FIZ CHEMIE BERLIN
Fachinformationszentrum Chemie GmbH

ARTIST

Structure Editor

Structure Data Base

Automatic Group Assignment

and

Property Estimation from Structure

FIZ CHEMIE Berlin

Franklinstraße 11

D-10587 Berlin

Tel.: +49 (0) 30 39977-118

Fax: +49 (0) 30 39977-132

E-Mail: info@infotherm.de

Web: www.chemistry.de

Contents

Introduction

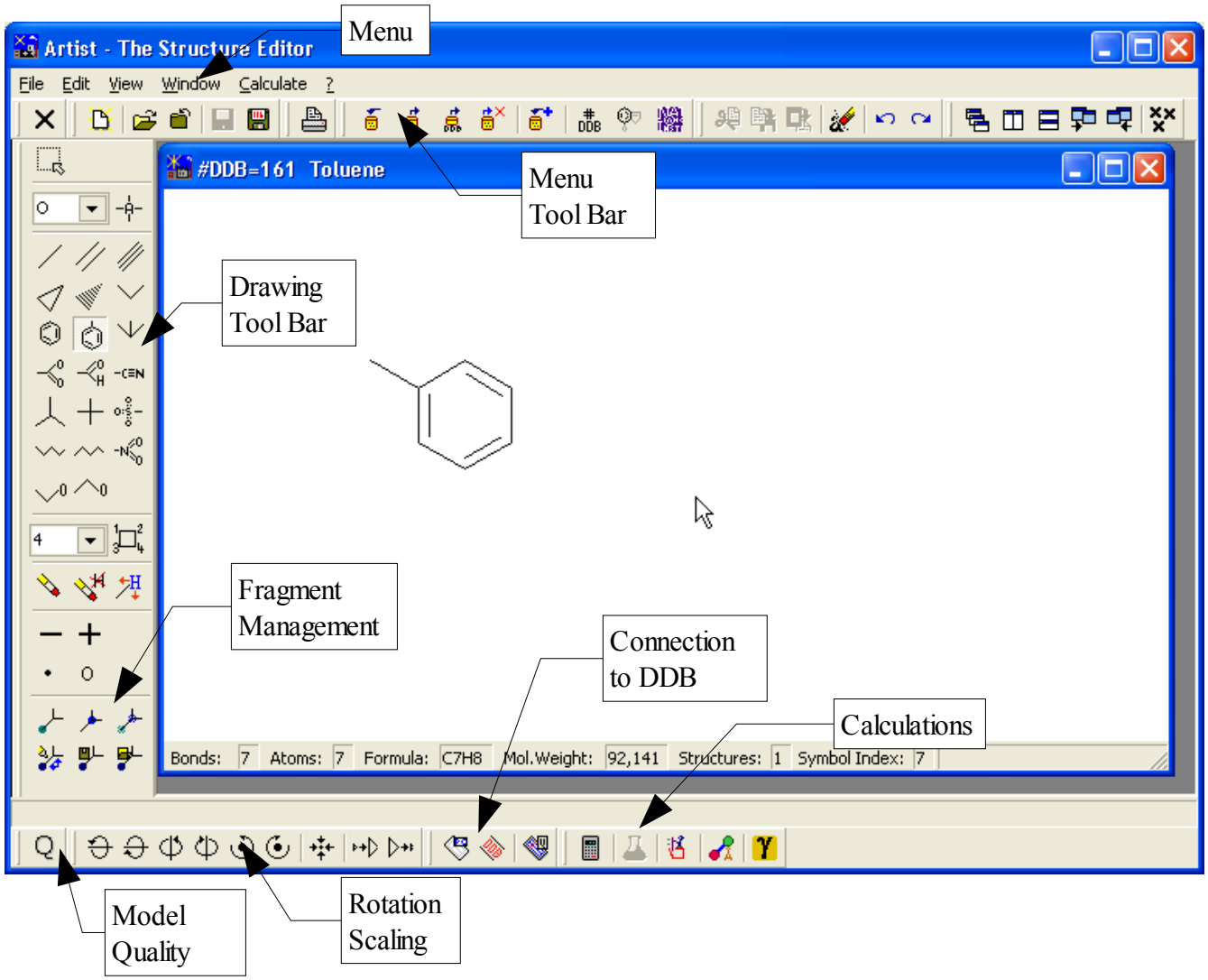
Artist is a tool for editing molecular structures and estimating thermo physical and transport properties from molecular structure directly. Artist contains a structure data bank (ChemDB) currently containing over 19000 components.

The core algorithm of Artist is an automatic fragmentation which allows to implement a wide range of group contribution methods in a quick and reliable manner. A list of the properties and methods is appended to this document.

The list of methods comprise mainly group contribution methods – because of the underlying fragmentation algorithm – but also some corresponding states methods (equation of states etc.).

Artist can additionally be used as retrieval program for components. The Dortmund Data Bank uses a component number index and this number has to be found either by name, formula, CAS registry number and several other techniques. Artist adds the ability to determine the DDB number from the drawn structure.

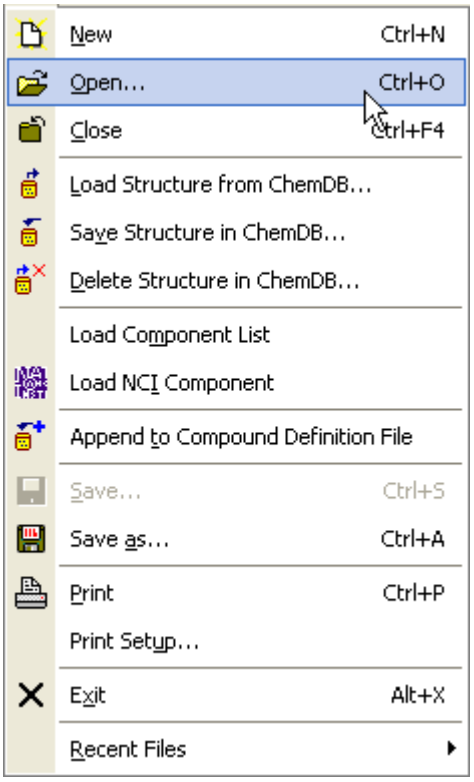
The Main User Interface



Artist is a typical example for a multi-document program contain multiple drawings in a parent window. All functions either in the menus, in the tool bars, or in the context menus are operating on the drawing in the topmost window.

The Menus

The File Menu



New: Open a new and empty drawing window

Open: Opens a molecular structures file. The list of supported file formats is shown in the appendix.

Close: Close current drawing window. If the structure has been modified the program asks if the drawing should be saved.

Load Structure from ChemDB: Read a molecular structure from the structure data bank.

Save Structure in ChemDB: Save a molecular structure in the the structure data bank. If the data bank already contains a structure a dialog will be displayed for either replacing a structure or adding the new structure.

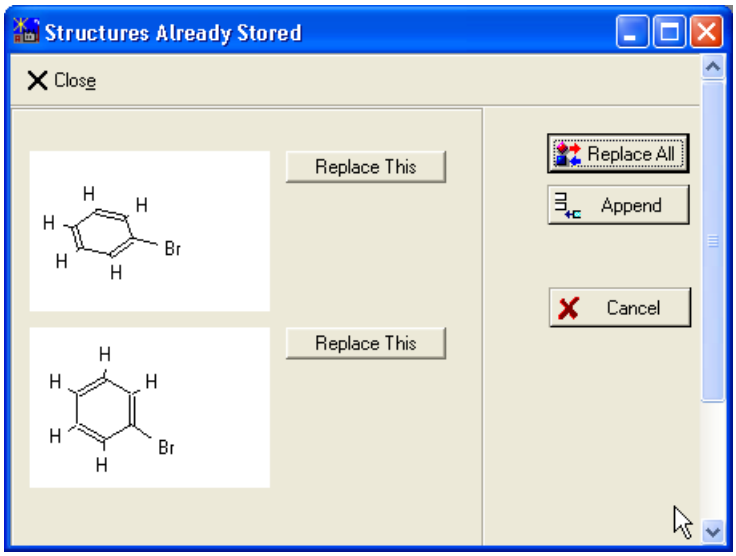


Figure 1: Saving Structures in ChemDB

Delete Structure in ChemDB: Removes a structure from the data bank. Artist displays a dialog from which the structure to be deleted can be selected.

Common note for ChemDB access functions: Artist opens the standard DDB component selection dialog for obtaining the component's DDB number because the structures are indexed (main index) by this component number.

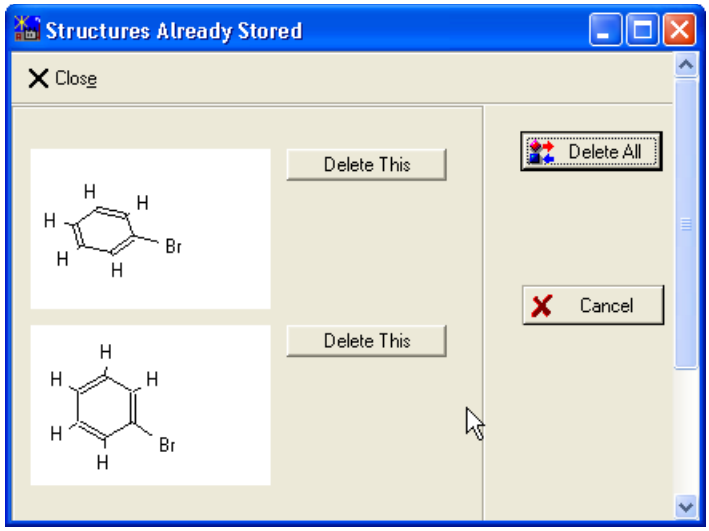


Figure 2: Deleting Structures From ChemDB

Load Component List: Component lists are containing DDB code numbers. Files with component lists have the extension ".stl".If a line is selected the molecular structure of the component will be loaded and displayed in a new drawing window.

