COSMOthermX

A Graphical User Interface to the COSMOtherm Program

Tutorial for version C30_1201
December 7th 2011

COSMOlogic GmbH & Co. KG
Burscheider Str. 515, D-51381 Leverkusen, Germany
# Table of Contents

Abbreviations ............................................................................................................. 4

1 Introduction: COSMO-RS theory ........................................................................... 5

2 COSMOtherm and COSMOthermX ...................................................................... 7
   2.1 Quality levels and parameterizations ............................................................... 7
   2.2 Flowchart of a property calculation with COSMOthermX ............................ 9
   2.3 The main window ............................................................................................ 10
   2.4 Global options .................................................................................................. 13
   2.5 Mixture options .............................................................................................. 17
   2.6 Output files ...................................................................................................... 18
   2.7 Job list ............................................................................................................. 19

3 Compounds .............................................................................................................. 20
   3.1 Compound selection ........................................................................................ 20
   3.2 The compounds context menu ........................................................................ 21
   3.3 Selecting compounds from databases ............................................................ 23
   3.4 New Molecule ................................................................................................ 23
      Example: Glycofurol ....................................................................................... 26
   3.5 Conformers ..................................................................................................... 26
   3.6 Visualization of σ-surfaces, σ-profiles, and σ-potentials .................................. 27
      Example: COSMO charge surface visualization, σ-profile and σ-potential ...... 27
   3.7 Using COSMOview ......................................................................................... 27

4 Property input .......................................................................................................... 29
   4.1 Mixture: Calculation of compound properties in mixture ............................ 29
      Special Application: Cocrystal screening with COSMOthermX .................... 30
   4.2 Vapor pressure ............................................................................................... 34
      Example: Calculation of a vapor pressure curve ........................................... 34
   4.3 Boiling point .................................................................................................... 35
   4.4 Activity coefficient calculation ....................................................................... 35
      Example: Calculate the activity coefficient of aspirin in water ....................... 36
   4.5 Henry law coefficient calculation ................................................................... 36
   4.6 Gas solubility .................................................................................................. 37
      Example: Gas solubility of methane in water .................................................. 37
   4.7 Solubility ......................................................................................................... 37
      Example: Solubility of glycol in hexane ............................................................ 39
   4.8 Solvent screening ........................................................................................... 39
   4.9 Salt solubility .................................................................................................. 41
      Example: Prediction of the solubility of NaCl in several solvents ...................... 42
   4.10 Salt solubility screening .................................................................................. 42
   4.11 Partition coefficient calculation (log P / log D) .............................................. 43
      Example: Prediction of octanol / water partition coefficients ........................... 44
   4.12 Calculation of pKA ....................................................................................... 44
      Example: Calculation of the aqueous pKₐ of pyrrolidine ............................... 45
   4.13 Vapor liquid-equilibria (VLE) and liquid-liquid equilibria (LLE) ................. 45
      Example: Calculate the solubility of 2-chlorophenol in hexane ...................... 47
   4.14 Solid liquid equilibria .................................................................................... 47
      Example: Solid-liquid equilibrium curve of toluene and ethylbenzene ............ 48
   4.15 FlatSurf: surface activity ............................................................................... 49
      Example: Calculate the air-water surface partition energy ............................. 50
4.16 Density.......................................................................................................................... 51
Example: Liquid density and liquid molar volume of pure compounds .................................. 51
4.17 Viscosity.......................................................................................................................... 52
4.18 The o-moment approach and QSPR calculations ............................................................. 52
4.19 Similarity......................................................................................................................... 54
4.20 Liquid extraction ............................................................................................................. 54
Example: Three-phase equilibrium between water, hexane, and [bmim+][pf6-] with an additional solute ...................................................................................................................... 56
4.21 Reaction......................................................................................................................... 56
Example: Solvent choice for a simple ester reaction ............................................................... 57
4.22 Ionic Liquid Screening.................................................................................................... 58
4.23 Ionic Liquid Properties .................................................................................................. 58
4.24 Environmental / Safety .................................................................................................. 58
5 Using your own COSMO files............................................................................................ 59
5.1 Using the file manager ..................................................................................................... 59
5.2 Adding your own databases ............................................................................................ 59
5.3 Extend the existing databases .......................................................................................... 60
6 Atom weighting .................................................................................................................. 61
Index ..................................................................................................................................... 63
Abbreviations

2D: Two dimensional
3D: Three dimensional
AM1: A semiempirical quantum chemical method
BP-SVP / BP-TZVP: Becke-Perdew$^{1, 2, 3}$ (BP) functional for density functional theory calculations with a split valence plus polarization function (SVP) or triple valence plus polarization function (TZVP) basis set. The necessary parameterization file does always correspond to one functional and basis set. The term “BP-TZVP parameterization” is thus sometimes used and refers to the COSMOtherm parameterization not the basis set specification.
CAS-Number: The Chemical Abstracts Services registration number is a unique identifier for compounds.
COSMO: COnductor like Screening MOdel
COSMO-RS: COnductor like Screening MOdel for Real Solvents
DB: Database, usually used for the COSMOtherm compound databases.
DFT: Density Functional Theory: A quantum chemical theory used in several software packages for molecular or lattice calculations.
HB: Hydrogen Bond
IL: Ionic Liquid
LFER: Linear Free Energy Relationship
LLE: Liquid Liquid Equilibrium
MW: Molar Weight
QC / QM: Quantum Chemical / Quantum Mechanical
QSPR: Quantitative Structure Property Relationship, also QSAR (A = activity)
$\sigma$ (sigma): The COSMO screening charge
SLE: Solid Liquid Equilibrium
SMILES: Simplified Molecular Input Line Entry Specification, notation to describe molecules e.g. methane is “C” ethane is “CC” and ethanol is “CCO”
SMS: Sigma Match Similarity
VLE: Vapor Liquid Equilibrium
1 Introduction: COSMO-RS theory

COSMO-RS is a predictive method for thermodynamic equilibria of fluids and liquid mixtures that uses a statistical thermodynamics approach based on the results of quantum chemical calculations. The underlying quantum chemical model, the so called “Conductor-like Screening MOdel” (COSMO)\(^4\), is an efficient variant of dielectric continuum solvation methods. In COSMO calculations the solute molecules are calculated in a virtual conductor environment. In such an environment the solute molecule induces a polarization charge density \(\sigma\) on the interface between the molecule and the conductor, i.e. 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 (QC) self-consistency algorithm, the solute molecule is thus converged to its energetically optimal state in a conductor with respect to electron density. The molecular geometry can be optimized using the same methods for calculations in vacuum. The time consuming QC calculation has to be performed only once for each molecule of interest. The polarization charge density of the COSMO calculation (also called screening charge density), which is a good local descriptor of the molecular surface polarity, is used to extend the model towards “Real Solvents” (COSMO-RS)\(^5,6\). The (3D) polarization density distribution on the surface of each molecule \(i\) is converted into a distribution-function, the so called \(\sigma\)-profile \(p_i(\sigma)\), which gives the relative amount of surface with polarity \(\sigma\) on the surface of the molecule. The \(\sigma\)-profile for the entire solvent of interest \(S\), which might be a mixture of several compounds, \(p_S(\sigma)\) can be built by adding the \(p_i(\sigma)\) of the components weighted by their mole fraction \(x_i\) in the mixture.

\[
p_S(\sigma) = \sum_{i \in S} x_i p_i(\sigma)
\]  

The most important molecular interaction energy modes, i.e. electrostatics \(E_{\text{misfit}}\) and hydrogen bonding \(E_{\text{HB}}\) are described as functions of the polarization charges of two interacting surface segments \(\sigma\) and \(\sigma'\) or \(\sigma_{\text{acceptor}}\) and \(\sigma_{\text{donor}}\), if the segments are located on a hydrogen bond donor or acceptor atom. Electrostatic energy arises from the misfit of screening charge densities \(\sigma\) and \(\sigma'\), as illustrated above. The less specific van der Waals \(E_{\text{vdW}}\) interactions are taken into account in a slightly more approximate way.

\[
E_{\text{misfit}}(\sigma, \sigma') = a_{\text{eff}} \frac{\sigma' \alpha}{2} (\sigma + \sigma')^2
\]  

\(a_{\text{eff}}\) is a characteristic interaction parameter.
E_{HH} = a_{eff} c_{HH} \min(0; \min(0; \sigma_{\text{donor}} + \sigma_{HH}) \max(0; \sigma_{\text{acceptor}} - \sigma_{HH})) \quad (3)

E_{vdW} = a_{eff} (\tau_{vdW} + \tau'_{vdW}) \quad (4)

Eqs. 2-4 contain five adjustable parameters, an interaction parameter $\alpha'$, the effective contact area $a_{eff}$, the hydrogen bond strength $c_{HH}$, the threshold for hydrogen bonding $\sigma_{HH}$, and the element specific vdW interaction parameter $\tau_{vdW}$. In order to take the temperature dependency of $E_{HH}$ and $E_{vdW}$ into account, temperature dependent factors are applied, each with one adjustable parameter.$^7$

The transition from microscopic molecular surface charge interactions to macroscopic thermodynamic properties of mixtures is possible with a statistical thermodynamic procedure. The molecular interactions in the solvent are fully described by $p_S(\sigma)$, and the chemical potential of the surface segments can be calculated solving a coupled set of non-linear equations.

$$
\mu_S(\sigma) = -\frac{RT}{a_{eff}} \ln \left[ \int p_S(\sigma') \exp \left( \frac{a_{eff}}{RT} \left( \mu_S(\sigma') - E_{\text{misfit}}(\sigma, \sigma') - E_{HH}(\sigma, \sigma') \right) \right) d\sigma' \right] \quad (5)
$$

The $\sigma$-potential $\mu_S(\sigma)$ is a measure for the affinity of the system $S$ to a surface of polarity $\sigma$. The vdW energy, which does not appear in Eq. (5), is added to the reference energy in solution (energy of the COSMO calculation). The chemical potential of compound $i$ in the system $S$ (the solvent) can now be calculated by integration of $\mu_S(\sigma)$ over the surface of the compound.

$$
\mu_i' = \mu'_{C,S} + \int p'(\sigma) \mu_S(\sigma) d\sigma \quad (6)
$$

To take into account size and shape differences of the molecules in the system an additional combinatorial term, $\mu'_{C,S}$, which depends on the area and volume of all compounds in the mixture and three adjustable parameters is added. For information on the exact expression for the combinatorial term refer to the COSMOtherm User’s Manual. The chemical potential can be used to calculate a wide variety of thermodynamic properties, e.g. the activity coefficient:

$$
\gamma_i' = \exp \left( \frac{\mu_i' - \mu_i}{RT} \right) \quad (7)
$$

where $\mu_i'$ is the chemical potential in the solvent $S$, and $\mu_i$ is the chemical potential of the pure compound $i$.

For a more detailed introduction to COSMO and COSMO-RS, refer to the COSMOtherm User’s Manual and the cited literature.
2 COSMOtherm and COSMOthermX

COSMOtherm is a command line/file driven program which can be run directly from a UNIX or DOS shell. It allows for the calculation of any solvent or solvent mixture and solute or solute system at variable temperature and pressure. COSMOtherm uses the chemical potentials derived from COSMO-RS theory to compute all kinds of equilibrium thermodynamic properties or derived quantities, e.g.:

- Vapor pressure
- Free energy of solvation
- Activity coefficients
- Partition coefficients
- Solubility and solid-liquid equilibria (SLE).
- Liquid-liquid equilibrium (LLE) and vapor-liquid equilibrium (VLE)
- Phase diagrams, azeotropes, miscibility gaps, excess enthalpies and excess free energies
- $pK_a$ of acids and bases
- Various QSAR models
- Reaction constants
- Liquid extraction equilibria

COSMOthermX is a Graphical User Interface to the COSMOtherm command line program. It allows for the interactive use of the COSMOtherm program, i.e. selection of compounds, preparation of property input, program runs and display of calculation results.

2.1 Quality levels and parameterizations

The input for the compounds is read from the COSMO files, identified by the extensions .cosmo or .ccf, which are result files from quantum chemical COSMO calculations. COSMOtherm extracts the relevant information directly from the COSMO files. The compressed COSMO files (.ccf) use significantly less disk space than conventional COSMO files. At least one COSMO file or compressed COSMO file has to be selected as compound input. Note that COSMO file names with blanks are not allowed by COSMOthermX.

Because the quality, accuracy, and systematic errors of the electrostatics resulting from the underlying COSMO calculations depend on the quantum chemical method as well as on the basis set, COSMOtherm needs a special parameterization for each method / basis set combination. All of these parameterizations are based on molecular structures quantum chemically optimized at the given method / basis set level.

COSMO files shipped with COSMOtherm are available on various quantum chemical levels. Recommendations for which method to use depend upon the required quality and the later usage of the predictions.

The application of COSMOtherm in chemical and engineering thermodynamics (e.g. prediction of binary VLE or LLE data, activity coefficients in solution or vapor pressures) typically requires high quality of property predictions of mixtures of small to medium sized molecules (up to 25 non-Hydrogen atoms). The recommended quantum chemical method for such a problem is a full TURBOMOLE BP-RI-DFT COSMO optimization of the molecular structure using the large TZVP basis set$^8$, in the following denoted BP-TZVP, and the corresponding parameter file BP_TZVP_C30_1201.ctd. A similar quality can be reached with the DMOL3 program package using full COSMO and gas phase geometry optimization with the PBE DFT functional and the numerical DNP basis set and the corresponding parameter file DMOL3_PBE_C30_1201.ctd in
COSMOtherm.

With the current version, a new quantum chemical calculation level, BP-TZVPD-FINE, is introduced to COSMOtherm. This method is based on a Turbomole BP-RI-DFT COSMO single point calculation with TZVPD basis set on top of an optimized BP/TZVP/COSMO geometry. The COSMO single point calculation is using the TZVPD basis set with diffuse basis functions and a novel type of molecular surface cavity construction (fine grid marching tetrahedron cavity, FINE\(^9\)), which creates a COSMO surface whose segments are more uniform and evenly distributed compared to the standard COSMO cavity. The gas phase energy files of this level are optimized on Turbomole BP-RI-DFT level with TZVP basis set, followed by a single point BP-RI-DFT calculation with the TZVPD basis set. The associated *BP_TZVPD_FINE_HB2012_C30_1201.ctd* parameter set additionally incorporates a novel hydrogen bonding term (HB2012), which shows improved thermodynamic property prediction results for compound classes where the classical COSMO-RS hydrogen bonding term showed weaknesses, e.g. with secondary and tertiary aliphatic amines and polyether compounds.

The wider realm of organic liquid and gas phase thermodynamics is predicted with roughly the same quality as the standard BP-TZVP-COSMO and DMOL3_PBE methods. Until further experience is gained, this level of theory and parameterization are provided in an "as is" state. COSMOtherm users should feel encouraged to try this level: it probably fits their special demands in property prediction better than the recommended standard levels BP-TZVP-COSMO and DMOL3-PBE, in particular if the mentioned compound classes are involved.

Screening a large number of compounds, e.g. prediction of solubility of compounds in various solvents, typically requires a predictive quality that is somewhat lower than for chemical engineering applications. The molecules involved are often larger (>100 atoms) and an overall large number of compounds has to be computed by quantum chemistry. Thus a compromise between computational demands and quality of the predictions has to be made: A very good compromise is the optimization of molecular geometry on the computationally very cheap semiempirical MOPAC AM1-COSMO level\(^10\) with a subsequent single point COSMO calculation on Turbomole BP-RI-DFT COSMO level using the small SVP basis set. This method is named BP-SVP-AM1 in the following, and the corresponding parameterization is available in the *BP_SVP_AM1_C30_1201.ctd* parameter file.

For more information on available quantum chemical levels and parameterizations refer to the COSMOtherm User’s Manual, section 3.
2.2 Flowchart of a property calculation with COSMOthermX

Start

Select compounds from the Databases or the File Manager

Found all compounds?

Yes

Select settings for property calculations

Add

Another property?

Yes

No

Run

Add compounds to database

Use the Compound Wizard to generate COSMO files and/or gas phase structures.

COSMOthermX

external

TURBOMOLE
2.3 The main window

The main window is separated into 4 parts: the pull down menu and shortcuts section in the upper region, the compound list with database access and job list on the left side, the input preparation and the property section on the right and right bottom side.

The COSMOthermX main window has several menus:

File:
NEW JOB: Create a new input file. Type the filename and press “Open”. Also available as shortcut □.

OPEN JOB: Open an existing input file. Select a file from the directory or type the filename into the “File name” text field and press “Open”. The panel also allows for changing the directory, also available as shortcut □.

SAVE: Save the input file to the current directory with the actual name. Also available as shortcut □.

SAVE AS…: Choose a directory and a name for the input file to be saved.

CLOSE JOB: The job selected from the job list is removed from the list. All input and output files of the job are kept.

DELETE JOB: The job selected from the job list will be deleted, together with all input and output files of the job.

QUIT: Exit the program.

Edit:
EDIT INPUT FILE: Open the actual input file in a text editor window. The input file can be changed manually, e.g. for additional input in the compound lines (further information: COSMOtherm User’s Manual). The input file can be saved or COSMOtherm can be run. Note that not all changes can be read in when the input file is re-opened with COSMOthermX. Also available as shortcut □.
SAVE COMPLETE COMPOUND LIST: With this option it is possible to save the complete compound list to a file for later use. The concentrations are also saved. To save only the selected compounds to a list use the SAVE TO COMPOUND LIST option from the compound list context menu has to be used.

OPEN COMPOUND LIST: A list of previously saved compounds can be loaded with this option. The saved concentrations of the compounds will also be loaded but not automatically applied. Please select RESTORE LIST CONCENTRATIONS from the extras menu for this feature.

RESTORE LIST CONCENTRATIONS: Apply the saved list concentrations after loading a compound list. The saved concentrations of the compounds will not be applied automatically when a list is opened.

