperature T (Kelvin) range. If the Antoine coefficients are specified this will have an effect on the calculated vapor pressures or boiling points. Alternative is to give input values for the pure compound vapor pressure at a fixed temperature. If both are not specified the pure compound vapor pressure is approximated using the COSMO-RS method.

`meltingpoint, hfusion, cpfusion`

Pure compound melting point `meltingpoint` (Kelvin), pure compound enthalpy of fusion `hfusion` (kcal/mol), and pure compound heat capacity of fusion `cpfusion` (kcal/(mol K)). Only used if both `meltingpoint` and `hfusion` are specified (`cpfusion` optional), and will then have an effect in solubility calculations if the temperature of the solvent is below the melting point.

Temperature

```
TEMPERATURE temperature {temperature_high ntemp}
```

`temperature`

Temperature (Kelvin) at which temperature the COSMO-RS calculation should take place. Default room temperature 298.15. The first temperature in case of a range of temperatures.

`temperature_high`

The last temperature (Kelvin) in case of a range of temperatures. Only used in case of solvent vapor pressure calculations or solubility calculations.

`ntemp`

The number of temperatures in case of a range of temperatures.

Pressure

```
PRESSURE pressure {pressure_high npress}
```

`pressure`

Pressure (bar) at which pressure the COSMO-RS calculation should take place. Default 1.01325 bar (1 atm). The first pressure in case of a range of pressures.

`pressure_high`

The last pressure (bar) in case of a range of pressures. Only used in case of solvent boiling point calculations.

`npress`

The number of pressures in case of a range of pressures.

Properties

Solvent vapor pressure

The mole fraction of each compound of the solvent should be given with the subkey `FRAC1` of the key `COMPOUND` for each compound. It is possible to calculate the vapor pressure for a temperature range, see key `TEMPERATURE`.

```
PROPERTY vaporpressure
End
```

The input pure compound vapor pressure will be used in the calculation of the partial vapor pressure of this compound in the mixture if it is supplied with the key `COMPOUND` for this compound. If it is not specified then it will be approximated using the COSMO-RS method.

Solvent boiling temperature

The mole fraction of each compound of the solvent should be given with the subkey `FRAC1` of the key `COMPOUND` for each compound. It is possible to calculate the boiling temperature for a pressure range, see key `PRESSURE`.

```
PROPERTY boilingpoint
End
```

The input pure compound vapor pressure will be used in the calculation of the partial vapor pressure of this compound in the mixture if it is supplied with the key `COMPOUND` for this compound. If it is not specified then it will be approximated using the COSMO-RS method.

Log partition coefficients solute

The mole fraction of each compound of the solvent 1 and solvent 2 should be given with the subkey `FRAC1` and subkey `FRAC2` of the key `COMPOUND` for each compound, respectively. The partition coefficients are calculated for all compounds.

```
PROPERTY logp
{VOLUMEQUOTIENT volumequotient}
End
```

`volumequotient`

> If the subkey `VOLUMEQUOTIENT` is included the `volumequotient` will be used for quotient of the densities of solvent 1 and solvent 2 instead of calculated values.

Activity coefficients solute

The mole fraction of each compound of the solvent should be given with the subkey `FRAC1` of the key `COMPOUND` for each compound. The activity coefficients are calculated for all compounds.

```
PROPERTY activitycoef
{DENSITYSOLVENT densitysolvent}
End
```

`densitysolvent`

> If the subkey `DENSITYSOLVENT` is included the `densitysolvent` will be used for the density of the solvent (kg/L) instead of calculated values. Relevant for the calculation of the Henry's law constant.

The input pure compound vapor pressure will be used in the calculation of the partial vapor pressure of this compound in the mixture if it is supplied with the key `COMPOUND` for this compound. If it is not specified then it will be approximated using the COSMO-RS method. Relevant for the calculation of the Henry's law constant.

Solubility solute in solvent

The mole fraction of each compound of the solvent should be given with the subkey `FRAC1` of the key `COMPOUND` for each compound. The solutes should have zero molar fraction in the solvent. It is possible to calculate the solubility of a solute at a temperature range, see key `TEMPERATURE`.

```
PROPERTY solubility
End
```

For solubility calculations of a solid compound one should add the pure compound melting point, pure compound enthalpy of fusion, and optionally the pure compound heat capacity of fusion using the subkeys `meltingpoint`, `hfusion`, and `cpfusion`, respectively, of the key `COMPOUND` for this compound.