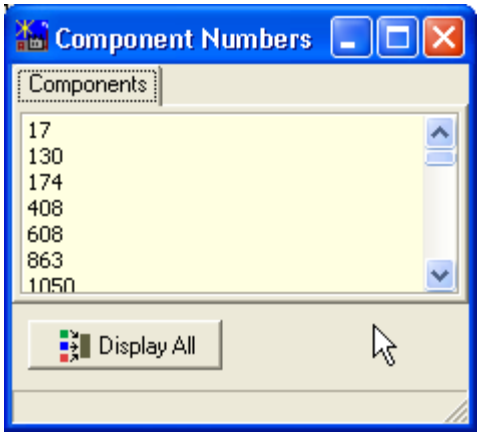


Figure 3: Component List

Load NCI Component:Artist comes with the NCI (National Cancer Institute, U.S.A.) structure data bank which contains approx. a quarter of a million structures mainly for pharmaceutical components. This menu entry allows to load a NCI component by its NCI number.

Append to Compound Definition File: For adding new components Artist provides a dialog which allows to enter name, formula, CAS registry number and molecular weight.

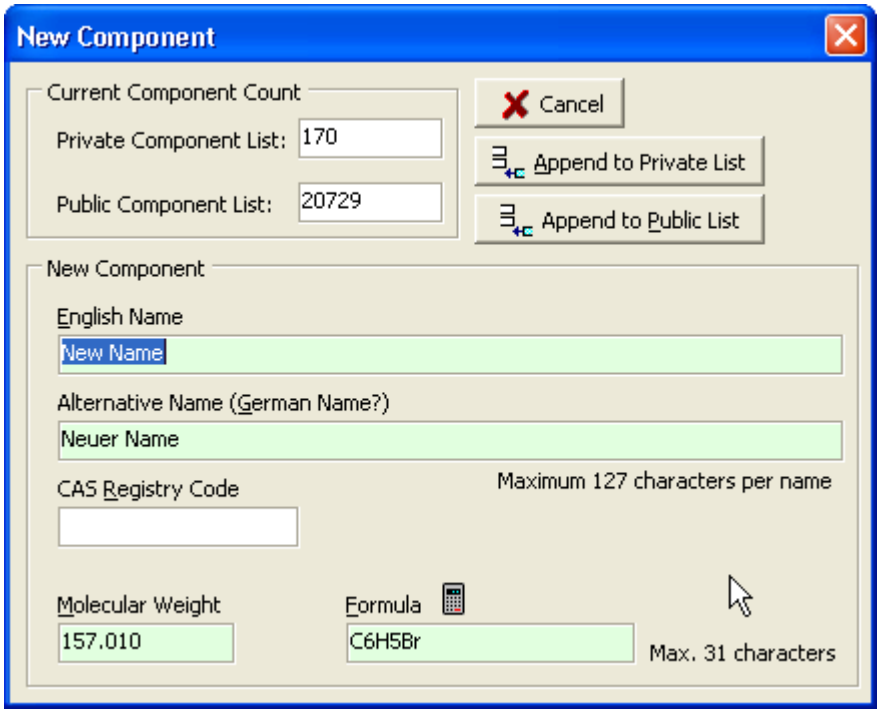


Figure 4: Adding a New Component

Artist can provide the molecular weight and the formula. After entering the necessary values (two names) the component can be appended either to the private (customer) or to the public (DDBST) data bank. Artist automatically updates the necessary files including the component basic file and the structure data bank.

Save: Stores the current structure in the file it has been loaded from.

Save as: Stores the current drawing in a new file.

Print: Prints the current drawing. This function always shows a print preview first.

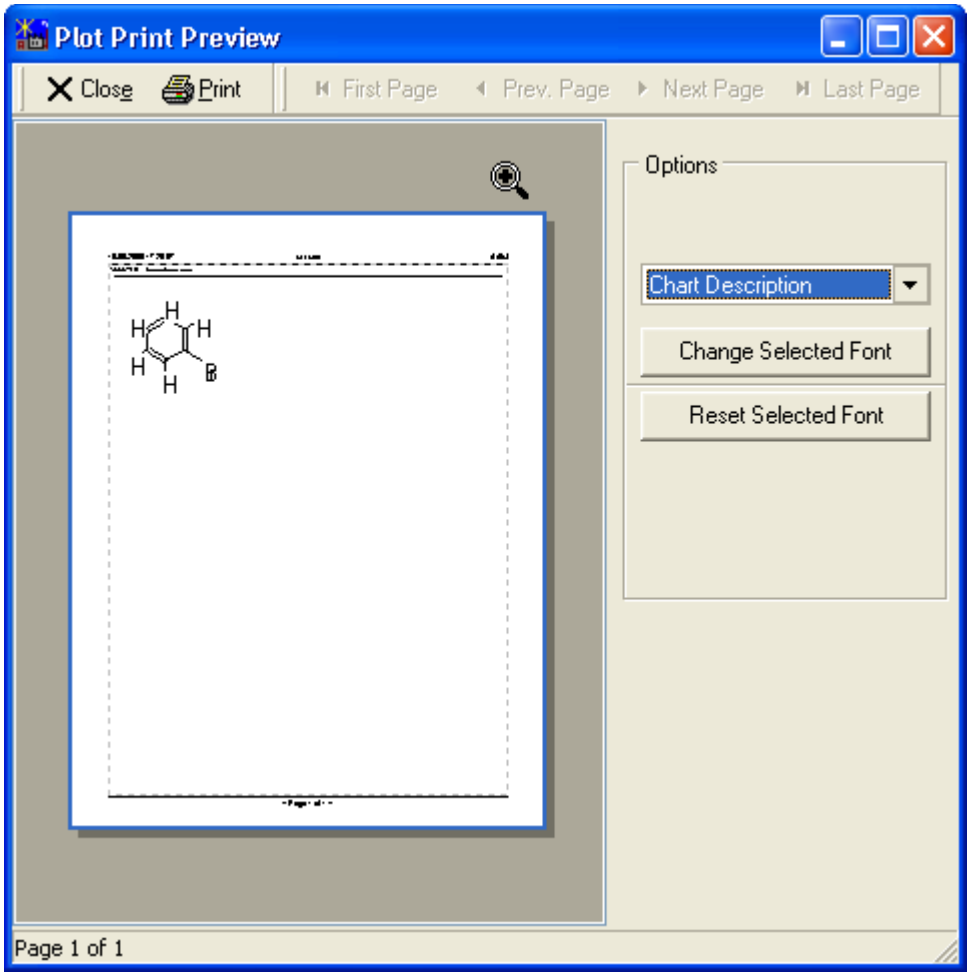
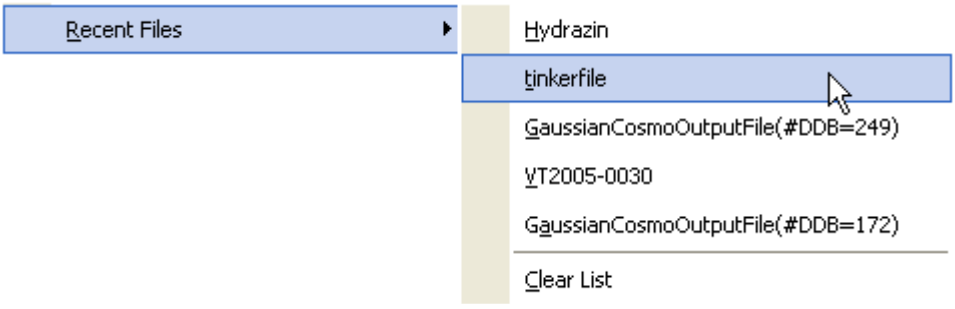


Figure 5: Print Preview Dialog

Print Setup: Allows to select the printer and its properties

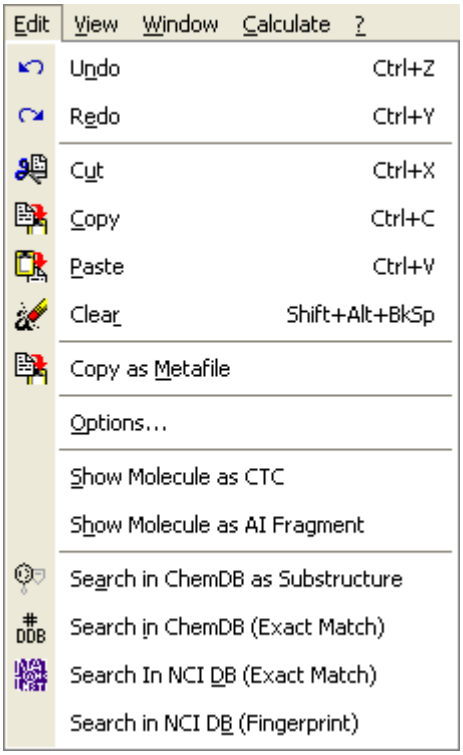
Exit: Closes Artist.

Recent Files: The sub menu is only visible if at least a single file has been loaded and contains the list of the recently load files.



"Clear List" removes all entries from this sub menu.

The Edit Menu



Undo: Undoes the latest action

Redo: Redoes the latest undone action

Cut: Copies a marked area in the Artist internal buffer and deletes the source.

Copy: Copies a marked area in the Artist internal buffer

Paste: Paste a marked area from the Artist internal buffer

Clear: Removes the current drawing

Copy as Meta file: Copies the current drawing as Windows meta file to the Windows clipboard

Options: Displays a option dialog.

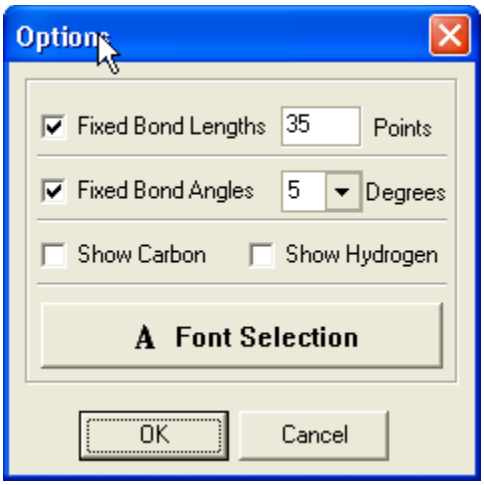


Figure 6: Options

Show Molecule as CTC: Displays the current structure in CTC format (see description in appendix)

Show Molecule as AI Fragment: Displays the current structure in a format usable for the automatic fragmentation algorithm.