Compounds:
SVP: open database index files of the BP-SVP-AM1 level of theory in tabulated form.
TZVP: open database index files of the BP-TZVP level of theory in tabulated form.
TZVPD-FINE: open database index files of the BP-TZVPD-FINE level of theory in tabulated form.
DMOL3: open database index files of the DMOL3-PBE level of theory in tabulated form.
SEARCH IN ALL DB: Search a compound in all databases that are listed in the EXTRAS/GENERAL SETTINGS DATABASES pane, irrespective of the parameterization which the DBs are linked to.
NEW MOLECULE: This option allows for searching for a compound by name, CAS-Number, SMILES or structure. It also offers the possibility to generate a complete new structure with TURBOMOLE (if installed).

FILE MANAGER: Opens a compound from an arbitrary directory.

Run:
RUN: The actual input file is saved automatically with the current file name and COSMOtherm is run. Also available as shortcut 🍺.

RUN (EXTERNAL) INPUT: A COSMOtherm calculation can be run with an external input file, without opening it in the GUI. Any valid COSMOtherm input file can be run.

Result:
OPEN OUTPUT FILE: Select a COSMOtherm output file from a directory or type the filename into the “File name” text field and press “Open”.

OPEN TABLE FILE: Select a COSMOtherm table file from a directory or type the filename into the “File name” text field and press “Open”.

OPEN COSMOOMIC FILE: : Select a COSMOmic file from a directory or type the filename into the “File name” text field and press “Open”.

OPEN ACTUAL OUTPUT FILE: Select a COSMOtherm output file from a directory or type the filename into the “File name” text field and press “Open”.

SAVE COMPLETE COMPOUND LIST: With this option it is possible to save the complete compound list to a file for later use. The concentrations are also saved. To save only the selected compounds to a list use the SAVE TO COMPOUND LIST option from the compound list context menu has to be used.

OPEN COMPOUND LIST: A list of previously saved compounds can be loaded with this option. The saved concentrations of the compounds will also be loaded but not automatically applied. Please select RESTORE LIST CONCENTRATIONS from the extras menu for this feature.

RESTORE LIST CONCENTRATIONS: Apply the saved list concentrations after loading a compound list. The saved concentrations of the compounds will not be applied automatically when a list is opened.

Compounds:
SVP: open database index files of the BP-SVP-AM1 level of theory in tabulated form.
TZVP: open database index files of the BP-TZVP level of theory in tabulated form.
TZVPD-FINE: open database index files of the BP-TZVPD-FINE level of theory in tabulated form.
DMOL3: open database index files of the DMOL3-PBE level of theory in tabulated form.
SEARCH IN ALL DB: Search a compound in all databases that are listed in the EXTRAS/GENERAL SETTINGS DATABASES pane, irrespective of the parameterization which the DBs are linked to.
NEW MOLECULE: This option allows for searching for a compound by name, CAS-Number, SMILES or structure. It also offers the possibility to generate a complete new structure with TURBOMOLE (if installed).

FILE MANAGER: Opens a compound from an arbitrary directory.

Run:
RUN: The actual input file is saved automatically with the current file name and COSMOtherm is run. Also available as shortcut 🍺.

RUN (EXTERNAL) INPUT: A COSMOtherm calculation can be run with an external input file, without opening it in the GUI. Any valid COSMOtherm input file can be run.

Result:
OPEN OUTPUT FILE: Select a COSMOtherm output file from a directory or type the filename into the “File name” text field and press “Open”.

OPEN TABLE FILE: Select a COSMOtherm table file from a directory or type the filename into the “File name” text field and press “Open”.

OPEN COSMOOMIC FILE: : Select a COSMOmic file from a directory or type the filename into the “File name” text field and press “Open”.

OPEN ACTUAL OUTPUT FILE: Select a COSMOtherm output file from a directory or type the filename into the “File name” text field and press “Open”.
OPEN ACTUAL TABLE FILE: Select a COSMOtherm table file from a directory or type the filename into the “File name” text field and press “Open”.

OPEN TABLE FILE WITH EXCEL: Select a COSMOtherm table file from a directory or type the filename into the “File name” text field and press “Open”. If possible, the table file will be opened in Microsoft Excel.

EXPORT AS EXCEL FILE: The table file can be saved in MS Excel format.

Extras:
GLOBAL OPTIONS: A panel with selections for gas phase energy input, units, COSMOtherm output print options, additional output files and program control. For a detailed description of the global options dialog see below.

MIXTURE OPTIONS: A panel with options applying to settings for the mixture calculation. A shortcut is also available in the main window. Mixture options are used to apply certain options to a mixture only. With this options it is possible to use different options in a single COSMOtherm run. Mixture options will only be used if the “Use Mixture Options” checkbox is activated in the property panel. If several mixtures or properties are calculated in a single run, the mixture options have to be activated each time the property settings are transferred to the property selection window, otherwise they will not be used for the respective property calculation. For a detailed description of the mixture options dialog see below.

VIEWER SETTINGS: Change atom / bond color, show or hide label, change sphere / cylinder smoothness and so on.

GENERAL SETTINGS: Opens the GENERAL SETTINGS dialog where some settings are already specified: Paths for the COSMOtherm executable and the CTDATA directory of the COSMOtherm installation are set, and the parameter files for the quantum chemical levels (extension .ctd) are specified. Additionally, you can set paths for the Adobe Acrobat Reader and a web browser. If you intend to use COSMO metafiles (extension .mcos) for the fragment approach, you should also specify the fragment directory. When the path for the QSPR property directory is set QSPR coefficient files (extension .prop) can be read from there. A standard USER DIRECTORY is specified by default, which will be used automatically. If you intend to save input files the user directory will be used as starting point for the file select dialog. After saving a file, the last used directory will be used as long as COSMOthermX is not closed. USE NETBOOK OPTION (SCROLLBARS) will enable global scrollbar for the main window. This is helpful for small displays. When SHOW MEMORY USAGE is enabled, the amount of memory used will be displayed in the menu bar in COSMOthermX. DO NOT LOAD PREVIOUSLY SAVED JOBS ON START UP will start COSMOthermX with an empty job list. With OPEN OUTPUT FILE FOR EVERY RESULT, the .out file of COSMOtherm will be opened in a text editor window for every job. USE NEW PLOT LIBRARIES will result in the plots in the table view be drawn in a different style than before. Since the new plot libraries have not been tested extensively, this is not set by default.
**DATABASE SETTINGS:** Opens the Settings dialog with the Databases tab. The databases that come with the COSMOTHERM release are specified here. Additional databases can be added with ADD DATABASES. For detailed information on adding your own databases, refer to the section “Using your own COSMO files”.

**Tools:**
- **COSMOVIEW:** Opens the COSMOview tool which allows for the visualization of .wrl files generated by COSMOtherm. For more information, please refer to the section “Using COSMOview”.
- **COSMOWEIGHT:** Opens the COSMOweight tool. For information on atom weighting and the COSMOWeight tool, please refer to the section “Atom Weighting”.

**Help:**
- **PHYSICAL CONSTANTS:** Displays information about some physical constants and conversion factors.
- **COSMOTHERM MANUAL:** Open the COSMOTHERM User’s Manual with the Adobe Acrobat Reader.
- **COSMOTHERMX TUTORIAL:** Open the COSMOTHERMx Tutorial (this document) with the Adobe Acrobat Reader.
- **COSMOMIC MANUAL:** Open the Manual for the COSMOMIC extension.
- **ABOUT:** Information about the current COSMOTHERMx version is displayed.
- **LICENSE AGREEMENTS USED:** A list of external libraries and software tools employed and the corresponding license terms.

### 2.4 Global options

Global options can be set from the “GLOBAL OPTIONS” panel. The panel offers selections for gas phase energy input, units, COSMOTHERM output print options, additional output files, program control and QSPR options. The panel can be opened from “Extras/Global Options” or directly from the main window.

Global options apply for the whole COSMOTHERM run, i.e. all temperature / mixture / property lines in the input.

**Gas phase estimation:**
- **USE ENERGY FILE:** The gas phase energy is required for certain properties (e.g. the calculation of the chemical potential in the gas phase). It can be taken from a gas phase quantum chemical calculation (.energy file) or empirically estimated by COSMOTHERM. If possible, the quantum chemical calculated value of the gas phase energy should be used. Energy files from COSMObase use Hartree units. If this option is checked COSMOTHERM will use the energy file if no adequate pure compound properties are available.

- **USE COMPOUND PROPERTY FILE:** For some properties experimental vapor pressure data can be used. The relevant pure compound property data are taken from the .vap file. The Antoine or Wagner coefficients might be used in this case.
Units:
The units for the gas phase energy input and the general COSMOtherm output of the calculated properties can be chosen here.

Print options for the COSMOtherm output and table files:
PRINT COMPOUND/MIXTURE NUMBER: Write compound / mixture number to the right side of the output file. This may be useful for efficient processing of the output file.

PRINT CONFORMER INFO: If a compound input consists of several conformers this option causes the output of the calculated COSMOtherm mixture information to be written for each individual conformer. By default, only the results for the mixed compound are written to the output file.

SUPPRESS PURE COMPOUNDS INFO: Pure compound information will not be written to the output file.

SUPPRESS MIXTURE OUTPUT IN .OUT FILE: Mixture information will not be written to the output file.

PRINT 15 DIGIT LONG NUMBERS TO .OUT-FILE: Print all real numbers in scientific exponent number format with 15 significant digits to the output file.

PRINT FULL LENGTH ATOMIC WEIGHT STRING: Print complete atomic weight or real weight string to the compound section of the output file. If you toggle this option, the line for the atomic weights may become very long.

PRINT MOLECULAR SURFACE CONTACTS: Print statistics of molecular surface contacts for all compounds in all mixtures to the output file. For a detailed description see section 5.7 of the COSMOtherm User’s Manual.

PRINT DETAILED SEGMENT MOLECULE CONTACTS: Print statistics of the molecular surface contacts for all segments of all compounds in all mixtures to the output file and to the contact statistics table file name.contact. Refer to the COSMOtherm User’s Manual, section 5.7, for details.

PRINT DERIVATIVES OF CHEMICAL POTENTIAL: Print the values of the temperature and composition derivatives of the chemical potentials of all compounds in all mixtures to the output file. See section 5.6 “Chemical Potential Gradients” of the COSMOtherm User’s Manual for further information.

PRINT FULL COMPOUND NAMES IN TABLE ENTRIES: Print full compound and/or molecule names to all tables in the table file and the sigma-moments file. By default the compound/molecule names are cropped after 20 characters.

Additional output files:
σ-MOMENTS (.MOM): Write the σ-moments of all processed compounds in tabulated form to filename.mom. In addition some other molecular information will be written to filename.mom, including volume V, molecular weight, dielectric energy $E_{\text{die}}$, average energy correction $dE$, van der Waals energy in continuum $E_{\text{vdW}}$, ring correction energy $E_{\text{ring}}$ and the standard chemical potential of the molecule in the gas phase with respect to the ideally screened state $\mu_{\text{QSPR gas}} = E_{\text{COSMO}} - E_{\text{gas}} + dE + E_{\text{vdW}} + E_{\text{ring}} - H_{\text{gas}}RT$, using $T = 25^\circ C$. Please note that $\mu_{\text{QSPR gas}}$ is only used as special QSPR descriptor and is only loosely connected to the gas phase chemical potential used otherwise). Refer also to sections 5.4 and 5.5 of the COSMOtherm User’s Manual.
**ATOMIC σ-MOMENTS (.MOMA):** Write the atomic σ-moments of all processed compounds to filename.moma. If this option is used, σ-moments will be calculated for each atom of the compounds.

**σ-PROFILES (.PRF):** Write the σ-profiles of all processed compounds to file filename.prf. A summary of the σ-profiles will be written in tabulated form to the table file filename.tab.

**σ-POTENTIALS (.POT):** Write the σ-potentials of all calculated mixtures to filename.pot. A summary of the σ-potential information will be written in tabulated form to the table file filename.tab.

**GENERATE VRML (GEOMETRY):** Create VRML files name_mol.wrl of the molecular geometry of all compounds. The files will be written to the directory of the selected input file in the next COSMOtherm run.

**GENERATE VRML (σ-SURFACE):** Create VRML files name_sig.wrl of the molecular COSMO surface charges of all compounds. The files will be written to the directory of the selected input file in the next COSMOtherm run.

VRML, the "Virtual Reality Modeling Language", is a script language allowing for the interactive examination of virtual three-dimensional objects. VRML files are usually identified by the extension .wrl. All VRML files generated by COSMOtherm can be viewed by the internal VRML viewer of COSMOtherm. Otherwise, VRML files can be viewed with common World-Wide-Web browsers such as Mozilla Firefox or Microsoft Internet Explorer if an appropriate VRML browser plug-in has been installed. Plug-ins are freely available, e.g. the Cortona VRML client (http://www.parallelgraphics.com/products/cortona) COSMOtherm offering advanced functionality like colour correction and charge picking. For more information refer to the section 'Using COSMOview'.

**QSPR Property Computation:**

Only one of the following options can be chosen at a time. These options allow for the calculation of special properties in addition to normal COSMOtherm calculations. If different QSPR properties should be calculated in a single run, the Mix QSPR panel has to be used. Note that fitting for QSPR is done on the BP-SVP-AM1 parameterization level. Only compounds from the SVP Database should be used.

**NO SELECTION:** This option is preselected, no QSPR (Quantitative Structure Property Relationship) will be calculated.

**BLOOD BRAIN PARTITIONING:** The penetration of the blood brain barrier will be calculated.

**INTESTINAL ABSORPTION COEFFICIENT:** The intestinal absorption will be calculated.

**PLASMA PROTEIN (HUMAN SERUM ALBUMIN) PARTITIONING:** The binding to human serum albumin will be calculated.

**SOIL SORPTION COEFFICIENT:** The soil water partition will be calculated.

**OCTANOL WATER PARTITIONING:** The octanol water partition will be calculated.

By default, the computed property value will be listed in the compound section of the COSMOtherm output file. An additional file with the extension .mom will be written, listing the
molecular \( \sigma \)-moments and, in the last column, the computed property. Note that QSPR property calculations can also be done from the Mix-QSPR card, which allows for a larger variety of settings.

**Program control settings:**

**SWITCH OFF TEMP. DEPENDENCY OF HYDROGEN BOND CONTRIB.:** Switch off temperature dependency of the hydrogen bond contribution to the total interaction energy of the compound for the complete COSMO\_therm run.

**SWITCH OFF TEMP. DEPENDENCY OF VAN DER WAALS CONTRIB.:** Switch off temperature dependency of the van der Waals contribution to the total interaction energy of the compound, active for the complete COSMO\_therm run.

**SWITCH OF HYDROGEN BONDING:** Switch off hydrogen bonding (HB) contribution to the chemical potential.

**SWITCH OFF VAN DER WAALS CONTRIBUTIONS:** Switch off van der Waals (vdW) interaction energy contribution to the chemical potential.

**SWITCH OFF COMBINATORIAL CONTRIB. TO CHEMICAL POTENTIAL:** Switch off combinatorial contribution to the chemical potential for the complete COSMO\_therm run.

**USE SUM OF FRAGMENT ENERGIES AS METAFILE ENERGY:** (Expert use only) Use the sum of the fragment COSMO file energies in the calculation of the total free energy (by default a zero value is used instead). Details on the usage of atomic weights and COSMO-metafiles are explained in section 5.3 of the COSMO\_therm Users manual.

**USE COMPOUND BOILING POINTS IN VAPOR PRESSURE PREDICTION:** The pure compound boiling points as given in the compound input section can be used as a reference points for scaling vapor pressure predictions.

**USE “EXTERNAL” QUANTUM CHEMICAL ENERGIES (COSMO):** The external quantum chemical energies and zero point energies from the energy files .energy (or property files .vap) are not used until this option has been checked. Conformer equilibria are computed using \( E_k^{\text{conf}} = E_k^{\text{QM}} + E_k^{\text{ZPE}} + \mu_k + dE_k \).

**USE “EXTERNAL” QUANTUM CHEMICAL ENERGIES (GAS):** similar to USE “EXTERNAL” QUANTUM CHEMICAL ENERGIES (COSMO), but conformer equilibria are computed from \( E_k^{\text{conf}} = E_k^{\text{QM}} + E_k^{\text{ZPE}} + \Delta G_k^{\text{Solv}} \).

**ALLOW NEGATIVE FREE ENERGY OF FUSION:** Allow for negative values of \( \Delta G_{\text{fus}} \). For more information, refer to sections 2.1 and 2.3.4 of the COSMO\_therm Users manual.

**USE HEAT CAPACITY OF FUSION ESTIMATE:** The heat capacity of fusion is approximated as \( \Delta C_{P_{\text{fus}}} = \Delta S_{\text{fus}} = \Delta H_{\text{fus}}/T_{\text{melt}} \). The value of \( \Delta G_{\text{fus}}(T) \) thus obtained can be used to compute the solubility of solid compounds or salts as well as solid-liquid-equilibria.

**SWITCH OFF USAGE OF GAS PHASE CONFORMER MINIMUM ENERGY:** Switch off the usage of gas phase minimum conformer energies in single-conformer calculations.

**USE OLD (SLOW) CONFORMER EQUILIBRIUM:** Use the old conformer equilibration algorithm. See COSMO\_therm User’s Manual, section 2.1 “Global Command and Comment Lines” for further information.
CHANGE THRESHOLD FOR THE ITERATIVE SELF-CONSISTENCY: Change threshold for the iterative self-consistency cycle for the determination of the chemical potential. A smaller value leads to higher accuracy of the COSMOtherm results but also to a longer computational time due to an increasing number of iterations. Default value: $10^{-8}$.