Binary mixture

Exactly two compounds should be given in the input file.

```
PROPERTY binmixcoef
{Nfrac nfrac}
{isotherm | isobar}
End
```

`nfrac`

Number of different mixtures for which the binary mixture is calculated. default value 10.

```
isotherm | isobar
```

If the subkey `isotherm` is included (default) the binary mixture will be calculated at a fixed temperature. If the subkey `isobar` is included the binary mixture will be calculated at a fixed vapor pressure.

The input pure compound vapor pressure will be used in the calculation of the partial vapor pressure of this compound in the mixture if it is supplied with the key `COMPOUND` for this compound. If it is not specified then it will be approximated using the COSMO-RS method.

Analysis

Sigma-profile

```
PROPERTY sigmaprofile
{Nprofile nprofile}
{SigmaMax sigmamax}
{pure | mixture}
End
```

```
nprofile
```

Number of data points for which to calculate the sigma-profile. default value 50.

```
sigmamax
```

The sigma-profile is calculated between -sigmamax and sigmamax. Default value 0.025.

```
pure | mixture
```

If the subkey `pure` is included the pure compound sigma-profiles are calculated. If the subkey `mixture` is included the sigma-profile of the mixture is calculated. The mole fraction of each compound in the mixture should be given with the subkey `FRAC1` of the key `COMPOUND` for this compound.

Sigma-potential

```
PROPERTY sigmapotential
{Nprofile nprofile}
{SigmaMax sigmamax}
{pure | mixture}
End
```

```
nprofile
```

Number of data points for which to calculate the sigma-potential. default value 50.

```
sigmamax
```

The sigma-potential is calculated between -sigmamax and sigmamax. Default value 0.025.

```
pure | mixture
```

If the subkey `pure` is included the pure compound sigma-potentials are calculated. If the subkey `mixture` is included the sigma-potential of the mixture is calculated. The mole fraction of each compound in the mixture should be given with the subkey `FRAC1` of the key `COMPOUND` for this compound.

COSMO-RS command line utilities

The two COSMO-RS command line utility programs *kf2cosmo* and *cosmo2kf* convert COSMO kf files from binary to ASCII and vice versa.

```
kf2cosmo file.coskf file.cosmo
```

kf2cosmo reads from the kf file file.coskf (should exist) the section 'COSMO' and writes to the ASCII file file.cosmo (should not exist). Instead of a .coskf file, the file can also be a TAPE21 file which is a result file from an ADF COSMO calculation.

```
cosmo2kf file.cosmo file.coskf
```

cosmo2kf reads from the ASCII file file.cosmo (should exist) and writes a section 'COSMO' to the kf file file.coskf (should not exist). Note that only a section 'COSMO' is written to the kf file, which means that this file can not be used like an ordinary TAPE21 (.t21) file.

```
cpkf file.t21 file.coskf COSMO
```

With the ADF kf utility *cpkf* one can copy the section 'COSMO' from a file.t21 (should exist) to a file.coskf (should not exist). The file file.t21 should be a result of an ADF COSMO calculation. The file file.coskf is much smaller than file.t21. This file file.coskf can not be used like an ordinary .t21 file, but it contains all necessary information such that it can be used as input for a COSMO-RS calculations.

References

1. A. Klamt, *Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena.* J. Phys. Chem. **99**, 2224 (1995)

2. A. Klamt, V. Jonas, T. Bürger and J.C. Lohrenz, *Refinement and Parametrization of COSMO-RS.* J. Phys. Chem. A **102**, 5074 (1998)

3. A. Klamt, *COSMO-RS From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design, Elsevier.* Amsterdam (2005), ISBN 0-444-51994-7.

4. C.C. Pye, T. Ziegler, E. van Lenthe, J.N. Louwen, *An implementation of the conductor-like screening model of solvation within the Amsterdam density functional package. Part II. COSMO for real solvents.* Can. J. Chem. **87**, 790 (2009)

5. B. Delley, *The conductor-like screening model for polymers and surfaces.* Molecular Simulation **32**, 117 (2006)