Search in ChemDB as Substructure: This function allows to search the molecular structure data bank for components containing the current structure as substructure.

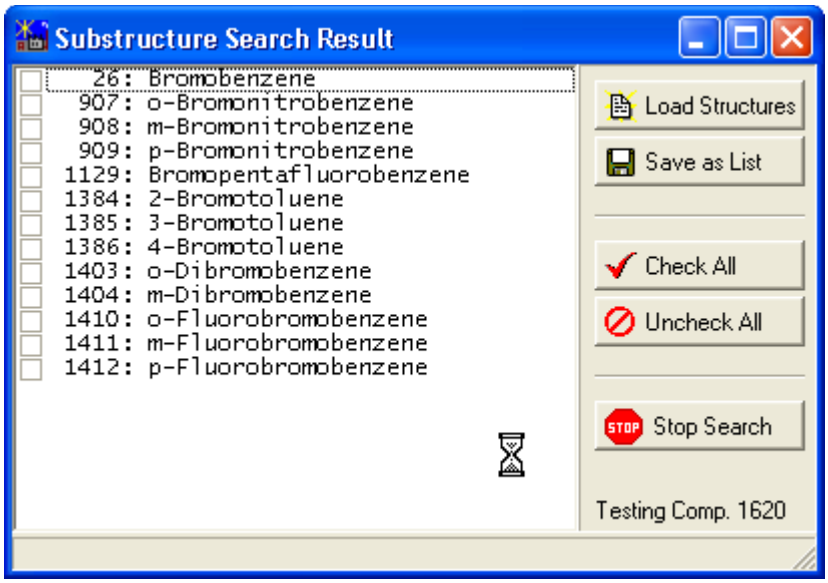


Figure 7: ChemDB Substructure Search

Search in ChemDB (Exact Match): This function allows to search the structure database for exactly matching structures.

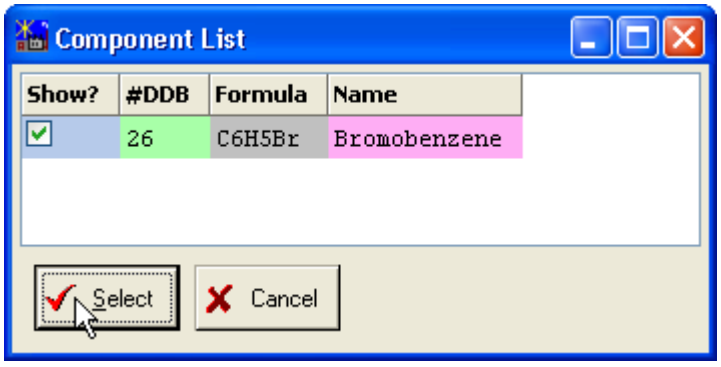


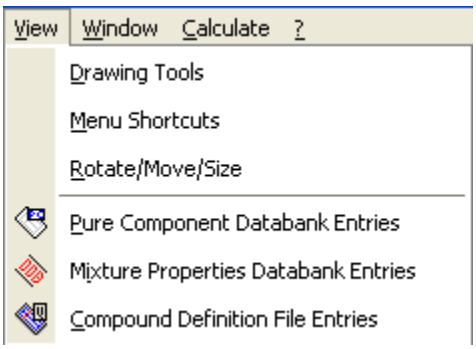
Figure 8: Search for Exact Match - Result

This allows to identify the DDB component number for a molecule which is the main index number in the Dortmund Data Bank. This list can contain multiple structures if conformers are available.

Search in NCI DB (Exact Match): This function searches the NCI data bank for matching structures.

Search in NCI DB (Fingerprint): This allows to search the NCI data bank for components with same (or at least similar) formula.

The View Menu



The first three entries ("Drawing Tools", "Menu Shortcuts", "Rotate/Move/Size") allow to make tool bars visible if they have been teared off and closed.

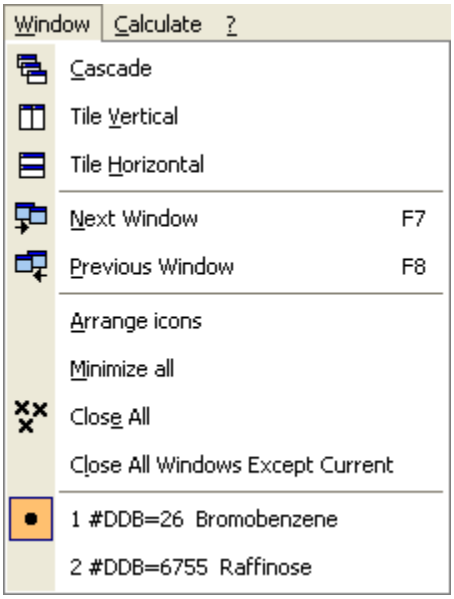
The three other functions allow to start other programs in the DDB Software Package by OLE access. These function are only available if the DDB number of the component is available (see page 10 for "Search in ChemDB – Exact Match")

Pure Component Properties: Calling the Dortmund Data Bank and start a search for pure component properties of the current molecule.

Mixture Properties Data bank Entries: Calls the Dortmund Data Bank program and retrieves all data from mixture properties data banks.

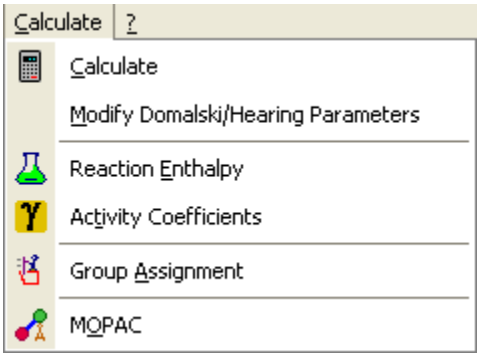
Compound Definition File Entries: Calls the Component Editor and displays the component basic file entries (like names, formula, CAS registry number)

The Window Menu



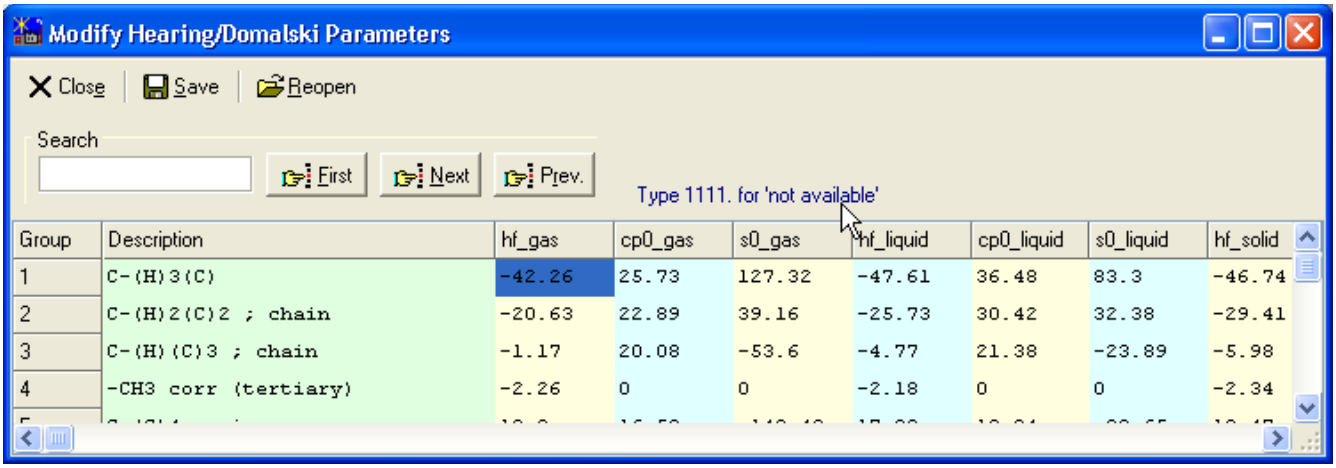
The Window menu contains the standard Windows functions for windows like arrangement, switching to the previous and next window, closing all or some windows, and a list of all windows.

The Calculate Menu



Calculate: Display the calculation dialog – see details in Estimation of Properties chapter.

Modify Domalski/Hearing Parameters: Displays a dialog where group contribution parameters for the Domalski/Hearing models can be modified. Please



Group	Description	hf_gas	cp0_gas	s0_gas	hf_liquid	cp0_liquid	s0_liquid	hf_solid
1	C-(H)3(C)	-42.26	25.73	127.32	-47.61	36.48	83.3	-46.74
2	C-(H)2(C)2 ; chain	-20.63	22.89	39.16	-25.73	30.42	32.38	-29.41
3	C-(H)(C)3 ; chain	-1.17	20.08	-53.6	-4.77	21.38	-23.89	-5.98
4	-CH3 corr (tertiary)	-2.26	0	0	-2.18	0	0	-2.34

Reaction Enthalpy: Display the dialog for estimating the reaction enthalpy – see chapter Reaction Enthalpy for details.

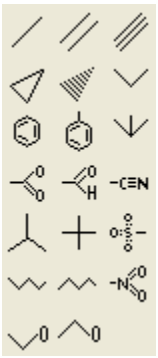
Activity Coefficients: Displays the dialog for estimating activity coefficients – see chapter Activity Coefficients for details.

Group Assignment: Display the dialog for the automatic group assignment see chapter Automatic Group Assignment for details.

MOPAC: Display the dialog for calling MOPAC, Tinker, and RasMol. See chapter MOPAC, Tinker, RasMol for details.