Compound input settings:
USE ALL CONFORMER FILES FOUND FOR A COMPOUND: Use all .cosmo, .ccf or .mcos files as conformers that are found in the database or directory from where the compound is selected. The COSMO file filenames must follow the name convention of conformer COSMO files in COSMObase (i.e. conformer COSMO files are named by subsequent numbers starting with zero: name0.cosmo, name1.cosmo, …, name9.cosmo). When this option is checked, you will be enabled to make a selection from compounds 0-9.

Fine tuning of COSMO-RS parameters
Several options for adjusting the parameters of the COSMO-RS method to specific requirements. For detailed information, refer to the COSMOtherm User's Manual, section 5.11 “Fine Tuning of COSMO-RS parameters”.

2.5 Mixture options
Settings from the MIXTURE OPTIONS panel can be applied to individual temperature / mixture / property lines. The panel can be opened from “Extras/Mixture Options” or directly from the main window. Mixture options will only be used if the USE MIXTURE OPTIONS checkbox is activated in the property panel. If several mixtures or properties are calculated in a single run, the mixture options have to be activated each time the property settings are transferred the property selection window, or they will not be used for the respective property calculation.

Print options for the COSMOtherm output file:
SUPPRESS MIXTURE OUTPUT IN .OUT-FILE: Mixture information will not be written to the output file.

PRINT DERIVATIVES OF CHEMICAL POTENTIAL: Print the values of the temperature and composition derivatives of the chemical potentials of all compounds in all mixtures to the output file. See COSMOtherm User’s Manual, section 5.6 “Chemical Potential Gradients” for further information.

PRINT FULL COMPOUND NAMES IN TABLE ENTRIES: Print full compound and/or molecule names to all tables in the table file and the sigma-moments file. By default the compound/molecule names are cropped after 20 characters.

Program control settings:
SWITCH OFF TEMP. DEPENDENCY OF HYDROGEN BOND CONTRIB.: Switch off temperature dependency of the hydrogen bond contribution to the total interaction energy of the compound for the complete COSMOtherm run.

SWITCH OFF TEMP. DEPENDENCY OF VAN DER WAALS CONTRIB.: Switch off temperature dependency of the van der Waals contribution to the total interaction energy of the compound, active for the complete COSMOtherm run.

SWITCH OF HYDROGEN BONDING: Switch off hydrogen bonding contribution to the chemical potential.
**Switch off van der Waals contributions**: Switch off van der Waals (vdW) interaction energy contribution to the chemical potential.

**Switch off combinatorial contrib. to chemical potential**: Switch off combinatorial contribution to the chemical potential for the complete COSMOtherm run.

**Do not check for charge neutrality**: Overrides the check for charge neutrality of a given mixture composition and allows you to compute non-neutral mixtures.

**Use old (slow) conformer equilibrium**: Use the old conformer equilibration algorithm. See COSMOtherm User’s Manual, section 2.1 “Global Command and Comment Lines” for further information.

**Select compounds printed in .out file**:
Write to the COSMOtherm output file the evaluated information only for the selected compounds. Helps to shorten the output file if not all evaluated information is required by the user.

**Advanced settings**: Switch of combinatorial contribution for specific compounds:
The combinatorial contribution is switched off for the selected compounds only.

### 2.6 Output files

Files produced by COSMOtherm and COSMOthermX are

- The regular input file `.inp`
- The regular output file with the extension `.out`
- The summary table file `.tab`
- If the corresponding options are set:
  - σ-profiles `.prf`
  - σ-potentials `.pot`
  - σ-moments `.mom`
  - Atomic σ-moments `.moma`
- Files and folders for special cases such as solvent screening

The summary table file will be displayed in a table editor. If the print options for σ-profiles or σ-potentials have been set, the data for the σ-profiles or σ-potentials will also be displayed in the table editor in individual tabs. Moreover, there will be tabs where σ-profiles or σ-potentials can be plotted. Note that the regular output file will only be opened in a text editor window if the corresponding option in the **General Settings** dialog is checked.
2.7 Job list

The Job list is hold in the second tab on the side of the main window. Jobs opened via the OPEN button in the main window will be displayed in the job list. A right mouse button click on one or more jobs inside the compound list opens a context menu with several options for the highlighted job. All jobs in the job list (saved jobs from previous sessions and temporary jobs from the current session) can be closed or deleted individually. CLOSE ALL JOBS will remove all jobs from the job list but retain the files belonging to the jobs. With SHOW RESULTS, the summary table file of the selected job will be opened. MANAGE JOBS will open the context menu for the the highlighted job.
3 Compounds

Apart from the menu and shortcut bars, the COSMOthermX main window has two sections. The section on the left may show one of two different tabs. The **COMPONDS** tab contains a list of selected compounds. At the bottom of this section, there are buttons to open the **FILE MANAGER**, database files or the **NEW MOLECULE** option from which the compounds can be selected. The **JOBS** tab shows a list of recent calculations.

3.1 Compound selection

The compound section of the COSMOthermX main window offers various functionalities for compound selection. There are four buttons which open the database index files of the corresponding level of theory in tabulated form. Inside the **FILE MANAGER** or the databases, a list of compounds can be highlighted by using the “Ctrl” or “Shift” keys together with the mouse.

- **MANAGE COMPOUNDS**: Opens the context menu for the highlighted compound.
- **ACTIVATE CONFORMERS TREATMENT**: If this checkbox is marked and you have selected more than one conformer for a compound, the conformers will be weighted internally by COSMOtherm using their COSMO energy and their chemical potential.
- **SVP**: COSMO files are calculated on the BP/SVP quantum chemical level, based on AM1 molecular structures.
- **TZVP**: Same as SVP-Database, but COSMO files are calculated on the BP/TZVP quantum chemical level.
- **TZVPD-FINE**: COSMO files are calculated as single points using the BP-DFT functional and the TZVPD basis set with the novel fine grid marching tetrahedron cavity (FINE), based on BP/TZVP structures. The parameter set used for this level includes a novel Hydrogen Bond interaction term (HB2012 term). This parameter set is delivered in an “as is” state for testing purposes.
- **DMOL3**: COSMO files are calculated on the PBE/DNP quantum chemical level with the DMOL3 program.
- **FILE MANAGER**: Opens the directory tree of your system and enables to choose COSMO files of any quantum chemical level directly from the file system. Make sure that you use an adequate parameterization for the COSMO files. If you do not plan to use any compounds other than those provided with your COSMOtherm installation, it is more convenient and fail-safe to use the database buttons. Also note that COSMO files with names containing blanks cannot be used in COSMOthermX.
- **OPEN LIST**: A list of previously saved compounds can be loaded with this option. The saved concentrations of the compounds will also be loaded but not automatically applied. Please select **APPLY LIST CONCENTRATIONS** from the extras menu for this feature.
- **NEW MOLECULE**: (formerly Compound Wizard) With this tool it is possible to search compounds in all databases by name, SMILES, CAS-Number, 2D structure and 3D structure. It is also possible to generate a 3D structure from smiles and to start COSMO calculations with TURBOMOLE directly from the New Molecule tool. See detailed description below.
- **CLEAR**: Clear all compounds from the selection window. Note that individual compounds can be removed using the Delete key.
3.2 The compounds context menu

A right mouse button click on one or more compounds inside the compound list opens a context menu with several options for the highlighted compound. The context menu is also available in the database search and the NEW MOLECULE tool and may vary by applicability. Options are

- compound properties
- energy file properties
- conformer properties
- decompose conformer sets / combine to conformer sets
- open
- remove from list
- view gas phase geometry
- view cosmo geometry
- view sigma surface
- view sigma profiles / potentials
- convert selection
- write to list
- show existing files

**Compound Properties:** Pure compound property data can be edited with this option. Data entries in the dialog come from the .vap file of the compound. Properties highlighted in green indicate that data entries are available, while for properties highlighted in blue no data entries exist so far. Data can be changed or added and can subsequently be used in the COSMOTHERM input for the current calculation only or saved permanently to the .vap file. Note that if applied to database compounds SAVE TO VAP will change the corresponding .vap files in the database permanently. The use of the saved pure compound property can be switched on and off in the GLOBAL OPTIONS dialog. By default the data are used where possible.

**Energy File Properties:** COSMOTHERM can read external values for the gas phase energy and the zero point vibrational energy of a molecule from the energy file. In combination with the corresponding global options, these energy values can be used for Boltzmann weighting of conformers or in the calculation of reaction equilibria. If the compound has several conformers, there will be a tab in the energy file panel for each one.
CONFORMER PROPERTIES: A conformer weight factor can be specified. The conformer weight factor is a degeneracy or symmetry factor that has to be used if it is possible to form this conformer in several different ways. Note that this option is only available in the compound list for compounds with conformers and only if the conformers treatment is activated.

COMBINE TO CONFORMER SET / DECOMPOSE CONFORMER SET: If several compounds are highlighted and this option is selected from the context menu, the highlighted compounds will be treated as conformers of a single compound. A name for the compound can be specified. If a single compound is selected the option will appear as DECOMPOSE CONFORMER SET and construct individual compounds of the different conformers of the selected compound. Note that this option is only available in the compound list and only if the conformers treatment is activated.

OPEN COSMO FILE: Opens the .cosmo or .ccf file of the compound in a text editor.

OPEN ENERGY FILE: Opens the .energy file of the compound in a text editor.

REMOVE FROM LIST: Deletes a compound from the compound list, same as “Del” key.

VIEW GASPHASE GEOMETRY: 3D ball-and-stick model of the molecular gas phase geometry.

VIEW COSMO GEOMETRY: 3D ball-and-stick model of the molecular geometry from the .cosmo file.

VIEW SIGMA-SURFACE: 3D preview of the molecular σ-surface. This graphic has a lower resolution than the graphic you get from a VRML of the σ-surface in a VRML viewer.

VIEW SIGMA-PROFILES / POTENTIALS: The σ-profiles and the σ-potentials of the selected compounds are plotted.

CONVERT SELECTION: The selected files can be converted into a variety of other file types like .xyz, .pdb or .ml2.

SAVE SELECTION TO COMPOUND LIST: The selected files can be written to a list which can be used for further processing. If the complete compound list should be saved, the SAVE COMPOUND LIST option from the FILE menu can be used alternatively. The compound concentrations are also saved. Please note: The saved concentrations of the compounds will not be loaded automatically when a list is opened. Please select RESTORE LIST CONCENTRATIONS from the EXTRAS menu to apply the saved list concentrations after loading a compound list.

EDIT WEIGHT STRING: Opens the .cosmo or .ccf file in the COSMOweight tool and allows for the setting of a weight string and correction charges which will be written to the compound list of the input file. The weight string will not be saved to the .cosmo or .ccf file.

EDIT .MCOS-FILE: Opens the .cosmo or .ccf file in the COSMOweight tool and allows for the creation of a .mcos file.

Note that the options EDIT WEIGHT STRING and EDIT .MCOS-FILE are available only in the compound list and only if the conformer treatment is deactivated. Refer to the section “Atom weighting” for information on the use of the COSMOweight tool.

SHOW EXISTING FILES: Opens a list of available files.
The options VIEW COSMO GEOMETRY, VIEW GASPHASE GEOMETRY, VIEW SIGMA-SURFACE and VIEW SIGMA-PROFILES / POTENTIALS from the context menu require a COSMOtherm run in the background. Output files of the runs are written to temporary files which will be removed when the display windows are closed. For the 3D ball-and-stick model of the molecular geometry or the σ-surface of the molecule, σ-profiles or σ-potentials to be written to permanent files check the corresponding checkboxes in the GLOBAL OPTIONS dialog.

3.3 Selecting compounds from databases

After opening one of the database parameterizations (e.g. TZVP) a window for the compound search and selection appears. Inside the window all implemented databases of the chosen parameterization will show as different tabs. If no additional databases are implemented, only one tab (e.g. Database-TZVP) will be visible. The database tables can be sorted with respect to number, name (which is the name of the .cosmo or .ccf file), CAS-Number, molecular weight (MW), charge and formula. For some compounds, there are several conformers with different σ-profiles to be considered. By default, all available conformers are selected. You can uncheck the selection to use only the lowest energy conformer. In case you should need a specific conformer other than the lowest energy conformer, you can use the “Del” key to delete the unwanted conformers from the selection or select it from the FILE MANAGER.

The database tables can also be searched for compounds. It is possible to enter a search string (name, formula, CAS-Number, smiles code) or open a text file with a list of compound names which will then be searched for in the database. The search can be processed in the current database only or in all databases which are opened in the databases dialog, e.g. all databases which are registered in the EXTRAS/DATABASE SETTINGs dialog for the corresponding quantum chemical level.

3.4 New Molecule

The NEW MOLECULE tool (formerly COMPOUND WIZARD) is a universal tool for adding known or unknown compounds to your calculation. It also offers some advanced search options for the database.

The wizard makes use of free external tools, which have not been developed by COSMOlogic. We cannot guarantee that all offered functionality will operate smoothly under all possible circumstances. The NEW MOLECULE tool will be very useful for setting up COSMOtherm calculations in many cases involving compounds that are not readily available from the COSMO databases.

The following list gives an overview of the general options:

- Search all databases by names, CAS-Number, SMILES, drawn structures or loaded structures.
- Search the free Internet database ChemSpider by compound name.
- Draw two dimensional or three dimensional structures for database searches or to start quantum mechanical COSMO calculations.
- Load structures from files.
- Convert SMILES or two dimensional structures to three dimensional structures. This option is only available in the search results window.
- Modify an existing molecule to generate a new one.

**Start Search:**
Inside the first window, the name, CAS-Number or SMILES can be given to start the corresponding database search. Alternatively a structure can be drawn or a molecule taken from file. If a quantum mechanical COSMO calculation has been started previously, you can also directly go the QM Results panel.

The search for name, SMILES, CAS-Number or a drawn 2D structure will always lead to the DB Search Results window. The 2D structure is therefore converted to a SMILES.

![Image of COSMO logic interface]

The above search for ethanol led to a number of perfect hits (green) and near hits (yellow). The right column shows in which database the compound was found. After selecting a compound from the results list, several options are available:

- Use one of the structures for a quantum mechanical COSMO calculation to obtain a .cosmo file for a different parameterization level.
- Use structure for 2D or 3D editing if only similar molecules have been found.
- Add the molecule to the compound list for the COSMOtherm calculation.
- The Internet database ChemSpider can be used to find a SMILES for a given name. This option is only available if the search was started by compound name.
- If the search was started by SMILES or a drawn structure, you can also choose to convert the SMILES to a 3D structure.

In case that a drawn 3D structure or a molecule file should be used for the search, the 3D molecule editor will open. Only two options are available after a 3D geometry has been build or opened from file. Either use the structure for a quantum mechanical COSMO calculation or restart the search with a SMILES generated from the 3D structure. The latter option is useful to search the databases for a given 3D Structure.
**Start quantum mechanical calculations:**

If the requested compound could not be found inside the database or not with the needed basis set level (e.g. TZVP), a QM calculation is necessary to generate the appropriate .cosmo file. The QM Calculation panel can either be accessed from the DB Search Results or from the 3D editor.

Inside the QM Calculation panel you can choose the parameterization level (SVP or TZVP), the type of calculation (gas phase or COSMO or gas phase and COSMO) and the output file type (.cosmo or .ccf). With **START CALCULATION (NETWORK)** the QM calculation can also be done on a remote machine (Linux only).

---

**General informations about a remote system are:**

1. The name or the IP address of a remote machine
2. User name on that machine
3. Password to log in (will not be saved to disk!)

Those three fields have to be filled in first. To check if the settings are correct, and if a connection can be established, click on check password settings. The Wizard will try to log in and determine the home directory of the user which has been given in the User field.

If the connection has been successful, the home directory will be added to the ‘Path’ field.

4. Path provides the information in which path on the remote machine the job shall run – a local fast disk should be chosen here.
5. TURBODIR has to be set to the TURBOMOLE installation directory on the target system. The default behaviour of ssh when starting remote jobs without an explicit shell or terminal is such that not all settings on the remote system are sourced. So it is very likely that your Turbomole settings (like $TURBODIR, $PATH, etc) are not available in such a case. It is therefore unavoidable to set the PATH to the Turbomole directory on the remote system by hand.
6. The number of CPUs can be left unchanged. Note that this field is not to tell TURBOMOLE how many CPUs or cores are available in general, but how many CPUs shall be used for the calculation of each job!
7. If the background option is activated, the jobs are started with nohup. In this case, TURBOMOLE does not get a notice when the job has finished, so it has to check actively if the job is still running or not. The frequency for those checks can be given in minutes.

The queueing system option is described in the next chapter.

After settings 1-5 are complete, click on save settings, and the machine with user name and paths will be added to the pull down list of machines for future usage.

---

After the QM calculation is finished you can transfer the molecules to your compound list to start COSMOtherm calculations. The QM Results panel can be directly accessed from the wizard Starting window. Note that for later use in COSMOtherm, blanks should not be used in compound names.
Example: Glycofurol

During a conference you heard that Glycofurol is a nice solvent for special purposes, but unfortunately you don’t know anything but its name. Open the New Molecule tool and enter glycofurol into the name field, press enter or the SEARCH BY NAME button to see if the substance is already inside your databases. Let’s assume it is not there and try the SEARCH THIS NAME IN CHEMSPIDER button.

A tutorial for using ChemSpider is also available directly within the New Molecule tool. ChemSpider will find the SMILES O(CCO)CC1OCCC1 for glycofurol. Enter it into the corresponding field and CONTINUE WITH THIS STRUCTURE/SMILES. Now a structure search based on the SMILES string will be performed. Assuming that the structure can again not be found in the database the GENERATE 3D STRUCTURE FROM SMILES option is your last resort. The 3D Structure will be generated and a new COSMO calculation can now be conducted by choosing USE FOR QM CALCULATION.

3.5 Conformers

Molecules often can adopt more than one conformation. For COSMO-RS, only conformers with different σ-profiles are relevant. For each of these conformers, an individual COSMO file is required for the compound input. If compounds are selected from the databases and the checkbox in the USE CONF. column is checked, all existing conformer COSMO files will be selected automatically. If the ACTIVATE CONFORMER TREATMENT checkbox in the compound section is checked, the conformers will be weighted internally by COSMOtherm using their COSMO energies and their chemical potentials.