Drawing Structures

Selection of atomic symbol



Drawing tools (single bond, double bond, triple bond, wedges, dimethyl group, benzene, methyl benzene, trimethyl group, carboxylic acids, aldehydes, nitriles, isopropyl, tert-butyl, sulfonate, alkyl chains, nitro, glycol chains)

Rings

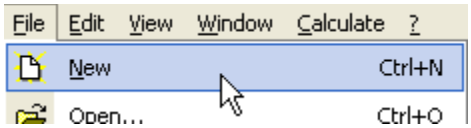
Delete bonds and atoms

Delete all hydrogens

Add hydrogens including a simple coordinates generation

Charges (minus, plus, radical, none)

After creating a new drawing windows (by File-New or ) an empty sheet is displayed.



The drawing tools allow to select some predefined fragments. The most simple fragment is the single bond  .

The drawing mode is 'click and drag'. After selecting the position of the start atom of the bond with pressing the left mouse key the structure can be drawn by moving the mouse while keeping the mouse pressed down.

Artist displays a small green box when the left mouse key is pressed down. This box is the area where drawing of the bond can be canceled. The mouse cursor has to be moved inside this box before the mouse key is released.

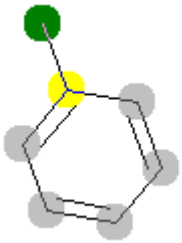
The bond itself follows the movement of the mouse while the left mouse key is kept down. When releasing the mouse key while the mouse pointer is outside the initial green box Artist fixes the bond.

The green box is also shown when the mouse cursor floats above already drawn atoms and bonds. If a new bond is started on an existing atom the bond is fixed to this atom.

If a fragment has more than a single bond and two atoms there's always one atom dedicated as anchor (center) atom which is set with the initial mouse click and an atom connected to the anchor which follows the mouse movement. The other fragment parts are defined by the position of this atom pair and their bond.

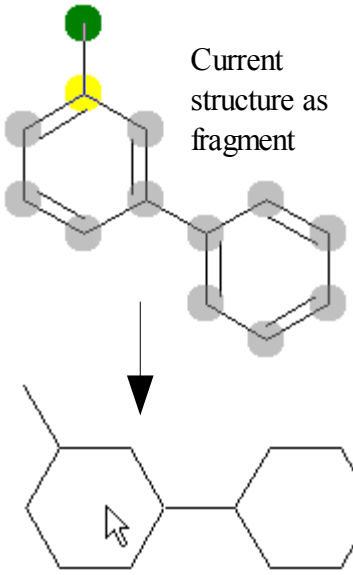
Fragment Management

Artist allows to use drawn structures as fragments.

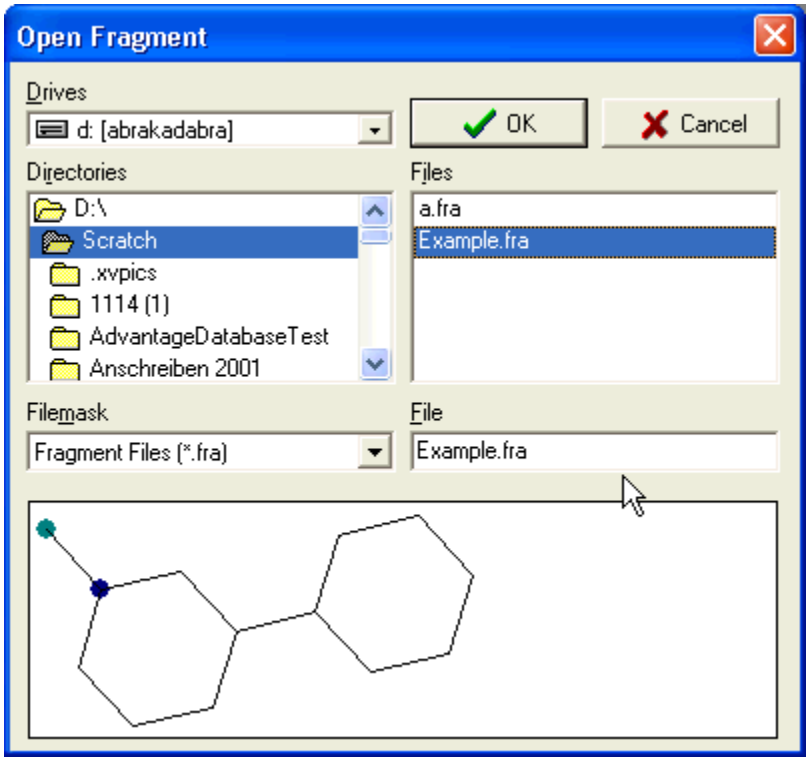


The single icons have the following functions.

- Define fragment anchor atom. This atom of a fragment is set with the mouse. The green circle defines the fragment center.
- Define an atom connected to the anchor atom which will follow the mouse movement. This is the yellow circle in the figure above.
- Directly use the current structure as fragment



- Save fragment to disk (".FRA file)
- Load fragment

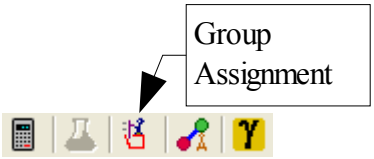

Open Fragment
Drives
d: [abrakadabra]
OK
Cancel
Directories
D:\
Scratch
.xvpics
1114 (1)
AdvantageDatabaseTest
Anschreiben 2001
Files
a.fra
Example.fra
Filemask
Fragment Files (*.fra)
File
Example.fra

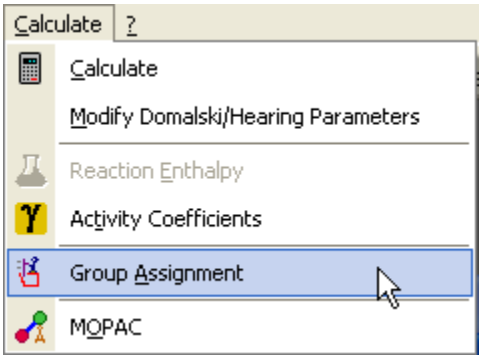
Automatic Group Assignment

The core technology of Artist is the automatic fragmentation algorithm which allows to break a molecules into groups. The groups are the basis for the implementation of group contribution methods.

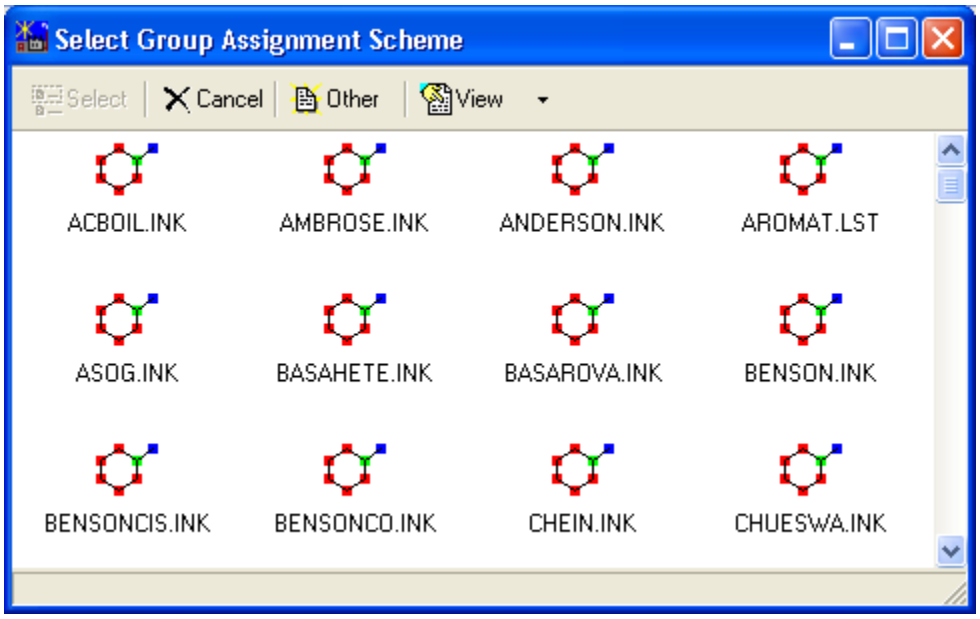
Since almost every method has its own specific list of groups Artist provides a bulk of group assignment schemes separately for each model.

The automatic group assignment can be called from the Calculation menu or by the icon from the tool bar.





The dialog list all files with group assignments (".INK" files). Many of these "files" are integrated inside the Artist programs.



A typical result for the fragmentation of bromo benzene by the UNIFAC model is shown here:

Molecule Description
#DDB=26 Bromobenzene
Comment:

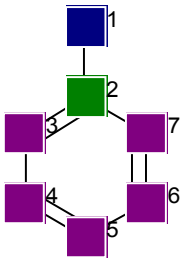
Group Assignment
Method: UNIFAC.INK

Group assignment was successful and complete

Subgroup number: 64 Maingroup number: 33 Groupname: Br
Atoms: 1 in 1 Group
Subgroup number: 10 Maingroup number: 3 Groupname: AC
Atoms: 2 in 1 Group
Subgroup number: 9 Maingroup number: 3 Groupname: ACH
Atoms: 3 4 5 6 7 in 5 Groups

DDB Encoded Group List:
      3   1064   1010   5009

List of rings
 6 membered aromatic ring found. Atoms: 2 3 4 5 6 7

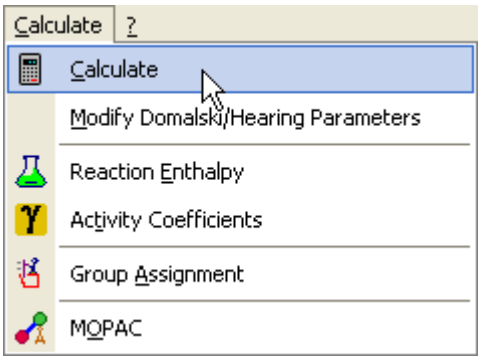


The group assignment found the UNIFAC subgroups 64 (Bromine), 10 (aromatic carbon with three bonds), and 9 (aromatic carbon with two bonds).

The DDB encoded groups ("3 1064 1010 5009") can be used in the component editor ("STOFFEditor"). The "3" determines the number of different groups, "1064" means 1 time the group number 64 , "5009" means that group 9 [ACH] has been found five times.