If you intend to use your own COSMO files for conformers please be aware that in order to be identified as conformers by COSMOthermX automatically, the names of the files must follow a convention. It is also possible to define conformers of a compound manually without following the name convention as described in section 2.5 of this document. Select the corresponding compounds from the databases of the File Manager, highlight them in the compound list using the Control key and the left mouse button, and choose COMBINE TO CONFORMER SET from the right mouse button menu. Note that this option is available only if the conformer treatment is activated.

Other useful options for calculations involving conformers from the GLOBAL OPTIONS dialog are the print option PRINT CONFORMER INFO (wconf) and the program control option USE ALL CONFORMER FILES FOUND FOR A COMPOUND (autoc). For more information on conformer input refer to the COSMOtherm User’s Manual, section 2.2.2.
3.6 Visualization of $\sigma$-surfaces, $\sigma$-profiles, and $\sigma$-potentials

The 3D screening charge distribution on the surface of a molecule $i$ can be used to qualitatively describe the molecule. Polarity, hydrogen bonding and lipophilicity or hydrophilicity can be visualized on the molecular surface. The surface screening charges can be converted into a distribution function, the $\sigma$-profile $\rho(\sigma)$, which gives the relative amount of surface with polarity $\sigma$ on the surface of the molecule. The $\sigma$-potential, as calculated from eq. 5, can also be visualized.

Example: COSMO charge surface visualization, $\sigma$-profile and $\sigma$-potential.

Select the compounds you would like to visualize. In the EXTRAS/GLOBAL OPTIONS/ADDITIONAL OUTPUT FILES menu, check GENERATE VRML ($\sigma$-SURFACE) for the COSMO charge surface visualization. Check $\sigma$-PROFILES (.PRF) and $\sigma$-POTENTIALS (.POT) for the generation of the $\sigma$-profile and $\sigma$-potential files. After the calculation has finished, use TOOLS/COSMOVIEW to display the $\sigma$-surfaces of the molecules. For more information on COSMOview, please refer to the section “Using COSMOview”. $\sigma$-Profiles and $\sigma$-potentials can be plotted in a spreadsheet program.

The COSMO charge surface, the $\sigma$-profile and the $\sigma$-potential of a compound can also be visualized from the FILE MANAGER or the databases. Highlight a compound with a left mouse button click. With a right mouse button click you can choose VIEW SIGMA SURFACE or VIEW SIGMA PROFILE/POTENTIAL. A COSMOtherm calculation is run in the background and a window displaying the chosen property will pop up.

Further information on $\sigma$-profiles and $\sigma$-potentials refer to the COSMOtherm User’s Manual, chapter 5.

3.7 Using COSMOview

COSMOview can be used to display .wrl files generated by COSMOtherm. COSMOview is included in COSMOthermX and can be accessed via TOOLS/COSMOVIEW from the main menu.

OPEN: Open a previously generated VRML. Alternatively, open a VRML file by right-clicking a compound and selecting view sigma surface or VIEW MOLECULE from the context menu.

SAVE GRAPHICS: Graphics can be saved, optionally with a transparent background.

HOME: Reset the camera to its initial position.
**Viewer Settings**: Change color, labels, atom settings, bond settings and so on.

**Arrange Windows**: The arrangement of multiple COSMOview windows can be changed.

**Display Sigma Surface**: If a surface is loaded, it can be hidden and shown again.

**Display Molecule**: The molecule structure can be hidden and shown again.

**Display Segments**: Instead of closed object surfaces it is possible to show only the wire frame. This option works for surfaces, atoms and bonds.

**Use Charge Density Picker**: To get an idea of the quantitative surface charge density at a given point, you can activate the charge picking mode and move the cursor over the σ-surface. A slider at the right-hand side will display the charge density at the spot you are pointing on. However these values can only be approximated and are not guaranteed to be entirely precise. This is mainly an effect of interpolation between the reduced grid size compared to .cosmo files. Please also note that since COSMOview uses an internal color correction, the legend produced will not be applicable to images obtained by other means than COSMOview, e.g. third-party browser plug-ins.

**Display Info**: Shows a few data on grid size and charge for

**Movement**: Molecules can be moved using the mouse buttons

  - Rotate the molecule by dragging the mouse with the left button pressed. If you move the mouse quickly, you can give the molecule a spin to have it turn by itself.
  - Zoom in and out with the right mouse button pressed or simply by turning the mouse wheel.
4 Property input

The larger section of the main window offers a selection of property cards. Inside each card you can adjust parameters like temperature, mole fraction etc. to your issue. Input settings from the property cards are transferred to the Property Selection panel with the ADD button. Changes in the MIXTURE OPTIONS dialog are taken into account for the property if the USE MIXTURE OPTIONS checkbox is activated. The COSMOTHERM calculation is started from the RUN button in the Property section, from the RUN menu or from the shortcut bar.

By default, COSMOTHERM produces two sorts of output files: The COSMOTHERM output file filename.out and a file filename.tab which contains the calculated property information in tabulated form. These files will automatically pop up after the calculation has finished. Additional output files will be written if the corresponding options in EXTRAS/GLOBAL OPTIONS are activated. These output files may contain $\sigma$-moments (.mom), atomic $\sigma$-moments (.moma), $\sigma$-profiles (.prf), or $\sigma$-potentials (.pot).

The phase definition subsections in all property panels have a context menu. With a right mouse button click, phase compositions can be copied and pasted into another property panel.

The explanations, examples and the results in this Tutorial refer to the full version of COSMOTHERM. Some examples cannot be calculated with the DEMO/EDU version. The example results might be calculated with different versions of COSMOTHERM and can deviate numerically from the current version.

4.1 Mixture: Calculation of compound properties in mixture

This option toggles the COSMOTHERM calculation of interaction energy terms at the given temperature and mixture composition. For all compounds in the compound list, the following terms will be calculated:

- Chemical potential $\mu_i$ of the compound in the mixture from eq (6).
- Log10(partial pressure [mbar])
- Free energy of the molecule in the mixture (E_COSMO+dE+Mu)
- Total mean interaction energy in the mix ($H_{\text{int}}$): The mean interaction enthalpy of the compound with its surrounding, i.e. the interaction enthalpy of the compound which can be used to derive heats of mixing and heats of vaporization.
- Contributions to the total mean interaction energy:
  - Misfit interaction energy in the mix ($H_{\text{MF}}$).
  - H-Bond interaction energy in the mix ($H_{\text{HB}}$)
  - VdW interaction energy in the mix ($H_{\text{vdW}}$)
  - Ring correction

For details on the calculation of the energy terms and contributions please refer to the COSMOTHERM User’s Manual, section 1.1.

Furthermore, COSMOTHERM allows for the computation of the contact probability of molecules and molecule surface segments in arbitrary mixtures. The checkbox COMPUTE CONTACT STATISTICS can be checked to obtain a more detailed contact interaction statistics of all segments of molecules A and B. For more information on the calculation of contact statistics please refer to the COSMOTHERM User’s Manual, section 5.7.
Special Application: Cocrystal screening with COSMOthermX

Co-crystals are organic crystals consisting of a drug (active pharmaceutical ingredient, API) and a coformer, usually a dummy compound with a good solubility, which may show overall improved physical and chemical properties as compared to the pure drug crystal itself. In the following it will be demonstrated how COSMOthermX may also be used for computational cocrystal screening and for estimation of cocrystal solubilities.

Calculation of the excess enthalpy $H_{\text{ex}}$ to predict cocrystal formation

To compute the likelihood of cocrystal formation we start from a virtually subcooled liquid of the cocrystallization components and neglect the long-range order in the crystal. It has been shown recently that this approximation gives rather good results on a set of potential cocrystal formers. An important quantity in this respect is the excess enthalpy $H_{\text{ex}}$ (mixing enthalpy) obtained when mixing the pure component A and B to yield the subcooled cocrystal liquid $A_n B_m$:

$$H_{\text{ex}} = H_{AB} - x_m H_{\text{pure.}A} - x_n H_{\text{pure.}B}$$

$H_{AB}$ and $H_{\text{pure}}$ represent the molar enthalpies in the pure reference state and in the m:n mixture, with mole fractions $x_m = m/(m+n)$ and $x_n = n/(m+n)$. The excess enthalpy $H_{\text{ex}}$ of an API and conformer pair gives a good estimate of the propensity to cocrystallize. Technically, using e.g. the MIXTURE panel within COSMOthermX, one has to perform three calculations to obtain $H_{\text{ex}}$: one for each of the pure components A and B, and one mixture calculation for A and B with the given stoichiometry in the subcooled liquid consisting of the mixture of A and B. Subsequently one will obtain 4 different enthalpies (abbreviated $H_{\text{int}}$ within COSMOtherm):

- $H_{\text{pure.A}}$: enthalpy of pure component A
- $H_{\text{pure.B}}$: enthalpy of pure component B
- $H_{AB}^A$: enthalpy of component A in liquid mixture AB
- $H_{AB}^B$: enthalpy of component A in liquid mixture AB

$H_{AB}$ is obtained additively from the enthalpies of the ingredients of the subcooled liquid AB. The excess enthalpy is then obtained via the following equations:

$$H_{AB} = x_m H_{AB}^A - x_n H_{AB}^B$$

$$H_{\text{ex}} = x_m H_{AB}^A + x_n H_{AB}^B - x_m H_{\text{pure.A}} - x_n H_{\text{pure.B}}$$

Example 1: Excess enthalpy of a 1:1 mixture of 4-pyridinecarbonitrile (=4-cyanopyridine) with 1,3,5-trihydroxybenzene

For the calculation of $H_{\text{ex}}$ for the compounds 4-pyridinecarbonitrile (4-pyridinecarbonitrile.cosmo), compound A, and 1,3,5-trihydroxybenzene (1,3,5-trihydroxybenzene0.cosmo and 1,3,5-trihydroxybenzene1.cosmo), compound B, with the TZVP parameterization (BP_TZVP_C30_1201.ctd) one would proceed as follows: First load the two components from the database with the right parameterization into your compound list and select the MIXTURE panel. Then 3 calculations have to be set up: one for the pure component A and one for pure B, and another one for the mixture AB (see figure below). Since we have a 1:1 stoichiometry the mole fractions for the mixture AB are $x_m = x_n = 0.5$. The temperature should be set to a value where the experiment takes place, usually this would be the room temperature.
When the calculation has finished, there will be several spreadsheets available showing the results of the calculation. We should get the following enthalpies from the columns labeled “H_int”:

\[
\begin{align*}
H_{\text{pure,A}} &= -5.13 \text{ kcal/mol (pure 4-pyridinecarbonitrile)} \\
H_{\text{pure,B}} &= -11.29 \text{ kcal/mol (pure 1,3,5-trihydroxybenzene)} \\
H_{\text{A-B}} &= -9.37 \text{ kcal/mol (4-pyridinecarbonitrile in mixture AB)} \\
H_{\text{B-A}} &= -13.32 \text{ kcal/mol (1,3,5-trihydroxybenzene in mixture AB)}
\end{align*}
\]

Using equation (2) we get:

\[
H_{\text{ex}} = 0.5 \times (-9.37) + 0.5 \times (-13.32) - 0.5 \times (-5.13) - 0.5 \times (-11.29) = -3.14 \text{ kcal/mol}
\]

The excess enthalpy for the mixture 4-pyridinecarbonitrile/1,3,5-trihydroxybenzene is highly negative, indicating that it will likely form a cocrystalline compound.

In a similar way the excess free energy \( G_{\text{ex}} \) for the compound \( A_mB_n \) can be calculated, which is obtained from the respective chemical potentials (labeled “\( \mu \)” in the COSMOtherm .tab file):

\[
G_{\text{ex}} = x_m \mu_A^A + x_n \mu_B^B - x_m \mu_{\text{pure,A}} - x_n \mu_{\text{pure,B}}
\]

For the system 4-pyridinecarbonitrile/1,3,5-trihydroxybenzene the excess free energy at the TZVP parameterization is \( G_{\text{ex}} = -0.85 \text{ kcal/mol} \). Different cocrystal systems may now be ordered according to their excess quantities, here \( H_{\text{ex}} \) and \( G_{\text{ex}} \) will give usually a rather similar ranking.

Alternatively, the VLE/LLE panel may be used to get a phase diagram of the System AB. In the VLE/LLE approach excess quantities are calculated at once for several compositions, which makes this procedure more time consuming.

**Calculation of Cocrystal Solubility**

Another useful quantity which may be estimated with the help of COSMOtherm is the solubility of the cocrystal. For the cocrystal solubility the same equations as for the salt solubility can be used.
For the solubility of a cocrystal/salt the following equation (eq. 2.3.1a from the COSMOTHERM manual) is valid:

$$\log_{10}(x_{AB}^{\text{SALT}}) = \left[ \mu_{AB}^{(0)} - \mu_{AB}^{(0)} - \max(0, \Delta G_{\text{fus}}) \right] / \nu_{\text{tot}} \right] / (RT \ln(10))$$

In this case $x_{AB}^{\text{SALT}}$ is the solubility of the AB cocrystal (salt) with stoichiometry $m:n$, referring to one mole of cocrystal (salt). Here $\mu_{AB}^{(0)} - \Delta G_{\text{fus}}$ is the chemical potential of the pure/salt crystal AB and $\mu_{AB}^{(s)}$ is the chemical potential of AB in the solvent, and $\nu_{\text{tot}} = m+n$. Please refer to the COSMOTHERM manual for further details.

The solubility of a cocrystal may be calculated with the help of the SALT SOLUBILITY panel as shown below.

**Example 2: Estimated solubility improvement of the cocrystal systems carbamazepine – niacinamide and carbamazepine – oxalic acid in water**

The following compounds have to be selected from the TZVP database: carbamazepine, niacinamide (=nicotinamide), oxalic acid and water. The SALT SOLUBILITY Panel should be selected, and the solvent water should be specified. As we do not know the free energy of fusion $\Delta G_{\text{fus}}$, we neglect it in this case, setting it to zero ($\Delta G_{\text{fus}}=0$). Of course we will not get absolute solubilities anymore, but we still can predict solubility trends with this approximation. Furthermore, we make the approximation to work in infinite dilution, i.e. non-iteratively.

After adding the systems carbamazepine-nicotinamide (CBZ-NCT, 1:1) and carbamazepine-oxalic acid (CBZ-OXA, 2:1), a solubility calculation can be started via the Run button (in a true predictive calculation the stoichiometry of the potential crystal would of course be unknown, and one should start from a 1:1 ratio). After the calculation has finished, we obtain for the logarithmic mole fraction solubility of the CBZ-NCT cocrystal $\log_{10}x = -1.93$ and for CBZ-OX $\log_{10}x = -2.86$ from the spreadsheet (for the different and sometimes intricate definitions of solubility please refer to the manual). The solubility of the component CBZ itself in the CBZ-OXA cocrystal is thus twice as large, because we have a 2:1 stoichiometry and the logarithmic mole fraction is approximately $\log_{10}x = -2.56$. The solubility of CBZ-NCT stays unchanged as we have 1:1 stoichiometry. Comparing
the calculated logarithmic solubilities with experimental literature data (Good, D. J. & Rodríguez-Hornedo, N. Crystal Growth & Design 2009, 9, 2252-2264) for some carbamazepine cocrystals results in the following plot:

The cocrystal system CBZ-NCT shows the highest improvement in solubility as compared to the solubility of the pure drug carbamazepine.

Please note that no correction for dissociation of the acids or the API was applied. The data for the remaining cocrystals of Good et al. is shown below with the solubility of the cocrystal (log10(x_CC)), the solubility of carbamazepine itself (log10(x_CBZ)) and the experimental ratio of the solubility of carbamazepine within the cocrystal to the solubility of pure carbamazepine (log10x_CC/x_CBZ).

<table>
<thead>
<tr>
<th>API</th>
<th>coformer</th>
<th>log10(x_CC)</th>
<th>log10(x_CBZ)</th>
<th>exp</th>
</tr>
</thead>
<tbody>
<tr>
<td>carbamazepine</td>
<td>nicotineamid (NCT)</td>
<td>-1.93</td>
<td>-1.93</td>
<td>2.182</td>
</tr>
<tr>
<td>carbamazepine</td>
<td>oxalic acid (OXA), 2:1</td>
<td>-2.86</td>
<td>-2.56</td>
<td>0.752</td>
</tr>
<tr>
<td>carbamazepine</td>
<td>glutaric acid (GTA)</td>
<td>-2.33</td>
<td>-2.33</td>
<td>2.070</td>
</tr>
<tr>
<td>carbamazepine</td>
<td>saccharin (SAC)</td>
<td>-3.14</td>
<td>-3.14</td>
<td>0.416</td>
</tr>
<tr>
<td>carbamazepine</td>
<td>succinic acid (SUC), 2:1</td>
<td>-2.80</td>
<td>-2.49</td>
<td>0.717</td>
</tr>
<tr>
<td>carbamazepine</td>
<td>malonic acid (MLN), 2:1</td>
<td>-2.78</td>
<td>-2.47</td>
<td>1.409</td>
</tr>
<tr>
<td>carbamazepine</td>
<td>salicylic acid (SLC)</td>
<td>-3.24</td>
<td>-3.42</td>
<td>0.296</td>
</tr>
</tbody>
</table>
4.2 Vapor pressure

COSMO-RS allows for the estimation of pure compound vapor pressures. The energy of the gas phase $E_{\text{gas}}^i$ is required for the calculation of the chemical potential in the gas phase, $\mu_{\text{gas}}^i$. $E_{\text{gas}}^i$ can be taken from a gas phase quantum chemical calculation or empirically estimated by COSMOtherm. If possible, the quantum chemically calculated value of $E_{\text{gas}}^i$ should be used: Check ENERGY FILE in GLOBAL OPTIONS (this is the default setting).

The vapor pressure option enables the computation of vapor pressures for a given temperature or a temperature range and a fixed mixture concentration. The number of points in a temperature range can be up to 100, default value is 10. Since the strength of COSMOtherm vapor pressure prediction is in the temperature dependency of the vapor pressure, and not in the prediction of the absolute vapor pressure at a given temperature, it is possible to scale the vapor pressure prediction using a pair of reference temperature and pressure. Data can be entered in the USE REFERENCE DATA TO SCALE VAPOR PRESSURE PREDICTION subsection of the vapor pressure panel.

For each temperature and compound in the mixture the partial vapor pressures, the chemical potential of the compound in the gas phase and its enthalpy of vaporization are computed and written to the COSMOtherm output file. The total vapor pressure of the mixture is written to the COSMOtherm table file in tabulated form $p_{\text{Vap}}$ vs $T$. In addition the total chemical potentials of the liquid $\mu_{\text{liquid}}^{\text{tot}}$ and of the gas phase $\mu_{\text{gas}}^{\text{tot}}$, as well as the heat of vaporization of the mixture $\Delta H_{\text{vap}}$, are written to the COSMOtherm table file. If the given temperature is below a compounds melting point, COSMOtherm will compute the partial and total vapor pressure of the subcooled melt. However, if experimental data on a compounds solid state properties are available ($T_{\text{melt}}$ and $\Delta G_{\text{fus}}(T)$), COSMOtherm will compute the compounds partial sublimation pressures ($p_{\text{S}}^i$/1 bar = $\exp[(\mu_{\text{gas}}^i - \mu_{\text{S}}^i + \Delta G_{\text{fus}}^i)/RT]$), and write the computed partial sublimation pressures and the according enthalpies of sublimation to the COSMOtherm output and table files.

If three or more temperature points were calculated in a vapor pressure curve, the total vapor pressure will be fitted to Antoine’s vapor pressure equation $\ln(p) = A - B / (C + T)$, where $T$ is the temperature in [K] and $A$, $B$ and $C$ are coefficients. The coefficients are written to the COSMOtherm output and table files.

Note that since the Vapor Pressure option in COSMOtherm is intended for the prediction of vapor pressures data from .vap files are not used even if this is indicated in the options. However, if vapor pressure data are available from a .vap file, they will be printed to the last column of the table file for comparison.

Example: Calculation of a vapor pressure curve

Select methanol from the TZVP DB. Make sure that “Energy file” in the GLOBAL OPTIONS is checked. Set the temperature range from 0°C to 70°C and the mole fraction as pure methanol. Pressing ADD transfers the selected vapor pressure settings to
the property panel at the bottom. Start the calculation by pressing RUN.

The total vapor pressures can be read from the output file and from the table file which will open in a separate window after the calculation has finished.

### 4.3 Boiling point

This option enables the iterative optimization of the equilibrium temperature for a given vapor pressure. The temperature of the system is varied and for each temperature the vapor pressure is calculated. This is repeated until the COSMOtherm prediction of the total vapor pressure and the specified pressure in the input file is below a certain threshold. During the procedure, the partial vapor pressures of the compounds are written to the COSMOtherm output file. When the required threshold is met, i.e. convergence is reached, the total vapor pressure of the mixture is written to the COSMOtherm table file.

### 4.4 Activity coefficient calculation

This option computes the activity coefficients of different compounds in the selected solvent or solvent mixture. For the calculation of the activity coefficient at infinite dilution, the mole or mass fraction of the compound of interest has to be set to zero in the composition of the solution. The chemical potentials $\mu_j^{(P)}$ of all pure compounds $j$ and the chemical potentials $\mu_i^{(L)}$ in the liquid phase (compound $i$ or compound mixture, respectively) are calculated. The activity coefficients are then calculated as $\ln(\gamma_j) = (\mu_j^{(L)} - \mu_j^{(P)}) / RT$.

It is also possible to calculate the activity coefficients at a given finite concentration. This is achieved by setting the mole or mass fraction of the compound of interest to the required value in the composition of the solvent. The compound in question is thus treated as part of the solvent.

By checking the ADVANCED SETTINGS a reference state for the activity coefficient can be chosen.
Example: Calculate the activity coefficient of aspirin in water

Choose the compounds, aspirin and water, from the TZVP database. By default, both aspirin conformers are selected from the database and the conformer treatment is activated to account for a conformer mixture. Then, set the temperature to the desired value, 25 °C, set the water mole fraction to 1.0 (check PURE), and transfer the selection to the property panel with the Add button. Windows displaying the output and table files will open after the calculation has finished. By checking the ADVANCED SETTINGS a reference state for the activity coefficient can be chosen.

4.5 Henry law coefficient calculation

This option allows for the computation of Henry law coefficients $H^0_j$ in compound $i$. The chemical potentials $\mu_j^0$ of all pure compounds $j$ and the chemical potentials $\mu_j^{(0)}$ at infinite dilution in compound $i$ are calculated. Then the Henry law coefficients $H_j^0$ for all compounds $j$ are calculated from the activity coefficients and the vapor pressures of the compounds are written to the COSMOtherm output and table files. It is also possible to calculate the Henry law coefficients at a given finite concentration, i.e. in a mixture of solvents.

The Henry law coefficient depends on the pure compound vapor pressure. For each compound, there are several possibilities to calculate or approximate this property. In order of increasing accuracy you might:

- Use the COSMOtherm approximation of the vapor pressure using the approximated gas phase energy of the compound. This is the default if no .energy file is present and requires no additional input.
- Use the COSMOtherm approximation of the vapor pressure using the exact gas phase energy of the compound from the .energy file. This option is set by default. (GLOBAL OPTIONS: check ENERGY FILE)
- Use the Wagner, DIPPR, or Antoine equation $\ln(p_j^0) = A - B / (T + C)$ to compute the vapor pressure at the given temperature. If available, data for these equations will be read from the .vap file if the VAPOR PRESSURE / PROPERTY FILE option is checked in the GLOBAL OPTIONS. Data can also be entered in the COMPOUND PROPERTIES dialog from the context menu of the compound list.
- Enter the exact value of the vapor pressure for this temperature via the COMPOUND PROPERTIES dialog from the context menu of the compound list

The Henry Law Coefficient option also allows for the calculation of the Gibbs free energy of solvation in different reference frameworks.
4.6 Gas solubility

With this option the solubility of a gas in a solvent or solvent mixture can be calculated in an iterative procedure. For each compound \( j \) the mole fraction \( x_j \) is varied until the partial pressure of the compound \( p_j = p_j^0 x_j \gamma_j \) (with the activity coefficient \( \gamma_j \) and the pure compound vapor pressure \( p_j^0 \)) is equal to the given reference pressure \( p \).

Like the calculation of the Henry Law Coefficient, the calculation of a gas solubility requires the knowledge of the pure compound vapor pressure. For options to give the pure compound vapor pressure please refer to the Henry law coefficient calculation section.

Example: Gas solubility of methane in water

Select the compounds, water and methane, from the TZVP database. Set the temperature, 25 °C, and the pressure, 1013.25 mbar, in the Gas-Solubility card. Set the solvent composition to pure water. Transfer the settings to the “Property Selection” by pressing ADD and run the program.

Windows displaying the output and table files will open after the calculation has finished.

4.7 Solubility

The SOLUBILITY option allows for the automatic computation of the solubility of a liquid or solid compound \( j \) in a solvent \( i \). Within the calculation all compounds are also considered solutes, even the solvents themselves. This approach is optimized for the calculation of many solutes in a limited number of solvents. If the solubility of a single solute should be calculated in a large number of solvents the SOLVENT SCREENING panel offers an appropriate solution.

The solubility is calculated from:

\[
\log_{10}(x_j) = \log_{10}\left(\exp\left((\mu_j^{\text{pure}} - \mu_j^{\text{solvent}} - \Delta G_{j,\text{fusion}})/RT\right)\right)
\]

COSMOtherm can directly calculate the chemical potentials \( \mu_j^{\text{pure}} \) of all pure compounds \( j \) and the chemical potentials \( \mu_j^{\text{solvent}} \) at infinite dilution. The free energy of fusion \( \Delta G_{\text{fusion}} \) is zero for liquid compounds and has to be given or estimated for solid compounds. The computed solubility \( x_j^{(0)} \) is a zeroth order approximation, which is valid only for small concentrations of the solute. If the solubility is large \( (x_j > 0.1) \), \( x_j^{(0)} \) is a poor approximation, but \( x_j \) can be refined iteratively by resubstitution of \( x_j^{(0)} \) into the solubility calculation. This procedure can be repeated until the differences in the computed value of \( x_j \) are below a certain threshold. In COSMOTHERM-X, this procedure is turned on by checking the ITERATIVE calculation type in the solubility panel.
As mentioned above the Gibbs free energy of fusion \( \Delta G_{\text{fus}} \) has to be taken into account for solid solutes. \( \Delta G_{\text{fus}} \) can be read from the vapor pressure / property file or from the compound line in the compound input section of the COSMOtherm input file. A temperature dependent heat of fusion can also be calculated if the compounds enthalpy or entropy of fusion \( \Delta H_{\text{fus}} \) or \( \Delta S_{\text{fus}} \) and melting temperature are known. This will be done automatically if the USE HEAT CAPACITY OF FUSION ESTIMATE checkbox is checked. To add \( \Delta G_{\text{fus}} \) (or \( \Delta H_{\text{fus}} \) or \( \Delta S_{\text{fus}} \) and \( T_{\text{melt}} \)) to the compound input lines open the COMPOUND PROPERTIES dialog for the compound. Alternatively, \( \Delta G_{\text{fus}} \) can be estimated by COSMOtherm using a QSPR approach. QSPR parameters are read from the parameter file, if possible, but can also be given explicitly when the ADVANCED SETTINGS checkbox is checked. Since one of the QSPR parameters is the chemical potential of the compound in water, water has to be included in the compound list even if it is not present in the system. For further information refer to the COSMOtherm User’s Manual, section 2.3.4.

By default, the output of the solubility option is in logarithmic mole fractions, \( \log_{10}(x) \). Additionally, the mass based solubility is written to the table file. The definition applied for the mass based solubility can be changed in the ADVANCED SETTINGS:

**Definition 1:**

\[
  w_j = x_j \cdot \frac{MW_j}{MW_{\text{water}}} 
\]

(popular in drug solubility screening)

**Definition 2:**

\[
  w_j = x_j \cdot \frac{MW_j}{(1-x_j)MW_{\text{water}}} 
\]

(popular in engineering, default)

**Mass fraction:**

\[
  w_j = x_j \cdot \frac{MW_j}{x_j MW_j + (1-x_j)MW_{\text{water}}} 
\]

(popular in chemistry)

Definition 1 is in principle an approximation to the default definition 2 for low solubilities.

For further information about the definition of the mass based solubility, refer to the COSMOtherm User’s Manual, section 2.3.

The decadic logarithm of the molar solubility (\( \log_{10}(S \ [\text{mol/l}] ) \)) will be written to an additional column in the table file. To calculate the solubility in the more commonly used g/L units the densities have to be known and will therefore be estimated by COSMOtherm. Please note, that this estimation introduces an additional error to the results.

In aqueous systems, acidic or basic solutes can dissociate. The dissociated species typically have properties different from the property of the undissociated compound, thus affecting the apparent solubility of the compound. Additional input for this suboption can be given in the COMPUTE DISSOCIATION CORRECTION subsection of the solubility panel. For the calculation of the dissociation correction either input of the solute pKa value or a cosmo file of the dissociated solute is required.
Example: Solubility of glycol in hexane

In principle, there are two ways to do this. If the liquid solubility of the compound in question is expected to be low, you can use the SOLUBILITY option. Alternatively, you can calculate the Liquid-Liquid Equilibrium (LLE) and search the phase diagram for the LLE point. The LLE point can also be printed to the end of the table file.

First, select the compounds, glycol and hexane (hexane.cosmo) from the TZVP database. Make sure that the ACTIVATE CONFORMER TREATMENT option is checked. In the Solubility card, set the temperature, 25 °C, the state of the solute to LIQUID and the calculation type, ITERATIVE. Then, check PURE in the “Solvent” paragraph for hexane. Add the settings to the property panel and run the calculation. The COSMOtherm output and table files will pop up after the calculation has finished.

4.8 Solvent screening

The solvent screening option works similar to the solubility option and is based on the same thermodynamics (Please note, that the definition for the mass based solubility output is fixed to the default, see solubility). It is especially adjusted to fit solvent screening purposes, i.e. to predict the solubility of one solute in a list of solvents. It is also possible to define a solvent mixture grid for the screening.

For solid compounds, a relative screening option is available in addition to the absolute solubility. As in general the Gibbs free energy of fusion $\Delta G_{\text{fus}}$ is one of the largest error sources for the solubility of solids, the relative screening option avoids this error by neglecting $\Delta G_{\text{fus}}$. The result is a relative solubility ranking. The numbers thus given contain all relevant information about solute-solvent interactions, which means that they represent trends of the solubility in different solvents, but they are no absolute solubility values.

In contrast to most other options the screening cannot be done in a single COSMOtherm run. The result is that many input files have to be generated and stored. For loading a solvent screening run with the OPEN JOB option from the FILE menu, the file type has to be set to COSMOtherm screening files.

Liquid solutes and relative solubility of solid compounds:
The use of the solvent screening option is straightforward in these cases. Select the solute, choose a non iterative (infinite dilution) or iterative (finite concentration) type of calculation and add all solvents or solvent mixtures to the Property Section. Please note, that the RELATIVE SCREENING is always calculated in infinite dilution (non iterative).
In addition to the logarithm of the solubility in mole fractions, the chemical potential of the pure solute \( \mu^\text{pure} \), the chemical potential in the solvent \( \mu^\text{solvent} \) and the solubility in gram per gram solvent are given (default definition, see also Solubility). The other values are included for general information purposes.

To illustrate the meaning of relative solubility the following figure shows benzoic acid in different solvents at 25°C. The graph is not automatically generated.

The logarithm of the best solubility is set to 0 and all other solvents are given relative to the best solvent. In the above example tetrahydrofuran (thf) is clearly the best solvent. A solvent with a log10(\( x^\text{RS} \)) value of -1.00 yields a solubility, which is decreased by a factor of 10 compared to thf, e.g. the solubility in water is about 2000 times worse.

Absolute solubilities for solid compounds:

When the absolute solubility of a solid solute should be computed, the free energy of fusion cannot be neglected. As the absolute solubility depends on the concentration of the solute itself, the iterative procedure (finite concentration) is always used. After selecting the corresponding options (State of solute = solid, and Absolute values is checked) three different options for the estimation of \( \Delta G^\text{fus} \) appear:

- An experimental solubility in one of the solvents or solvent mixtures can be used to determine \( \Delta G^\text{fus} \) of the solute. If the reference solvent is chosen carefully, i.e. it is known, that COSMOtherm results are usually in good agreement with experimental data for this solvent, this option might give results closest to measurements due to error cancellation. If,
however, COSMOTHERM fails to make good predictions for the reference solvent, the errors might add up instead of canceling. The experimental solubility can be given in mass fraction, mole fraction or mg/g solvent according to definition 2 ($w_j = x_j \cdot \frac{MW_j}{(1-x_j)MW_{solvent}}$).

In rare cases it might happen, that the reference solvent leads to a negative $\Delta G_{fus}$. This unphysical value is by default allowed for by applying ALLOW NEGATIVE FREE ENERGY OF FUSION in the GLOBAL OPTIONS dialog. If this option is not set, COSMOTHERM will use $\Delta G_{fus} = 0.0$ instead. The original reference solubility cannot be reproduced then and all resulting solubilities will be shifted.

- An experimentally measured value for $\Delta G_{fus}$ or (or $\Delta H_{fus}$ or $\Delta S_{fus}$ and $T_{melt}$) can be entered or taken from the property (.vap) file of the solute. The accuracy of this method depends on extrapolation of $\Delta G_{fus}$ at the melting point to the temperature where the solution is calculated.
- A quantitative structure property relationship (QSPR) can be used. The predictive accuracy of this option is on average the worst of the three choices.

**Screening of solvent mixtures**

Solvent mixtures are defined by their start and end compositions and the total number of mixtures required. Start and end compositions can be pure solvents or solvent mixtures. For each component of the solvent mixture, the concentration is evenly distributed between the respective concentrations in the start composition and in the end composition.

**4.9 Salt solubility**

In COSMOTHERM a salt $A^-C^+$ always is treated by means of its anion $A^-$ and cation $C^+$. To obtain the solubility of a salt, the chemical potentials of the individual ions $A^-$ and $C^+$ and the free energy of fusion $\Delta G_{fus}$ of the salt $A^-C^+$ have to be determined. The salt solubility is then computed from the mean chemical potentials of the ions and the free energy of fusion. COSMOTHERM will calculate a temperature dependent free energy of fusion if the enthalpy or entropy of fusion of the salt ($\Delta H_{fus}$ or $\Delta S_{fus}$ respectively) and the melting temperature ($T_{melt}$) are known. Note that the value of a salt’s free energy of fusion $\Delta G_{fus}^{AC}$ (as computed from salt enthalpy or entropy of fusion and melting point) by default is considered to be the “total” or “net” free energy of fusion of the salt. This means that the given salt free energy of fusion value is scaled by the total stoichiometry of the salt in a salt solubility (or salt SLE) computation. In the Salt Solubility card, the free energy of fusion can be entered directly or, alternatively, the enthalpy or entropy of fusion and the melting temperature can be entered. If the heat capacity of fusion $\Delta C_{p fus}$ is entered, it will be used to describe the temperature dependency of the Gibbs free energy of fusion. The computed mass and mole fraction salt solubility can be compared directly to experimental data. The default output of the salt solubility option is the mass based solubility. The definition applied for the mass based solubility can be changed in the ADVANCED SETTINGS:
Definition 1: \[ w_j = x_j \cdot \frac{MW_j}{MW_{solvent}} \] (popular in drug solubility screening)

Definition 2: \[ w_j = x_j \cdot \frac{MW_j}{(1-x_j)MW_{solvent}} \] (popular in engineering, default)

Mass fraction: \[ w_j = x_j \cdot \frac{MW_j}{x_jMW_j + (1-x_j)MW_{solvent}} \] (popular in chemistry)

Definition 1 is an approximation to the default definition 2 for low solubilities.
For further information about the definition of the mass based solubility, refer to the COSMOtherm Manual.