The group assignment can deliver also a list of ortho/meta/para pairs, a list of rings, a list of chains, and the graphical representation of the assignment. The list of rings, chains, and o-,m-,p-pairs are suppressed if no such structural part are present..

Estimation of Properties



Artist was originally designed for the estimation of pure component properties using group contribution methods, Artist now contains several dozens of different estimation methods. A list of models together with the references can be found in the appendix.

The calculation dialog can called either from menu or from an icon in the tool bar.

The calculation dialog contains three major parts

1. The list of properties in a tree view with the different models grouped below the properties.
2. A dialog part where additionally needed properties (besides the model specific group contributions) can be entered. Artist collects result from previous calculations (a "history" of results) and allows to use these results as input for other models.
3. The list of results either in a table format or as – in some cases more detailed – text output.

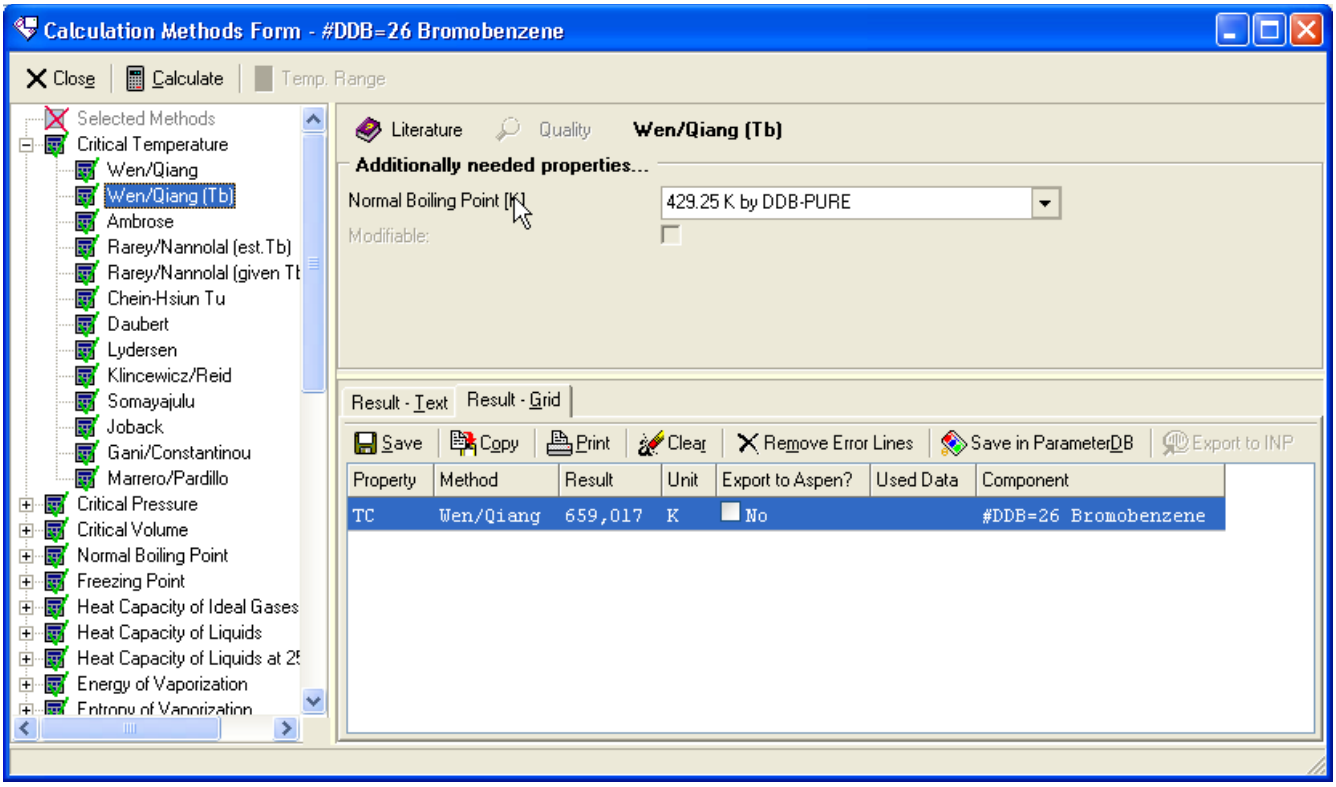
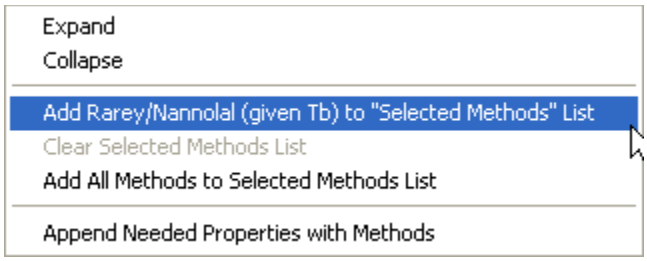


Figure 9: Property Estimation

The branch "Selected Methods" can be used to collect methods from different properties branches. This

allows to make a “selected methods list” for a even quicker access. The different methods can be added by the tree view's context menu entries.



Artist allows to calculate properties by single methods and by all methods for a property. If the property line is selected all methods are calculated in a row. If a single method is selected only this single method is used.

Some methods need additional input besides the structural information. The list of additionally needed properties is displayed in the dialog and the user has to fill the missing information. If data are missing and a complete property calculation is started the methods with missing values will fail.

Artist provides a history of all results calculated or entered before and from the pure component basic file (which are e. g. critical data, see component editor for more details).

Artist presents the estimation results in two different modes – a table and a text output.

The text output

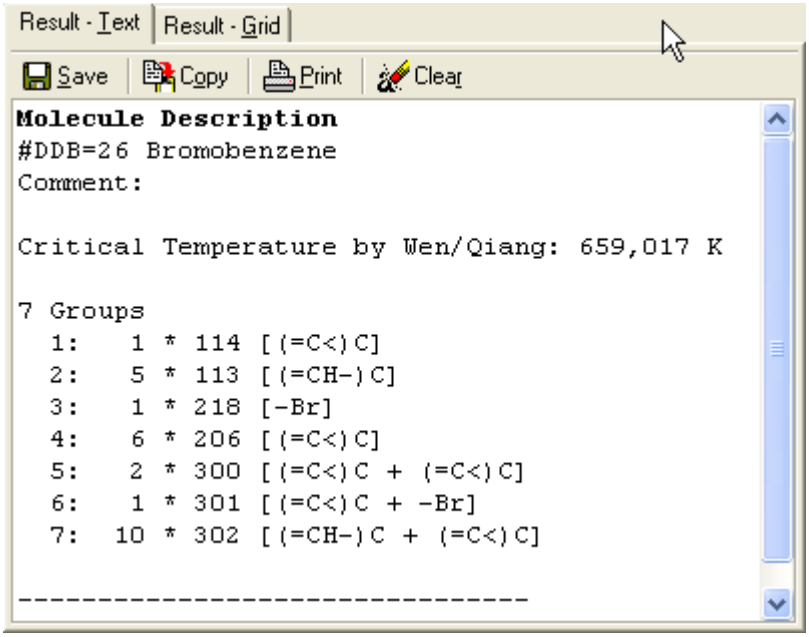


Figure 10: Property Estimation - Text Result

contains additional information – especially the list of groups from the automatic fragmentation.

The table output

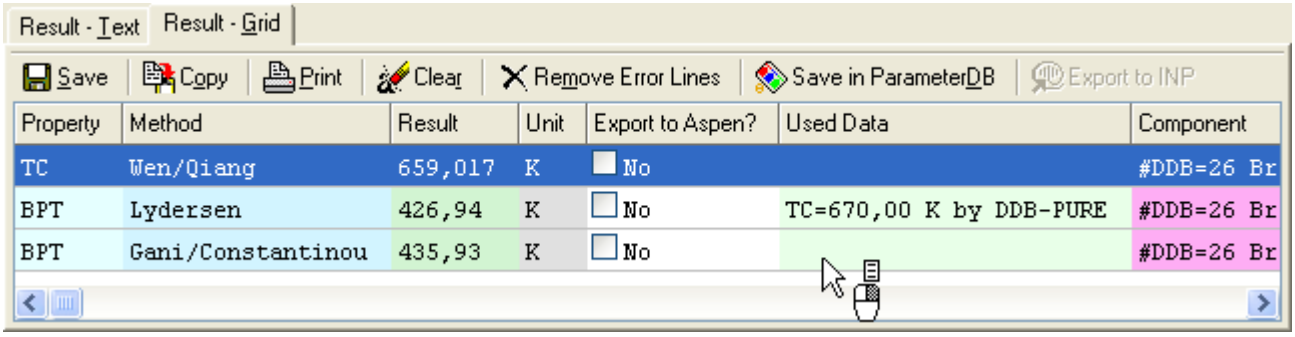
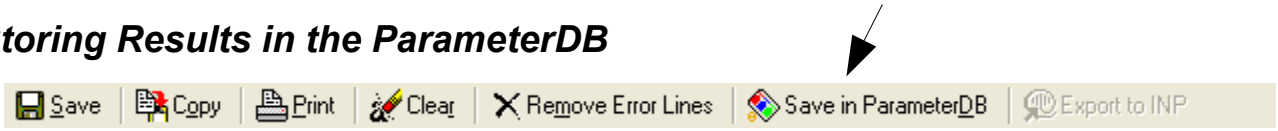


Figure 11: Property Estimation - Table Output

is more concise and allows to copy the result to a spread sheet program. Results can also be stored in the DDB parameter data bank (ParameterDB) or as Aspen INP file (selected properties only).