Example: Prediction of the solubility of NaCl in several solvents

Select the compounds, water (h2o.cosmo), ethanol (2 conformers), 1-octanol, Na (na.cosmo), and Cl (cl.cosmo), from the TZVP database. In the “Solvent” definition box, set the solvent composition to pure water. Then define the salt composition in the “Salt definition” box: Check the checkboxes in the cl and na compound lines. The stoichiometric coefficients are set to 1 automatically, which is correct for NaCl. Enter the value for the enthalpy of fusion (28.16 kJ/mol) and the melting point temperature (800.7 °C). Transfer the settings to the “Property Section” by pressing ADD. Then, change the solvent composition to pure ethanol (check PURE in the compound line for ethanol) and transfer the settings to the “Property Section” by pressing ADD. Change the solvent composition to pure 1-octanol and transfer the settings again to the “Property Selection” by pressing ADD and run the program.

4.10 Salt solubility screening

Like the solvent screening option, the salt solubility screening option is especially adjusted to fit screening purposes. The thermodynamic background is the same as in the Salt Solubility option. Both relative screening and absolute solubility applications can be done. Like in the solvent screening option, the screening cannot be done in a single COSMOtherm run. Instead, many input files have to be generated and stored. For loading a salt solubility screening run, special screening files are generated.
Liquid solutes:
The use of the salt solubility screening option is straightforward in these cases. Define the salt, choose a non iterative (infinite dilution) or iterative (finite concentration) type of calculation and add all solvents or solvent mixtures to the Property Section.

Relative solubility of solid compounds:
This option works similar to the calculation of liquid solutes. Note that the RELATIVE SCREENING is always calculated in infinite dilution (non iterative).

Absolute solubility of solid compounds:
When the absolute solubility of a salt should be computed, the free energy of fusion has to be taken into account. Since the absolute solubility depends on the concentration of the solute itself, the iterative procedure (finite concentration) is used.

After selecting the corresponding options (State of solute = solid, and Absolute values is checked) there are two options for the estimation of $\Delta G_{\text{fus}}$:

- An experimental solubility in one of the solvents or solvent mixtures can be used to determine $\Delta G_{\text{fus}}$ of the solute. This option works comparable to the corresponding option in the solvent screening panel. It is based on the same thermodynamics and has comparable limitations.
- An experimentally measured value for $\Delta G_{\text{fus}}$ (or $\Delta H_{\text{fus}}$ or $\Delta S_{\text{fus}}$ and $T_{\text{melt}}$) can be used. The accuracy of this method depends on extrapolation of $\Delta G_{\text{fus}}$ at the melting point to the temperature where the solution is calculated.

4.11 Partition coefficient calculation ($\log P / \log D$)
Partition coefficients of solute $j$ between solvents $i_1$ and $i_2$ are defined as $P^{j,2} = c'_2 / c'_1$, with $c'_1$ and $c'_2$ being the concentrations of solute $j$ in $i_1$ and $i_2$, respectively. The calculation of the partition coefficient $\log P$ is accomplished via computation of the chemical potentials $\mu_j^{(1)}$ and $\mu_j^{(2)}$ of all compounds $j$ in infinite dilution in pure compounds $i_1$ and $i_2$, respectively:

$$\log_{10}(P^{j,2}) = \log_{10}[\exp((\mu_j^{(1)} - \mu_j^{(2)}) / RT) \cdot V_1 / V_2]$$

(8)

The ratio of molar $V_1/V_2$ (also called volume quotient) will be estimated from the COSMO volumes by default, unless a value is entered in the LOG P / LOG D panel. In case the COMPUTE OCTANOL-WATER PARTITION option is selected, the ratio of molar volumes and the phase compositions will be set automatically. With the THERMODYNAMIC PARTITION option, the ratio will be set to 1, which corresponds to a logP definition by mole fractions. The input of a volume quotient will be necessary if the experimental density of at least one of the two solvent phases differs substantially from a linear interpolation form the individual phase compounds. In this case the estimate from the COSMO volumes, based on the assumption of an incompressible liquid, will be poor. Furthermore, the mutual solubility of the solvents in each other has to be taken into account when computing $\mu_j^{(1)}$ and $\mu_j^{(2)}$. It is possible to give finite concentrations in the solvent mixture section.

In systems with one aqueous phase it is possible that acidic or basic solutes dissociate in the water phase. The dissociated species typically have partition properties different from the partition property of the undisassociated compound, thus affecting the apparent partition coefficient of the compound. The effect of the dissociated compound approximatively can be taken into account by applying a dissociation correction to the partition coefficient. Additional input for this suboption can be given in the COMPUTE DISSOCIATION CORRECTION (LOGD) subsection of the logP panel. For the
calculation of the dissociation correction either input of the solute pKa value or a cosmo file of the
dissociated solute is required.

Example: Prediction of octanol / water partition coefficients

Select your compounds, water, 1-octanol and aspirin from the File Manager. Select the Compute
Octanol-Water Partition Option:

COSMOtherm automatically chooses the correct phase compositions. As you can see the Wet
Octanol phase contains 0.274 mole fractions of water. Finally, add your settings to the property
panel and run the program. The partition coefficients can be read from the output and table files.

4.12 Calculation of pKa

The pK\textsubscript{A} of a solute j can be estimated from the linear free energy relationship: (LFER),

\[ pK_A = c_0 + c_1 (\Delta G_{\text{neutral}}^j - \Delta G_{\text{ion}}^j) \]  \hspace{1cm} (9)

where \( \Delta G^j \) are the free energies of the neutral and the ionic compounds.

The pK\textsubscript{A} option allows for the computation of the pK\textsubscript{A} value of a compound in a solvent i (usually
water). The free energies \( \Delta G^i \) in the solvent at infinite dilution are computed and the pK\textsubscript{A} is
estimated from the above LFER. Thus, to obtain a pK\textsubscript{A} value it is necessary to do quantum chemical
COSMO calculations of a molecule in its neutral and in its ionic state. Since the LFER is valid for both
anions and cations it is possible to estimate acidity as well as basicity. The LFER parameters \( c_0 \) and \( c_1 \)
are read from the COSMOtherm parameter file by default.

pK\textsubscript{A} prediction by COSMOtherm is not restricted to aqueous acid pK\textsubscript{A}. However, both aqueous base
pK\textsubscript{A} prediction and pK\textsubscript{A} in solvents other than water require reparameterization of the pK\textsubscript{A} LFER
parameters. LFER parameters for aqueous base pK\textsubscript{AV}, pK\textsubscript{A} in solvents dimethylsulfoxide (DMSO) and
acetonitrile at room temperature are shipped within the COSMOtherm parameter files
BP_TZVP_C30_1201.ctd and BP_SVP_AM1_C30_1201.ctd. The parameterizations will be used by
COSMOtherm if the corresponding options are selected from the pK\textsubscript{A} card. Please note, that the
solvent has to be set corresponding to the selected option for the LFER parameters. LFER
parameters for solvent-solute systems other than those provided by COSMOtherm or for
temperatures other than room temperature can be set by selecting the ADVANCED SETTINGS checkbox to enter the LFER parameters. For the computation of higher states of ionization, the neutral and singly charged ionic species have to be replaced by higher ionized species.

Example: Calculation of the aqueous pK$_A$ of pyrrolidine

Select the solvent (water), and the neutral and ionic compounds (pyrrolidine, pyrrolidine-cation), from the TZVP DB. Set the temperature, 25 °C, and set water as the solvent. Set the neutral and ionic compounds from the menus. Use the pK$_A$ parameters for bases in water.

Note that it is possible to reset the compounds and also add them to the input. In that case, COSMOtherm will do more than one property calculation and write the results to the output and table files. Since we have chosen room temperature and water as solvent for the calculation, no further settings are necessary. If you want to use your own LFER parameters, input is possible via the ADVANCED SETTINGS option. Save the input file and run the calculation. The COSMOtherm output and table files will open after the calculation has finished.

The table file lists the computed pK$_A$ (experimental pK$_A$ for Pyrrolidine = 10.27)

4.13 Vapor liquid-equilibria (VLE) and liquid-liquid equilibria (LLE)

COSMOtherm allows for the computation of phase diagrams (VLE and LLE) of binary, ternary or higher dimensional (“multinary”) mixtures. It is possible to calculate phase diagrams at fixed pressure (isobaric) or at fixed temperature (isothermal). The pressure or temperature has to be given in the input. The program automatically computes a list of concentrations covering the whole range of mole fractions of the binary, ternary or multinary mixture. At each point the following properties are calculated:

- the excess properties $H^e$ and $G^e$,
- the chemical potentials $\mu_i + RT \ln(x_i)$,
- the activity coefficient $\gamma_i$,
- the total vapor pressure of the system $p^{(tot)}$,
- and the concentrations of the compounds in the gas phase $y_i$. 
The total pressures used in the computation of a phase diagram are obtained from

\[ p^{(tot)}_i = \sum x_i p^0_i y_i \]  

(10)

The \( p^0_i \) are the pure compound vapor pressures for compounds \( i \). \( x_i \) are the mole fractions of the compounds in the liquid phase and \( y_i \) are the activity coefficients of the compounds as predicted by COSMOtherm. Ideal behavior in the gas phase is assumed. Thus, the computation of phase diagrams requires the knowledge of the pure compound’s vapor pressure \( p^0_i \) at a given temperature. There are several possibilities to calculate or approximate this property, as described in the Henry Law Coefficient section. By default, the COSMOtherm approximation of the vapor pressure, using the approximated gas phase energy of the compound, is employed, unless the use of energy files or vapor pressure / property files is specified in the GLOBAL OPTIONS. For other options, experimental data can be entered in the COMPOUND PROPERTIES dialog from the compound list context menu.

Vapor mole fractions \( y_i \) are obtained from the ratio of total vapor pressures and partial vapor pressures \( p_i = p^0_i x_i y_i \) :

\[ y_i = \frac{p^0_i x_i y_i}{p^{(tot)}} \]  

(11)

Phase diagrams can be calculated either at a fixed given temperature or at a fixed given pressure with variable temperatures. In an isobaric calculation, COSMOtherm will compute the mixture properties and vapor pressure for each concentration at different starting temperatures and iteratively converge to the temperature corresponding to the given pressure.

For binary phase diagram calculations, options for automatic post-processing are provided, such as fitting of the computed activity coefficients to activity coefficient models or acetropes detection. COSMOtherm also offers the possibility to compute phase diagrams of binary or ternary mixtures of an Ionic Liquid (IL) phase with additional solvent phases. In ternary calculations, up to 3 ionic liquid phases can be defined. Because in COSMO-RS theory any Ionic Liquid or dissolved salt phase has to be treated by means of the individual ions forming the IL or salt, a phase diagram computation has to be conducted in the form of a multicomponent phase diagram with the boundary condition of the anion and cation concentrations forming the IL or salt having a fixed ratio according to the IL / salt stochiometry. This means that within COSMOtherm, the IL / salt is treated by means of the individual ions, but on output, the results of the individual ion’s properties are combined to form a single IL or salt phase, i.e. in the output of COSMOtherm the “laboratory binary” definition is used for binary calculations involving Ionic Liquids. For information about the conversion from the “pseudo-binary” framework to the “laboratory-binary” framework refer to sections 2.3.7.3 and 5.9 of the COSMOtherm User’s Manual. Note that isobaric phase diagram calculations including an Ionic Liquid phase are unfeasible in COSMOtherm.

Miscibility gaps in binary, ternary and multinary mixtures can be searched for automatically. The liquid-liquid equilibrium properties are calculated from

\[ x^I_i y^I_i = x^II_i y^II_i \]  

(12)

where superscripts \( I \) and \( II \) denote the two liquid phases. If the SEARCH LLE POINT option is used, the COSMOtherm table file will be modified according to any miscibility gap that has been detected. In case of a binary mixture, the binodal LLE point (eq. 14) and the spinodal LLE point, that distinguishes the unstable region of the liquid mixture from the metastable region, will also be
printed in the table file. In ternary and multinary mixtures it is necessary to define a starting composition and end composition. A given number of points between these compositions is then used a starting point for the tie point search. If no good guess for the starting and end compositions is available, several concentrations should be tried. Please be aware that no guarantee can be given that all points are found.

Example: Calculate the solubility of 2-chlorophenol in hexane

First, select the compounds, 2-chlorophenol and water (h2o.cosmo) from the TZVP database. Check the ACTIVATE CONFORMER TREATMENT option. Set the condition of the LLE to ISOTHERMAL, set the temperature (25 °C), the type of the system (BINARY) and the components. Check SEARCH LLE POINT. Create the property input with ADD.

When the calculation has finished the COSMOTHERM output and table files will open in separate windows. The data of the binary phase diagram are tabulated in the .tab file, and can also be plotted in a plot tool. Go to the “graphics” card in the table file window. Choose a quantity from the left menu and plot it. Use the shift or Control keys to select another quantity for the same plot. A right mouse-button click in the plot opens a menu which allows you to add properties of the same or other table files, e.g. to compare VLEs at different temperatures, or to change the quantity on the x-axis.

4.14 Solid liquid equilibria

With the SLE option, COSMOTHERM will compute a range of mixtures and search for possible concentrations of solidification. The solid-liquid equilibrium properties are calculated from

\[
\mu_i^{Solid} = \mu_i^{Liquid} + RT \ln(x_i)
\]
The SLE search assumes that there is a simple eutectic point in the binary mixture. Complicated systems with several phase transitions in the solid state cannot be predicted by the SLE option. Since COSMO-RS is a theory for liquids only, the Gibbs free energy of fusion of the compound, $\Delta G_{\text{fus}}$, has to be taken into account for the solid-liquid equilibrium of a solid compound with a solvent. COSMOtherm will calculate a temperature dependent free energy of fusion if the compounds enthalpy or entropy of fusion ($\Delta H_{\text{fus}}$ or $\Delta S_{\text{fus}}$, respectively) and melting temperature ($T_{\text{melt}}$) are known. These data can be read from the .vap file or you can use the COMPUND PROPERTIES dialog from the compound context menu to enter the data. Note that the value of a salt’s free energy of fusion $\Delta G_{\text{fus AC}}$ (as computed from salt enthalpy or entropy of fusion and melting point) by default is considered to be the “total” or “net” free energy of fusion of the salt. This means that the given salt free energy of fusion value is scaled by the total stoichiometry of the salt in a salt solubility or salt SLE computation. Furthermore, the heat capacity of fusion can be used to improve the calculated temperature dependency.

Like in a VLE/LLE binary calculation, it is possible to compute solid-liquid phase equilibria (SLE) for pseudo-binary solutions with an Ionic Liquid or salt phase. The prediction mole fraction concentration of the ionic liquid or salt is converted to the "laboratory-binary" framework (i.e. the salt is considered to be one single compound) when printed to the table file. The thermodynamic properties are calculated at 325 mixture concentrations distributed on an even spaced grid, followed by an iterative refinement. Note that therefore, the calculation will take some time.

Example: Solid-liquid equilibrium curve of toluene and ethylbenzene

Select toluene and ethylbenzene from TZVP DB. For both compounds the necessary data for $\Delta G_{\text{fus}}$ are already contained in the database. Toluene and ethylbenzene are automatically used as compounds. To choose a temperature series check the COMPUTE SLE AT SEVERAL TEMPERATURES AND ENTER 140 K as starting temperature and 178 as end temperature. Use as many points as you like (beware that too many points will need a lot of calculation time).

The computed SLE points are stored in the .tab file for each calculated temperature. By selecting the GRAPH BIN tab, selecting all temperatures and selecting SLEx(1) and SLEx(2) a simple eutectic point can be revealed. The individual SLE points for each temperature are also printed in each card of the .tab file.
4.15 FlatSurf: surface activity

With the FLATSurf option, the surface interaction energy of all compounds is computed at the interface of the two solvents or solvent mixtures. This is possible under the idealized assumption of a flat interface. The position of the solute at the interface is described by the distance $z$ of the solute center from the interface, and orientation $\gamma$ of a fixed solute axis with respect to the surface normal direction. For such a given position of the compound a certain part of the molecular surface segments will be imbedded in phase $S$ and the rest in phase $S'$. By sampling all relevant positions and orientations the minimum of the free energy of the solute at the flat interface of $S$ and $S'$ can be found. The search for the optimal association of $X$ at the interface can be extended to conformationally flexible molecules when the free energy differences between different solute conformers are taken into account. The minimum of the free energy of the solute at the flat interface of $S$ and $S'$ and the total free energy of the solute at the flat interface of $S$ and $S'$ can both be used as significant and thermodynamically rooted...
descriptors for the determination of surface activity in a solution. More details about the method can be found in the COSMOtherm Users Manual, section 5.10.

COSMOtherm can use the experimental interfacial tension of the two solvent phases to improve the computed FlatSurf energies. This is possible with the IFT=value keyword. The value of the interfacial tension is expected to be in [dyne/cm]. Values for interfacial tensions of various solvent-solvent or air-solvent combinations can be found e.g. in the CRC Handbook of Chemistry and Physics.¹⁴

To visualize the immersion and geometric partition of a solute in the two phases the option “Create Flatsurf VRML charge surface” can be checked. With this option, a VRML file will be written where the immersion depth $z$ of the solute between the two solvent phases is represented graphically on the charge surface in the form of a black and white ring. The black part of the ring points towards FlatSurf solvent phase 1 and the smaller white part of the ring point towards FlatSurf solvent phase 2. Thus the ring indicates how the solute molecule is immersed in the two phases. The example on the right shows the immersion of a phenol molecule in a water (upper part) and a hexane phase (lower part).

**Example: Calculate the air-water surface partition energy**

Select the compounds water, benzene and chlorobenzene from the TZVP database. For air, select the vacuum.cosmo file. In the FLATSURF card, check PURE for vacuum in phase 1 and for water in phase 2. Enter the value for the interfacial tension at the air-water interface (72.8 dyne/cm) into the appropriate field. To visualize the immersion of the solute between the two phases, check CREATE FLATSURF VRML SIGMA SURFACE. Add the settings to the property selection window and run the program.

For each compound, the following descriptors are written to the output and table files:

- $\mu_{X,SS',res}(\text{Gmin})$: maximum of the free energy gain of the solute $X$ going from phase 1 to the interface of $S$ and $S'$. (Please note, that going from phase 1 to the interface makes a difference to going from phase 2 to the interface)
- $G_{X,SS'}(\text{Gtot})$: total free energy gain of the solute $X$ going from phase 1 to the interface of $S$ and $S'$.
- $a_{X,SS'}(\text{Min})$: contact area of the solute $X$ with phase 2 at the free energy minimum.
- $a_{X}(\text{A})$: total area of the COSMO-surface of solute $X$.
- $z(\text{Depth})$: distance of the center of solute $X$ from the interface at the free energy minimum.
• $k (\kappa)$: number of orientations that were used to determine the surface interaction energy minimum of solute $X$.

If several conformers were used to compute a compound’s surface interaction energy, \textit{COSMOtherm} will always write the name of the specific conformer to the table output, which was able to achieve the lowest value of $\mu_{ssmin}^X (\text{Gmin})$. I.e. From the list of all conformers of a compound the one with the lowest minimum free energy values at the flat interface of $S$ and $S'$ will be listed. In contrast, $G_{ss}^X (\text{Gtot})$, the total free energy gain of the solute $X$ at the flat interface is the thermodynamic average according to the interface partition sum of all conformers.

### 4.16 Density

The Density option uses the corrected molar liquid volume $\tilde{V}_i$ of the pure compounds to calculate the pure compound liquid density $\rho$ for all given compounds according to

$$\rho_i = \frac{MW_i}{\tilde{V}_i N_A}$$

where $MW_i$ is the molecular weight of the compound and $N_A$ is Avogadro’s constant. The corrected molar liquid volume $\tilde{V}_i$ is computed from a Quantitative-Structure-Property-Relationship (QSPR) which includes seven generic QSPR parameters and one element specific parameter. Generally, the QSPR model is valid for a specific temperature only, because it does not include a temperature dependency term. The QSPR parameters and element specific surface area parameters that can be read from the \textit{COSMOtherm} parameterization files \texttt{BP_SVP_AM1_C30_1201.ctd} and \texttt{BP_TZVP_C30_1201.ctd} were derived from room temperature data. Optimized element specific surface area parameters are available for elements H, N, C, O, F, S, Cl, Br and I. For other elements reasonable guesses for the element specific surface area parameters are provided. It is possible to specify user-defined values for the density/volume QSPR parameters and the element specific surface area parameters. Check the “Advanced Settings” checkbox to enter the parameters manually. Please refer to the \textit{COSMOtherm} Users Manual, section 2.3.12, for more details.

If several conformers of one compound are present, \textit{COSMOtherm} will compute the density descriptors of all individual conformers. Subsequently a thermodynamic average of the conformer descriptors at the given temperature condition will be calculated, from which the averaged density of the compound is predicted.

The computation of the liquid density of a pure Ionic Liquid (IL) compound can be done from the individual ions that form the specific IL. The composition of the IL has to be defined in the input. The IL is then pieced together from its anion(s) and cation(s). If the IL option is used, \textit{COSMOtherm} will compute the density of the given IL compound only (this is unlike the regular density option for individual compounds, which always computes the densities of all compounds present).

#### Example: Liquid density and liquid molar volume of pure compounds

Select the compounds from the TZVP Database: h2o, 1-butanol, butanone, diethylether, chcl3, chlorobenzene, benzene, toluene, and octane. Enter the temperature in the DENSITY card (25 °C) and transfer the settings to the “Property Selection” window with the ADD button. Run the program. The calculated densities and volumes can be read from the table file.
4.17 Viscosity

The pure compound liquid viscosity is another property that can be calculated from QSPR. The descriptors for the liquid viscosity are the compound surface area as read from the COSMO file $A_i$, the second $\sigma$-moment of the compound $M_i^2$, the number of ring atoms in the compound $N_i^{\text{Ring}}$ and the pure compound’s entropy times temperature $T_S$, which is computed from the difference of the total enthalpy of mixture of the pure compound $H_i$ and the chemical potential of the pure compound $\mu_i$: $T_S = -(H_i - \mu_i)$. This QSPR model, like the Density QSPR model, does not include a temperature dependency term, so that the model is valid at a specific temperature only. Currently the parameterization $\text{BP}_S\text{VP}_{\text{AM1}_C30_1201}.\text{ctd}$ and $\text{BP}_{\text{TZVP}}_{\text{C30_1201}}.\text{ctd}$ include the viscosity QSPR parameters for room temperature. User-defined values for the QSPR parameters can be specified manually within the ADVANCED SETTINGS.

For a compound with several conformers COSMOtherm will compute the viscosity descriptors of all individual conformers, followed by a thermodynamic averaging of the conformer descriptors at the given temperature condition to predict the averaged viscosity of the compound.

4.18 The $\sigma$-moment approach and QSPR calculations

The $\sigma$-potentials of liquids can be represented by a Taylor-series with respect to $\sigma$,

$$\mu_S^X = \sum_{\sigma_i} c_{S_i} M_i^X$$

where the coefficients $c_{S_i}$ describe the specific corrections required for matrix $S^{15}$. The $\sigma$-moments $M_i^X$ can be used to compute certain molecular properties via a Quantitative Structure Property Relationship (QSPR) approach$^{16,17}$. The COSMOtherm $\sigma$-moments can be correlated with properties such as lipophilicity, biological or environmental partition behavior like the octanol-water or soil-water partition, or the partition of a compound between the blood-brain barrier. The QSPR coefficients $c_{S_i}$ for a certain property can be determined from a multi-linear regression of the $\sigma$-moments with a sufficient number of experimental data. For a compound $X$ a property $\log(P)$ is calculated via:

$$\log(P) = c_0 M_0^X + c_1 M_1^X + c_2 M_2^X + c_3 M_3^X + c_4 M_4^X + c_5 M_5^X + c_6 M_6^X + c_7 M_{\text{HBdon}}^X + c_8 M_{\text{HBacc}}^X + c_9 M_{\text{HBdon2}}^X + c_{10} M_{\text{HBacc2}}^X + c_{11} M_{\text{HBdon3}}^X + c_{12} M_{\text{HBacc3}}^X + c_{13} M_{\text{HBdon4}}^X + c_{14} M_{\text{HBacc4}}^X$$

$$+ c_{15} M_{\text{HBdon}}^X + c_{16} M_{\text{HBacc}}^X$$

(13)
where $M_i^X$ is the $i^{th}$ $\sigma$-moment of compound $X$ and $M_{HBac} i^X$ and $M_{HBdon} i^X$ are the $i^{th}$ hydrogen bonding acceptor and donor moments of compound $X$. Thus, a maximum of 16 coefficients is available to do the $\sigma$-moments QSPR calculation. However, a multilinear regression can usually be done with only 5 descriptors ($M_0$ (area), $M_2$ (sig2), $M_3$ (sig3), $M_{HBac3}$, $M_{HBdon3}$) to avoid over-parameterization.

For a detailed description of $\sigma$-moments and property calculation via $\sigma$-moment QSPR refer to sections 5.4 and 5.5 of the COSMOTHERM User’s Manual.

The COSMOTHERM release includes QSPR coefficient files for the following properties, parameterized on the Turbomole BP-SVP-AM1 COSMO level:

- Blood-Brain partitioning.
- Intestinal absorption coefficients.
- Plasma-protein (human serum albumin) partition coefficients.
- Soil-Water partition coefficient.
- Octanol-Water partitioning.

Furthermore, the COSMOTHERM release includes a number of QSPR property files holding QSPR coefficients for the five Abraham parameters and the definition of thermodynamic partition properties via the six Abraham coefficients, for both computational COSMO levels BP-TZVP and BP-SVP-AM1.

For an automatic QSPR calculation of the selected compounds COSMOTHERM offers two options:

**Global QSPR option:** This option is activated when the required property in the subsection QSPR PROPERTY CALCULATION of the GLOBAL OPTIONS menu is checked. By default, the computed property value will be listed in the compound section of the COSMOTHERM output file. Additionally, a tabulated output file with the extension .mom will be written, listing the molecular $\sigma$-moments and, in the last column, the computed property. Currently, when this option is used, there can always one QSPR property only be calculated in a single run.

**Mix QSPR option:** This option can be selected from the Mix QSPR panel. It is closely related to the global QSPR option, but the Mix QSPR option writes the results to the mixture section of the COSMOTHERM output file as well as to the COSMOTHERM table file, but not to the molecules $\sigma$-moment files (.mom). If mixture composition and temperature are not specified (which is the default), COSMOTHERM calculates the QSPR property chosen from the menu for all molecules, i.e. for all conformers of the compounds individually, giving the same results as the global QSPR option. However, with the Mix QSPR option, temperature and mixture composition can be specified.
when the ADVANCED SETTINGS checkbox is checked. If this is done, the QSPR property will be calculated for all compounds by averaging the property according to the Boltzmann distribution of the conformers at the given temperature and mixture concentration. This will result in different values for the QSPR property only for compounds for which more than one conformer is present.

4.19 Similarity

With this option, COSMOtherm will calculate a molecular similarity of two compounds based on \( \sigma \)-profiles or \( \sigma \)-potentials. There are three possible methods which can be used for a similarity calculation, and which can be combined in a single calculation:

The \( \sigma \)-profile similarity factor \( S_{i,j} \) is calculated as the normalized overlap integral of the \( \sigma \)-profiles \( \rho_i(\sigma) \) and \( \rho_j(\sigma) \) of the two compounds \( i \) and \( j \):

\[
S_{i,j} = \left( \int_{-\infty}^{+\infty} \rho_i(\sigma) \rho_j(\sigma) d\sigma \right) / A_i A_j \tag{14}
\]

\( S_{i,j} \) will be small if the overlap between the compounds \( \sigma \)-profiles is small. In addition, the similarity factor given by eq. (14) is corrected by a factor \( S'_{i,j} \) taking into account the difference in the apparent hydrogen bonding donor and acceptor capacities of the two compounds and by a factor \( S''_{i,j} \) taking into account size differences between the two compounds \( i \) and \( j \).

COSMOtherm can also calculate a \( \sigma \)-potential based similarity factor for two compounds. This method is useful for comparing solvents. The COSMOtherm \( \sigma \)-potential similarity factor \( S''_{i,j} \) is defined as the sum of the differences between the two pure compound \( \sigma \)-potentials \( \mu_i(\sigma) \) and \( \mu_j(\sigma) \):

\[
S''_{i,j} = \exp \left( - \sum_{m=-0.02}^{m=0.02} \left[ \mu_i(\sigma_m) - \mu_j(\sigma_m) \right] \right) \tag{15}
\]

\( S''_{i,j} \) will be small if the overlap between the compounds \( \sigma \)-potentials is small. The \( \sigma \)-potential can also be weighted by the \( \sigma \)-profile of a third compound. Then the computed similarity is a solute-specific \( \sigma \)-potential similarity (see section 2.3.10 of the COSMOtherm User’s Manual).

Alternatively, COSMOtherm allows for the calculation of a molecular \( \sigma \)-profile similarity by the “Sigma-Match Similarity” (SMS) algorithm\(^{18} \).

The similarity factors \( S_{i,j} \), \( S'_{i,j} \) and \( S''_{i,j} \) are printed to the mixture output section of the COSMOtherm output file below the compound output block of the first compound for which the similarity factors are calculated and, additionally, to the COSMOtherm table file. If several conformers are present for a compound, the \( \sigma \)-profile based similarity factor will be computed for all possible combinations of the conformers and the overall compound similarity factor is averaged from the computed conformer similarity factors. It is possible to enter a solvent mixture, in which the different conformers are weighted. In benzene a different conformer might be relevant than in water, consequently the SMS can give different results if the conformers are weighted according to their abundance in water or benzene.

4.20 Liquid extraction

This option allows for the automatic computation of a multi-component multi-phase liquid-liquid extraction equilibrium. The mole or mass based equilibrium partitioning of an arbitrary number of compounds between a given number of liquid or, optionally, solid (precipitation) and gaseous (evaporation) phases, can be computed. The functionality of the option is described in the figure below for two liquid phases:
The two predefined phases I and II may be mixtures of compounds including ionic liquids or dissolved salts. Phases I and II are assumed to be immiscible and to separate in thermodynamic equilibrium. In the above example, phase I has a high water concentration and phase II is an ionic liquid. Compounds methanol, acetone, H2O and the ions are distributed between the two phases. Please note, that all compounds including water and the ionic liquid are equilibrated.

All compounds are distributed among the two phases according to their partition equilibrium constants:

\[ K_i^X = \exp \left( \frac{\mu_i^I - \mu_i^II}{RT} \right) \]

Since all chemical potentials themselves do depend on the composition of the phases, the final equilibrium distribution has to be solved iteratively. Starting from the given concentrations a set of new concentrations is calculated. In the next step the new concentrations are used to calculate the refined chemical potentials and thus another refined set of concentrations. This procedure is repeated until the concentrations of the two phases do not change any more. Thus the thermodynamic equilibrium, the mass balance and, if ions are present, the charge neutrality condition (as a boundary condition of the mass balance) of the two phases are solved simultaneously in an iterative self consistent manner until the system converges to a thermodynamic and mass equilibrium of two neutral phases.

Unlike almost all of the other COSMOTHERM computations, the liquid extraction can be done in the framework of the absolute amounts of substance, namely compound mole numbers (N) or compound masses (W). These options can be chosen alternatively to the common options of mole fraction and mass fraction.

In addition to the two starting phases in the Liquid Extraction panel, it is also possible define further liquid phases as solvent phases in the phase definition block. Furthermore, when absolute mole numbers or masses are used to define the liquid phases, one or several solutes can be added to the predefined phases from an external “solute” reservoir (ADVANCED SETTINGS). The additional “solute” is considered completely dissolved in the two phases I and II. This option may be helpful if it is not clear in advance, which phase will be preferred by a given compound when added to the system.

By default all phases are assumed to be liquid phases. It is, however, possible to define one of the phases defined in the phase definition block as a solid (precipitation) or gaseous (evaporation) phases. Solid phases can also be defined as salts, composed of cosmo files of the corresponding ions. All salts will precipitate into the phase that is defined as solid unless they are forced to precipitate into their own separate salt phases.

In the COSMOTHERM output file the equilibrium constants k, and the equilibrium concentration x, are printed. The table file gives the final absolute compositions of the phases in mole (N), gram (W) and the concentrations as mole fraction (x). If mass or mole fractions have been used as input, absolute mole numbers and masses are excluded from the output.
Example: Three-phase equilibrium between water, hexane, and [bmim+][pf6-] with an additional solute

Select the compounds from the TZVP Database: h2o, hexane, propanone, 1-butyl-3-methyl-imidazolium(+) and pf6(-). Enter the temperature (25 °C) and the phase compositions: Set the mole number for phase 1 to 35 mole of water, for phase 2 to 50 mole of hexane and for phase 3 to 15 mole of water, 25 mole of the cation 1-butyl-3-methyl-imidazolium and 25 mole of the anion pf6. Check the ADVANCED SETTINGS checkbox and enter 1 mole of propanone in the ADDITIONAL SOLUTE section. Transfer the settings to the “Property Selection” window with the ADD button and run the program.

4.21 Reaction

The reaction panel basically allows to calculate the equilibrium constant \( K_{\text{reac}} \), the free energy \( \Delta G_{\text{reac}} \) and the enthalpy \( \Delta H_{\text{reac}} \) of a given reaction in an arbitrary solvent.

For a hypothetical reaction \( A + B \rightarrow C + D \) the free energy is defined as:

\[
\Delta G_{\text{reac}} = G_C + G_D - (G_A + G_B)
\]

and the reaction equilibrium constant is defined as:

\[
K_{\text{reac}} = \exp(-\Delta G_{\text{reac}} / RT)
\]

In a standard calculation the values for \( G_A, G_B, G_C, \) and \( G_D \) are computed from the DFT gas phase energies modified by the free energy of solvation \( \Delta G_{\text{solvation}} \) as calculated by COSMOtherm.

\[
G_A = E_{A,\text{DFT}} + \Delta G_{A,\text{solvation}}
\]
In each calculation the solvent and both sides of the reaction have to be specified. If neither products nor educts are part of the solvent / solution, a reaction in infinite dilution is calculated. If products or educts are part of the solvent their concentration is taken into account, but remains constant.

Using the standard approach, a relative ranking of the reaction equilibrium in different solvents can be established from calculating the same reaction in different solvents or solvent mixtures, which makes it possible to determine the solvent suited best for a given reaction. For an absolute estimate of the reaction enthalpy or free energy higher level quantum mechanical energies are necessary. In general, the standard DFT energies are not accurate enough for absolute predictions of a reaction energy.

The ADVANCED SETTINGS dialog allows for entering high precision energies from external quantum mechanical calculations, zero point energies and experimental values for $\Delta G_{\text{solvation}}$ and $H_{\text{vap}}$. Please refer to the manual for a detailed description.

The result table file contains the individual values for the free energy, the enthalpy, the chemical potential in solvent, the free energy of solvation, enthalpy of interaction, enthalpy of vaporization and the external values for quantum mechanical or zero point energy.