Storing Results in the ParameterDB



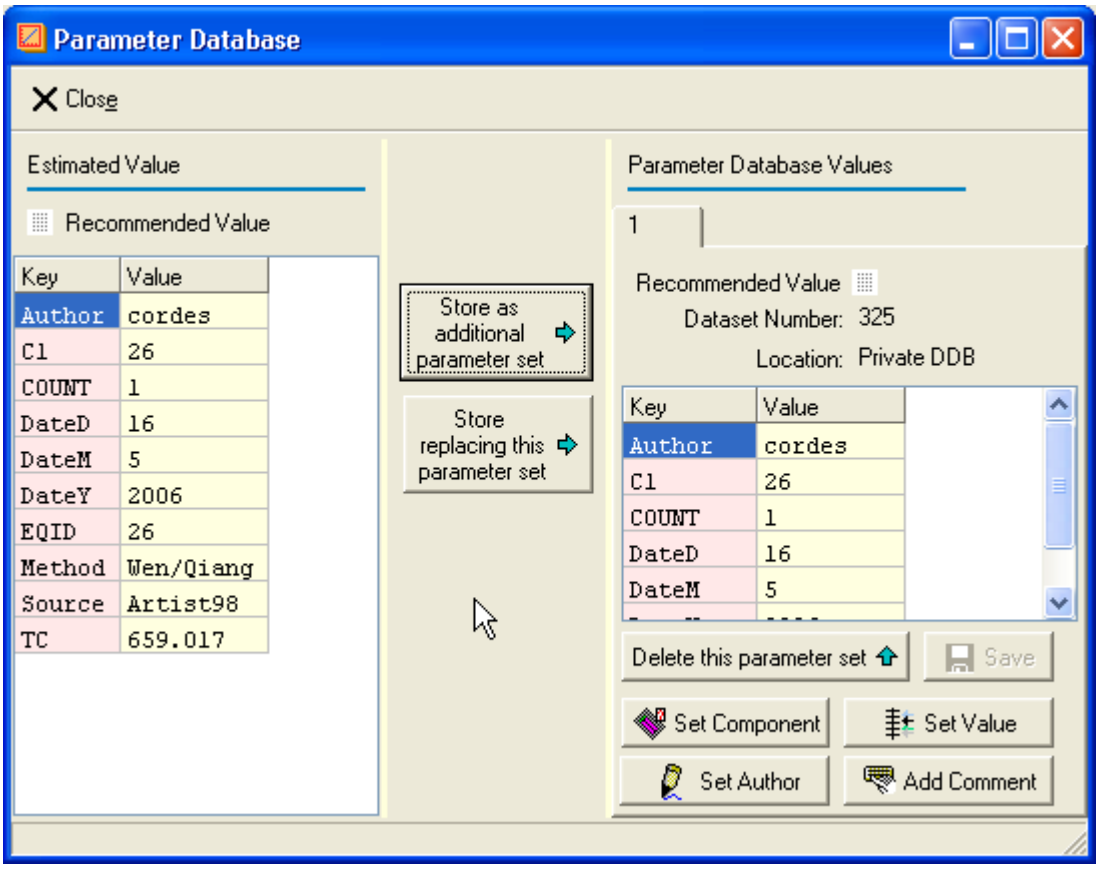


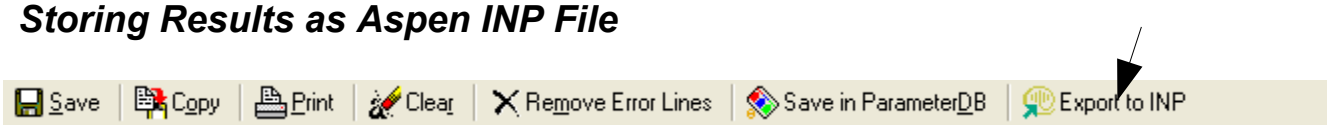
Figure 12: Property Estimation - Storing Results in Parameter Data Bank

The dialog allows to replace or delete already stored sets and to add the current result as a new data set. It also allows to alter some of the data set entries (Component, Author, Value, Comment).

This dialog only allows to store values in a private (customer) data base. The public datasets (delivered from DDBST GmbH) cannot be modified or removed here.

The values stored in the ParameterDB will be available for further calculations.

Storing Results as Aspen INP File

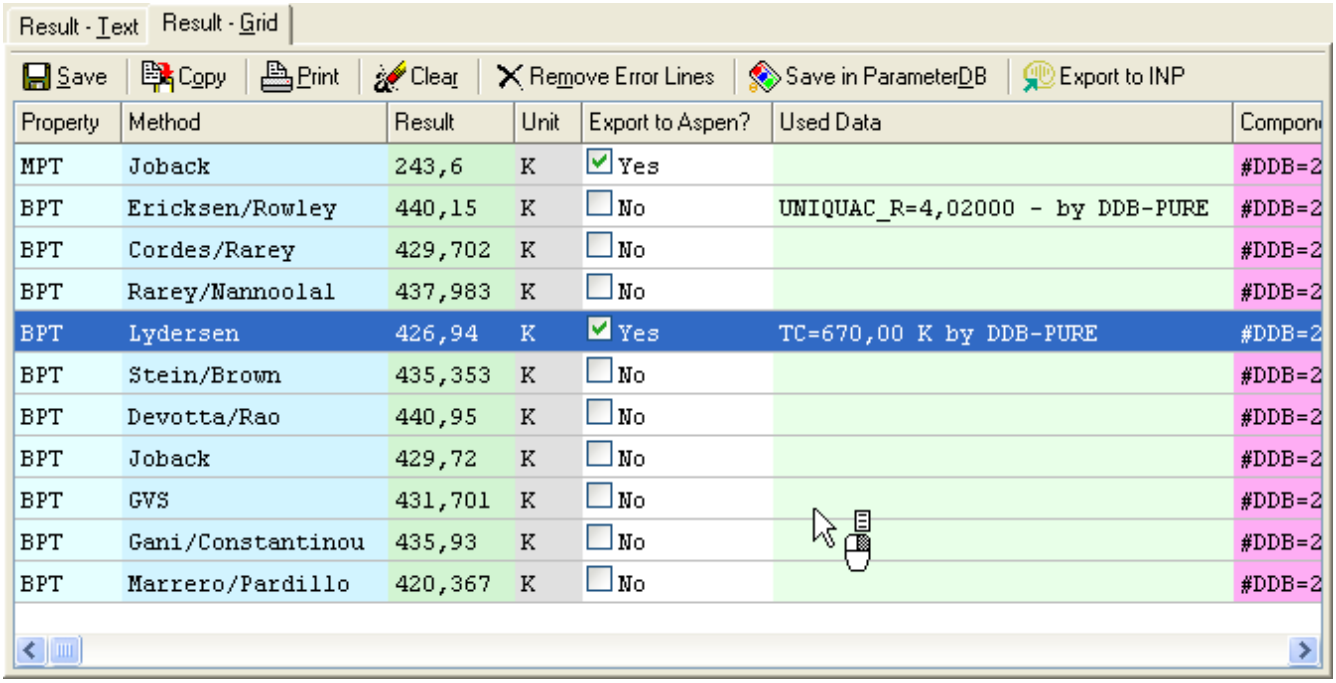


Artist allows to store

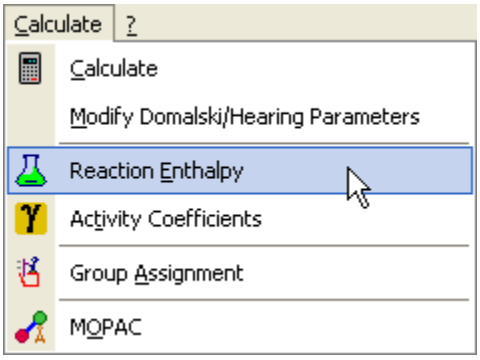
- Critical temperature, pressure, volume
- Normal boiling point
- Melting temperature
- Heat of formation

- Acentric Factor
- UNIQUAC r and q

Artist allows to export only single values of each property for a component. The result grid contains check boxes which allow to select the values intended for export.



Reaction Enthalpy



Artist allows to calculate the reaction enthalpy from estimated and known heats of formation. The calculation dialog is called from the 'Calculate' menu – 'Reaction Enthalpy' or by the icon in the tool bar. The menu entry is only activated if three of more structures have been drawn or loaded.

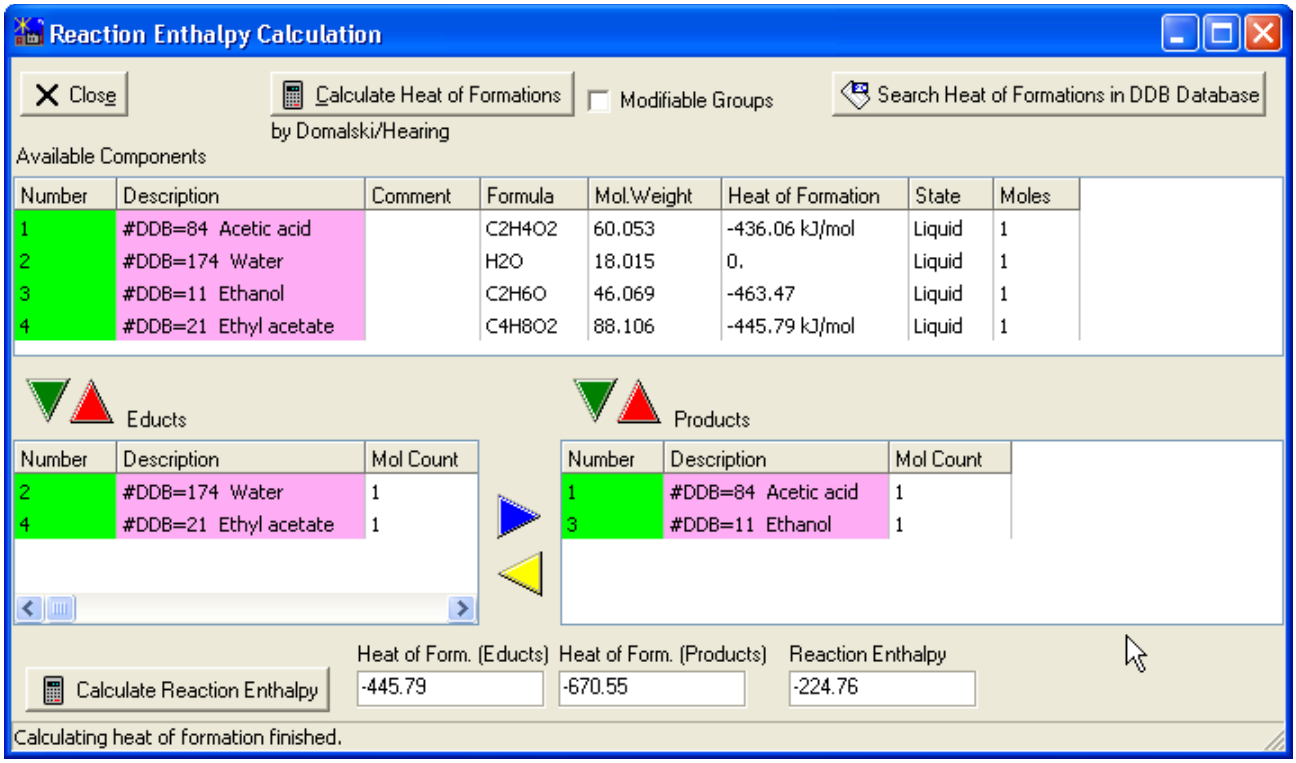


Figure 13: Reaction Enthalpy Calculation

All components are listed initially in the “Available Components” grid. Three types of data have to be provided.