Example: Solvent choice for a simple ester reaction

Load the educts and products methanol, formic acid, methylformate and water. Any solvent compound can also be added to the compound list. Enter methanol and formic acid with stoichiometry 1 into the reactant definition and methylformate and water with stoichiometry 1 into the product definition.

By varying the solvent composition, it can now easily be seen that the reaction free energy, and thus the equilibrium constant critically depend upon the choice of the solvent.
4.22 Ionic Liquid Screening

The IONIC LIQUID SCREENING panel enables the calculation of capacity and selectivity of Ionic Liquid solvents. The calculated solvent capacity \( C^\infty \) corresponds to the non iterative (infinite dilution) solubility of a solute in the Ionic Liquid solvent, and is a good first solubility guess for a fast screening procedure. The selectivity, i.e. the fraction of the capacities of two different solutes in the same IL solvent, can be calculated with the SELECTIVITY SCREENING suboption. The IONIC LIQUID SCREENING has been implemented for singly charged ions and the BP-TZVP level of theory. Note that .mcos files for the ions should not be used. Neutral solutes can be represented by .mcos files.

4.23 Ionic Liquid Properties

The IONIC LIQUID PROPERTIES panel allows for the computation of a range of properties of Ionic Liquids using a QSAR approach. The QSARs have been implemented for singly charged ions and the BP-TZVP level of theory, and some depend on specific COSMOtherm parameterizations. Note that for the ions, .mcos files should not be used. The TEMPERATURE DEPENDENT DENSITY and TEMPERATURE DEPENDENT LIQUID ENTROPY options require .energy files that provide gas phase frequency data at the BP-TZVP level. For information on the file format definition refer to the COSMOtherm User’s Manual, section 2.2.1. Missing information will lead to empty tab file entries. The prediction of the CRITICAL MICELLE CONCENTRATION is based on different interaction terms of ions and water. Therefore, water (h2o.cosmo) needs to be loaded to the compound list.

4.24 Environmental / Safety

In contrast to other panels in COSMOthermX, the ENVIRONMENTAL / SAFETY panel enables the calculation of a collection of physico-chemical properties, which are of importance for environmental and safety considerations, for a group of single (pure) compounds. Most of the properties can also be computed separately. For a description of these properties, refer to the corresponding sections in this document. Special to the ENVIRONMENTAL / SAFETY panel are the calculation of the flash point (FP) and the atmospheric lifetime. Note that usage of .mcos files is not recommended.

**Flash Point:**

Recently, we have turned prediction of the FP into prediction of a flash point pressure (FPP). The FPP appears to be a simple function of the surface area. Hence essentially FP prediction is predication of the boiling point at the predicted FPP. The accuracy is \( \pm 15 \) K. If an experimental BP point is provided in the respective sub-menu, a correction for the error in the vapor pressure prediction is applied, and the accuracy increases to \( \pm 10 \) K.\(^{19} \) Note that this option has to be considered as preliminary since it will be re-worked next year based on a new compilation of experimental data.

**Atmospheric life time:**

Since the atmospheric life time of organic compounds is mainly determined by the reaction with OH-radicals, it can be estimated based on the OH-radical reactivity \( k_{\text{OH}} \) of the compound. A molecular orbital based prediction method MOOH for \( k_{\text{OH}} \) has been published already in 1992, based on semi-empirical AM1 calculations. Recently Schüürmann et al.\(^{20} \) showed that the MOOH method is still up to date and at least as accurate (\( \pm 0.35 \) log-units) as the generally accepted increment method by Atkinson, with the advantage of having smaller maximum deviations, i.e. being more robust. The calculations are done with our MOOH implementation in MOPAC7, using the lowest energy gas-phase conformation provided.\(^{21,22} \)
5 Using your own COSMO files

There are several ways to make your own COSMO files available in COSMOthermX. You can

- select COSMO or compressed COSMO files from any directory on your system using the FILE MANAGER.
- add your own database(s).
- extend existing databases. (Note that extension of databases delivered by COSMOlogic is disabled.)

For all options there are a few prerequisites you have to take into account to ensure that the COSMOtherm calculations run correctly:

Ensure that all COSMO files you want to use come from the same quantum chemical level:

- TZVP: COSMO files from BP-TZVP COSMO or gas phase calculations should be combined with COSMO files from the BP-TZVP database only.
- SVP: COSMO files from BP-SVP-AM1 COSMO or gas phase calculations should be combined with COSMO files from the BP-SVP-AM1 database only.
- DMOL3: COSMO files from DMOL3-PBE COSMO or gas phase calculations should be combined with COSMO files from the DMOL3-PBE database only.
- TZVPD-FINE: COSMO files from BP-TZVPD-FINE calculations should be combined with COSMO files from the BP-TZVPD-FINE database only.

Note that .cosmo and .ccf files can be mixed in the databases as long as they come from the same quantum chemical level.

Gas phase energies, which should be used if properties involving a gas phase (VP, VLE, Henry law constant, gas solubility) are calculated and experimental vapor pressure data are not available, should be saved into an .energy file. The gas phase energies must be calculated at the quantum chemical level that has been used for the COSMO calculations, e.g. BP-TZVP for gas phase calculations and BP-TZVP-COSMO for COSMO calculations. Gas phase energy calculations for combination with BP-TZVPD-FINE COSMO files should be done on the BP-TZVPD level of theory. The default unit for gas phase energies is [Hartree].

Experimental vapor pressure data or other experimental data can be saved into a .vap file. Create the .vap file manually using any text editor, or use the "Compound Properties" menu from the right mouse button context menu for a selected compound to enter property data for a compound and save the .vap file permanently. Vapor Pressure / Property and energy files should be located in the same directory as the .cosmo files.

5.1 Using the file manager

COSMO or compressed COSMO files from quantum chemical calculations can be selected from any directory on your system using the FILE MANAGER in the compound section of COSMOthermX. Make sure that you select the appropriate parameterization for the .ccf/.cosmo files.

5.2 Adding your own databases

You can add your own database(s) to COSMOthermX in the EXTRAS/DATABASE SETTINGS dialog. When you select Add, you will be presented with a dialog where you have to enter a name for the database, point to the location of the database directory and the database index file, and specify a parameterization to be used with this database. If an index file with the name of the database
directory does not exist, COSMOthermX will create a database index file. For large databases, this may take some time.

Databases with BP-SVP-AM1 parameterization will be available with the SVP button in the compound section, databases with BP-TZVP parameterization will be available with the TZVP button, databases with BP-TZVPD-FINE parameterization will be available with the TZVPD-FINE button, and databases with DMOL3-PBE parameterization will be available with the DMOL3 button.

Note that, before you add your database to COSMOthermX, you have to make some preparations:

Collect all .cosmo / .ccf files that belong to the database in a single directory. If there are .energy or .vap files for the compounds, also put them into this directory. Ensure that all .cosmo / .ccf files are calculated on the same quantum chemical level, and that the .energy files come from corresponding gas phase calculations.

For conformers of a compound to be identified as conformers, the .cosmo or .ccf files have to be named with the compound name followed by a digit, e.g. ethanol0.ccf, ethanol1.ccf and so on. Then these .cosmo files will be treated as conformers of a single compound if the ACTIVATE CONFORMERS TREATMENT box is checked in the compound section. Note that COSMOtherm can treat up to 10 conformers automatically, i.e. with digits 0-9 in the cosmo file names.

5.3 Extend the existing databases

To extend databases with .ccf or .cosmo files, open your directory tree with the FILE MANAGER. Go to the directory where the .ccf / .cosmo files (and, eventually, .energy and .vap files) are located. Select the compounds you want to add to a database and click ADD TO DB. In the following dialog, select the appropriate parameterization for the .ccf / .cosmo files. Then, select the database to which the compounds shall be added. Take care to add only files to existing databases which match the quantum chemical level of the database. Also, note that database extension is disabled for databases that come with the COSMOtherm release and for databases that are part of COSMObase.
6 Atom weighting

Due to the additive nature of $\sigma$-profiles it is possible to approximately describe molecules which are not contained in a database or which are simply too large to be calculated by adjusting the ‘importance’ of other molecules’ atoms or even blending together multiple secondary molecules. The result of this process is either stored in a so-called weight string (for a single molecule) or in a COSMO metafile with the extension .mcos (for a molecule blended together from multiple molecules). For more information about atom weighting please refer to the COSMOtherm manual.

As the manual generation of weight strings usually involves quite a lot of counting and double checking, COSMOthermX includes COSMOweight, an interactive graphical tool, which can greatly improve your working speed when dealing with medium-sized molecules.

COSMOweight displays a ball-and-stick-model of the selected compound, where disabled atoms will be displayed in black, whereas atoms with higher importance than 1 will be labeled with their weight factor. For a brief description of navigation in 3D-space have a look at the section “Using COSMOview”.

General functionalities of COSMOweight are:

- Instead of manually assigning each unwanted atom a weight of 0, just click the “Set 0” button.
- Resetting all weights to their initial value of 1 is possible by pressing the “Set to 1” button.
- Select an atom with a mouse click and type the number you would like to assign to its weight on your keyboard.

COSMOweight can be operated in a navigation mode and a selection mode. Switch between the modes with the button. The following functionalities are available in navigation mode:

- Atom weights are increased by left-clicking and decreased by right-clicking single atoms.
- Rotate the molecule with the left mouse button.

In selection mode, you can:

- Select a group of atoms with the left mouse button.
- Set correction charges (Delta-charges) in the Delta-charge text field.

Please note that COSMOweight will sometimes display warning messages related to the structure of the molecule being edited. These are only supposed to help spotting chemical and/or logical mistakes - as you will know, the structure itself is not used in any COSMOtherm calculation.

Using COSMOweight to edit weight strings: Right-click a compound in your compound list and select EDIT WEIGHTSTRING. Assign the weights according to your needs, using the options described above.

Using COSMOweight as a meta file editor: This mode can either be accessed by right-clicking an .mcos file in your compound list and selecting EDIT .MCS FILE, by selecting one or more .cosmo or .ccf files from your compound list, right-click and select EDIT .MCS FILE or via TOOLS/COSMOWEIGHT. In the latter case, open an existing .mcos file by clicking OPEN or start from scratch by adding a first molecule. Additionally to editing single weight strings you can add further or remove...
molecules as well as switch between them in the tabbed pane. When you are done, click SAVE (≡) to generate a .mcos file.

**Saving .mcos files:** In the SAVE dialog, there are three different options to save the .mcos file. These options control where COSMOtherm will look for the .cosmo and .ccf files (henceforth “source files”) you used.

**ABSOLUTE:** You can always save your work with absolute paths. The absolute paths of the source files are written to the file, which means, that your sources can be spread over different directories and volumes and you are able to move your .mcos file wherever you please. However, if the source files themselves are moved, or - more importantly - you try to use the .mcos file on another computer with a different folder hierarchy, you will run into problems.

**RELATIVE TO FILE:** This is the default if you do not use a fragment directory. Source files will have to be located in the folder where you save the .mcos file or in subfolders. You can move this kind of .mcos file to different directories, even different computers, if you also copy the source files.

**RELATIVE TO FRAGMENT DIRECTORY:** This option gives you ultimate portability while still avoiding the need to copy data. All source files used have to be located in the folder set as “fragment directory” in the GENERAL SETTINGS dialog. You can move the .mcos file everywhere, as long as you keep with your FRAGMENT DIRECTORY settings. If you use a lot of metafiles, this is the most convenient way of handling the .mcos files. If you did not set a fragment directory, this option will be disabled. Please note that, you may want to set your fragment directory to the COSMO-DATABASE folder or one of its subfolders. The .mcos-files shipped with COSMOthermX will not work unless you either disable the use of a fragment directory or set the database with parameterization as a fragment directory, e.g. ‘<COSMOTHERM_HOME>/DATABASE-COSMO/BP-SVP-AM1>’.

**Problems when using .mcos files:** If COSMOtherm or COSMOweight complain about missing files when using previously saved .mcos files, you are likely trying to open a kind of .mcos file which is saved:

- **RELATIVE TO FRAGMENT DIRECTORY:** Please check if a fragment directory in COSMOthermX is correctly set and enabled.
- **RELATIVE TO FILE:** Make sure that no fragment directory is set in COSMOthermX
- **ABSOLUTE:** Please check the path. If the file was saved on a different computer or the path was changed, the file cannot be found.
Index

activity coefficient 6, 35, 36, 45
  finite concentration 35
AM1 8
Antoine equation 34, 36
atom weighting 61
atomic weight string 14, 22, 61
atomic $\sigma$-moments 15, 29
boiling point 35
BP-SVP-AM1 15, 59, 60
BP-TZVP 59
charge neutrality 18
chemical potential 6, 13, 14, 17, 20, 26, 29, 34
  combinatorial contribution 16, 18
  hydrogen bonding contribution 16, 17
  van der Waals contribution 16, 18
compound list 11
  concentrations 11
compound section 14, 59
remove compounds 20
compound selection 20, 59
Compound Wizard 23
compounds
  context menu 20
  pure compound properties 21, 36, 38, 48, 59
conformers 14, 23, 26
  conformer treatment 16, 18
  conformer weight factor 22
  conformers treatment 20, 22, 26, 36, 60
input 26
manual definition 26
name convention 17, 26, 60
contact probability 29
contact statistics 14, 29
surface contacts 14
COSMO 5
  compressed COSMO files 7, 59
COSMO charge surface visualization 27
COSMO energy 20, 26
COSMO file 7, 20, 59
COSMO metafile 12, 22, 61, 62
databases 13
FlatSurf charge surface 50
COSMO-RS 5
COSMOtherm 7, 11, 62
COSMOthermX 7, 10, 59, 60
COSMOview 13, 27
COSMOweight 13, 22, 61, 62
CTDATA directory 12
databases 13, 20, 23, 27, 59
  add database 59, 60
  conformers 60
  database directory 60
  database index file 59
  database table 23
  extension 59, 60
density 51
energy file 13, 34, 36, 59
excess properties 45
file manager 20, 27, 59
FlatSurf 49, 50
  interfacial tension 50
fragment directory 12, 62
free energy in mixture 29
gas phase energy 12, 13, 21, 34, 59
gas solubility 37
Gibbs free energy
  QSPR 38
heat capacity of fusion 16
Henry law coefficient 36
  finite concentration 36
input file 10, 11, 15
  compound lines 10
  external 11
interaction energy 5, 29
  combinatorial term 6
  electrostatic interaction 5, 29
  hydrogen bond interaction 5, 29
  temperature dependency 16, 17
ring correction 29
temperature dependency 6
van der Waals interaction 5, 6, 29
temperature dependency 16, 17
Liquid extraction 54
liquid-liquid equilibrium
miscibility gap 46
logD 38, 43

mcos-File editor 13
Metafile 61
misfit 5
mixture 29
mixture options 12, 17, 29
molar liquid volume 51
mole fraction 29
molecular geometry 22, 23
gas phase 22

New Molecule 23

Open List 11, 20
options dialog 12, 13, 27, 29, 46
output file 11, 12, 13, 14, 18, 29
additional output files 14, 18
print options 14, 17

parameterization 59
parameter files 12
partial pressure 29
partition coefficient 43
dissociation correction 38, 43
logD 38, 43

phase compositions 29
phase diagram 45
binary 45, 47, 48
isobar 46
liquid-liquid equilibrium 39, 45
LLE point 39
multinary 45
plot tool 47
solid-liquid equilibrium 47
ternary 45
vapor-liquid equilibrium 45

pKₐ 44
acetonitrile 44
acidity 44
aqueous acid pKₐ 44, 45
aqueous base pKₐ 44
basicity 44
DMSO 44
LFER parameters 44
linear free energy relationship 44
solvent-solute systems 44
program control settings 16, 17

pure compound vapor pressure 36, 37, 46
estimate 36, 46
exact value 36, 46
gas phase energy 36, 46
pure compound properties 46
vapor pressure / property file 46

QSPR 15, 52
Abraham parameters 53
density 51
directory 12
global option 53
logBB 15, 53
logKHSA 15, 53
logKIA 15, 53
logKOC 15, 53
logPOW 15, 53
Mix-QSPR 16, 53
molar liquid volume 51
parameterization 15, 53
QSPR-property files 12
viscosity 52
σ-moments 52
quantum chemical levels 7, 8, 12, 20, 59, 60
BP-SVP-AM1 8
BP-TZVP 7

reaction 21
Reaction 56

screening charge density 5
screening charge distribution 27
settings dialog 12, 62

similarity 54
based on σ-potentials 54
based on σ-profiles 54
molecular σ-profile match similarity 54
SMS 54
solute specific 54
solid-liquid equilibrium 47, 48
eutecic point 48
Gibbs free energy of fusion 48

solubility 37, 38, 39
absolute screening for solids 40, 43
Gibbs free energy of fusion 38
iterative refinement 37, 39
liquid 39
mass based 38, 41
relative screening 39, 43
salt solubility 41
salt solubility screening 42  
solid solutes 38  
zeroth order approximation 37  
surface contacts 14  
SVP 8  
table file 11, 12, 29  
print options 14  
temperature 29  
Turbomole 8  
TURBOMOLE 7  
TZVP 7  
TZVPD-FINE 8  

vapour pressure curve 34  
vapor pressure / property file 13, 21, 34, 38, 48, 59  
Antoine coefficients 13  
Wagner coefficients 13  
viscosity 52  
VRML 15, 22, 27  
molecular geometry 15  
surface charges 15  
VRML-Viewer 13  

zero point vibrational energy 21  
\( \sigma \) 5  
\( \sigma \)-moments 14, 16, 29, 52  
\( \sigma \)-potential 6, 15, 22, 23, 27, 29, 52  
\( \sigma \)-profile 5, 15, 22, 23, 27, 29, 61  
\( \sigma \)-profiles 23  
\( \sigma \)-surface 22, 23

---

11. Rational coformer or solvent selection for pharmaceutical cocrystallization or desolvation, Y.A. Abramov, C. Loschen, A. Klamt, . submitted.