1. The heat of formation of the components – which can be estimated by the Domalski/Hearing method, typed, or search in the Dortmund Data Bank.

2. The state of the component – liquid, gas, solid. This determines the result of the Domalski/Hearing method.
3. The number of moles for all components.

The next step is the assignment of molecules as educts and products. This can be done by selecting the component in the main grid and using the green and red triangle to add and remove them from the sub grids or by drag and drop. If all data are available the button starts the calculation and displays the result.

Search of Heats of Formation in the Dortmund Data Bank

The button 

starts a search in the pure component properties (PCP) branch of the Dortmund Data Bank for heats of formation of the components in the reaction enthalpy dialog.

Only a single heat of formation can be selected for a specific component. The links "[230]" etc. display the reference of the heat of formation.

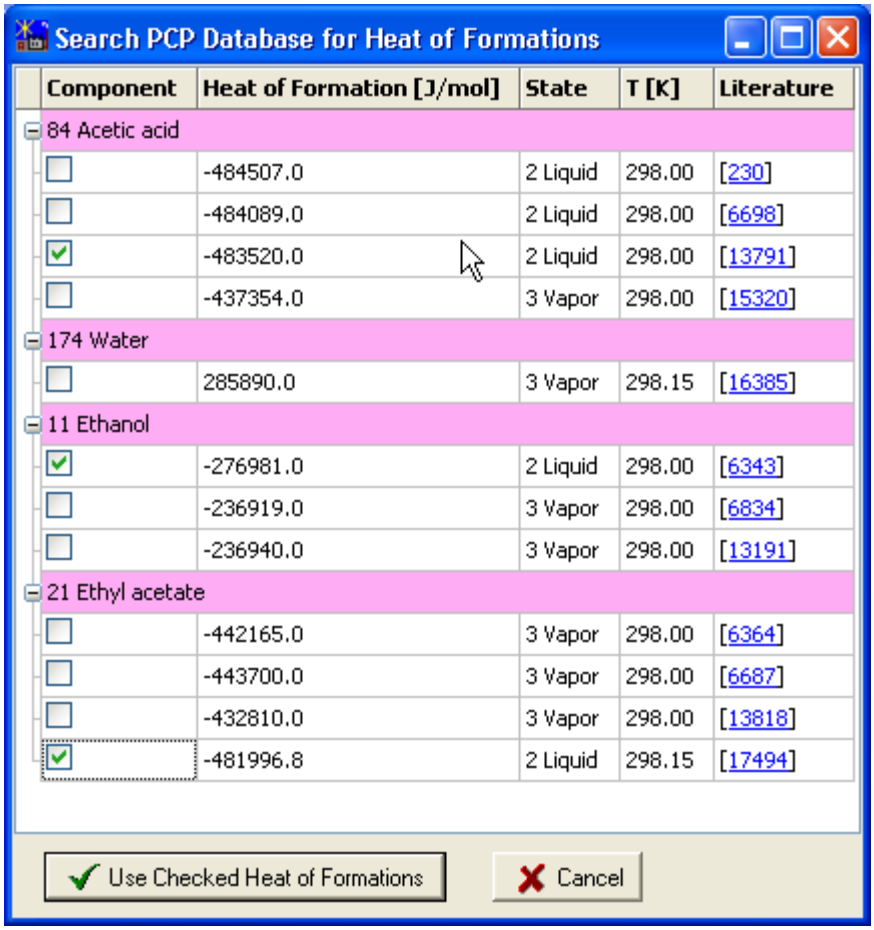
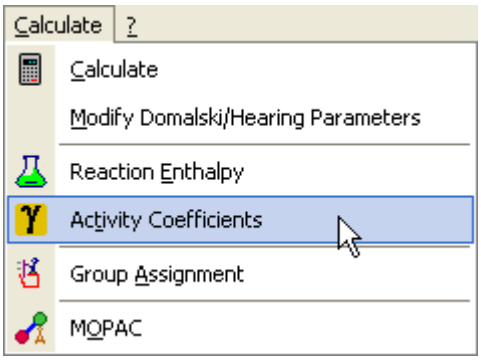


Figure 14: Search Heats of Formation in Dortmund Data Bank

Activity Coefficients



Artist allows to calculate activity coefficients of binary and higher mixtures with the activity coefficient models

- original UNIFAC
- modified UNIFAC (Dortmund)
- modified UNIFAC (Lyngby)
- ASOG

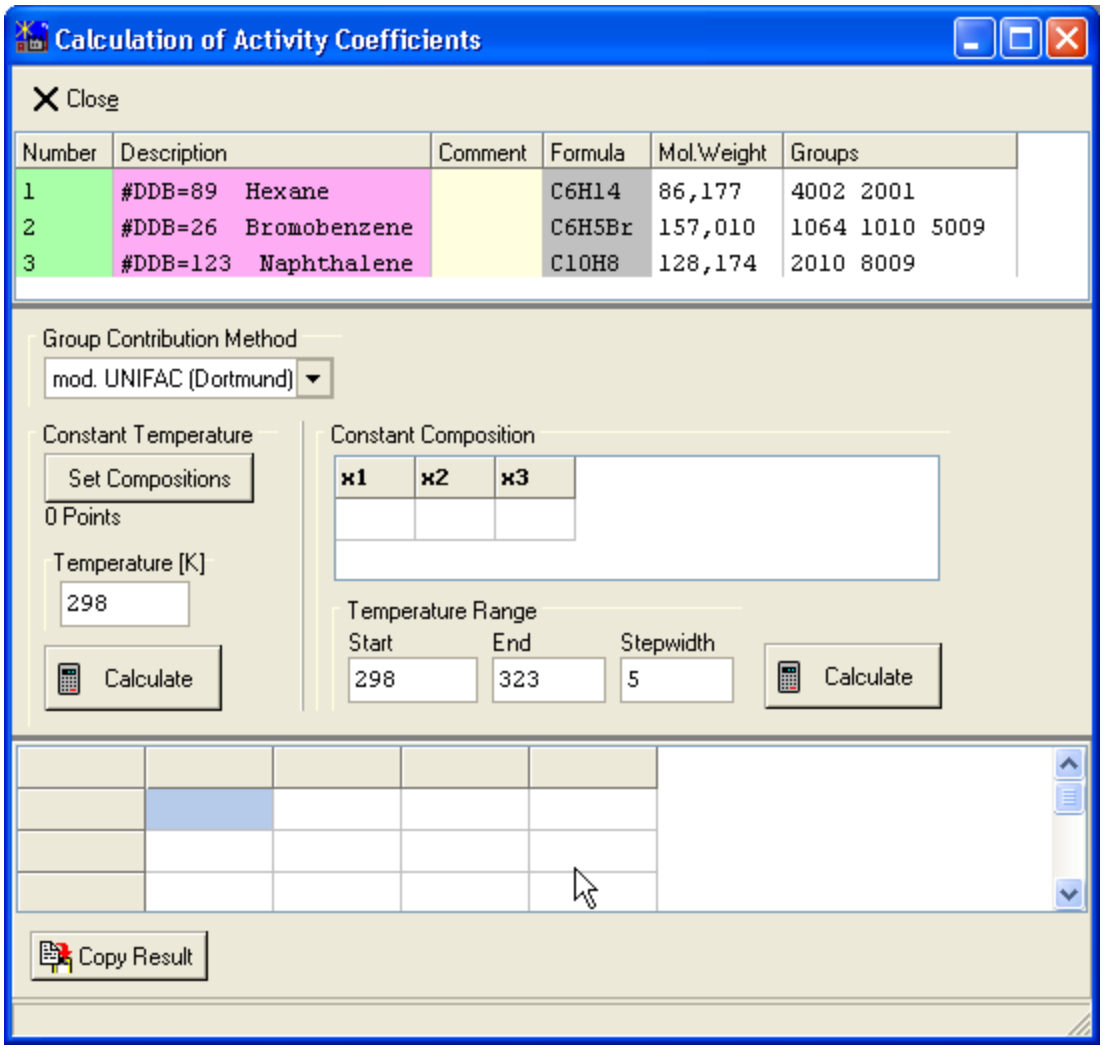


Figure 15: Activity Coefficient Calculation

The activity coefficient dialog automatically uses all components currently displayed and performs an

automatic fragmentation for the chosen model. The groups are shown encoded – e. g. “4002” means 4 times sub group no. 2.

Two calculation modes are available.

1. Calculate activity coefficients for a list of compositions for a constant temperature
2. Calculate activity coefficients for a list of temperatures for a constant composition

The Set Compositions button opens a dialog where the composition can be specified.

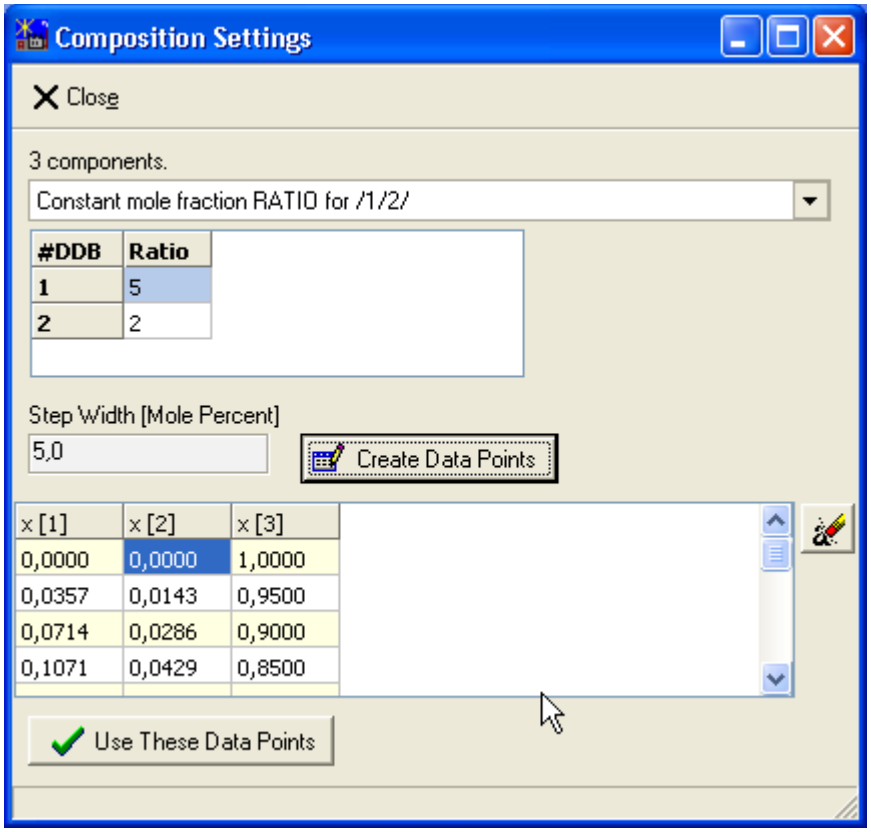
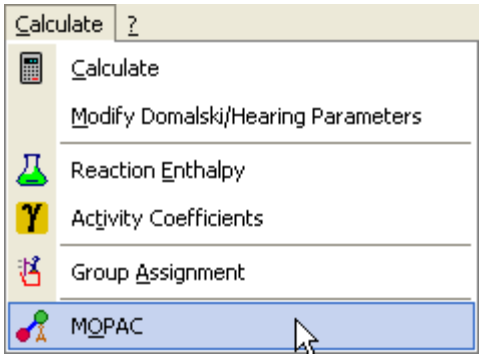


Figure 16: Composition Settings

This dialog allows to keep some compositions constant or to set some composition in a specified ratio. The other compositions are calculated as equally distributed points.

MOPAC, Tinker, RasMol



We cannot distribute neither MOPAC nor Tinker and RasMol the user has to download the programs from the Internet. All these programs are freely available (MOPAC at least up to version 7).

MOPAC

Artist can use MOPAC for some calculations and for structure optimization. Artist has been tested to work MOPAC version 7, the last freely available version.

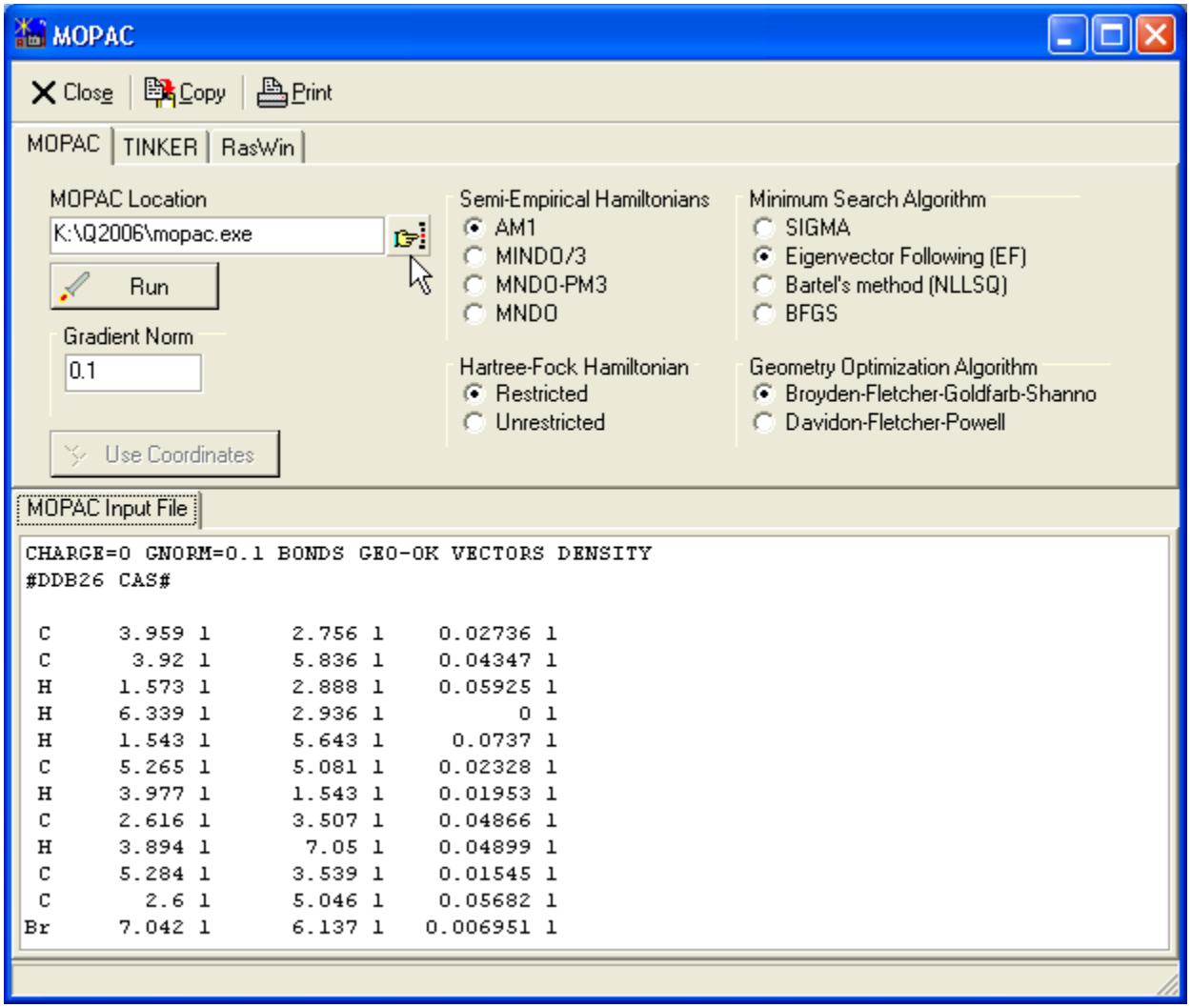
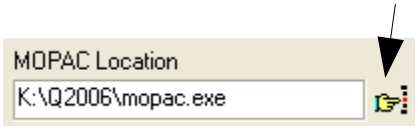


Figure 17: MOPAC Interfacing

Since MOPAC is a standalone program the first step for using the program is defining its location in Artist.

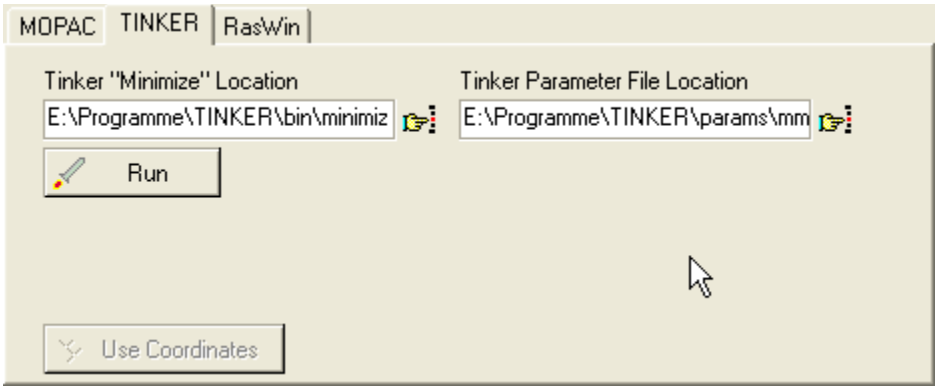


Artist does not much more than writing a MOPAC compatible molecular structure file (shown in the dialog), calling MOPAC by command line and presenting the results. For a description of the different options please use the MOPAC manuals.

The button 

allows to load the coordinates generated by MOPAC.

Tinker

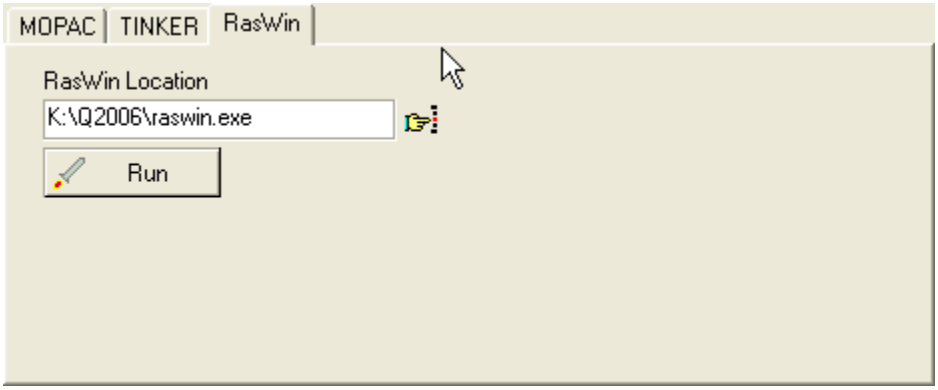


Tinker is a molecular mechanics programs used in Artist for generating 3D structures. The location of the “minimize.exe” program and the location of the “mm3.prm” have to be specified before Tinker can be used.

The button 

can be used to read the coordinates generated by Tinker.

RasMol

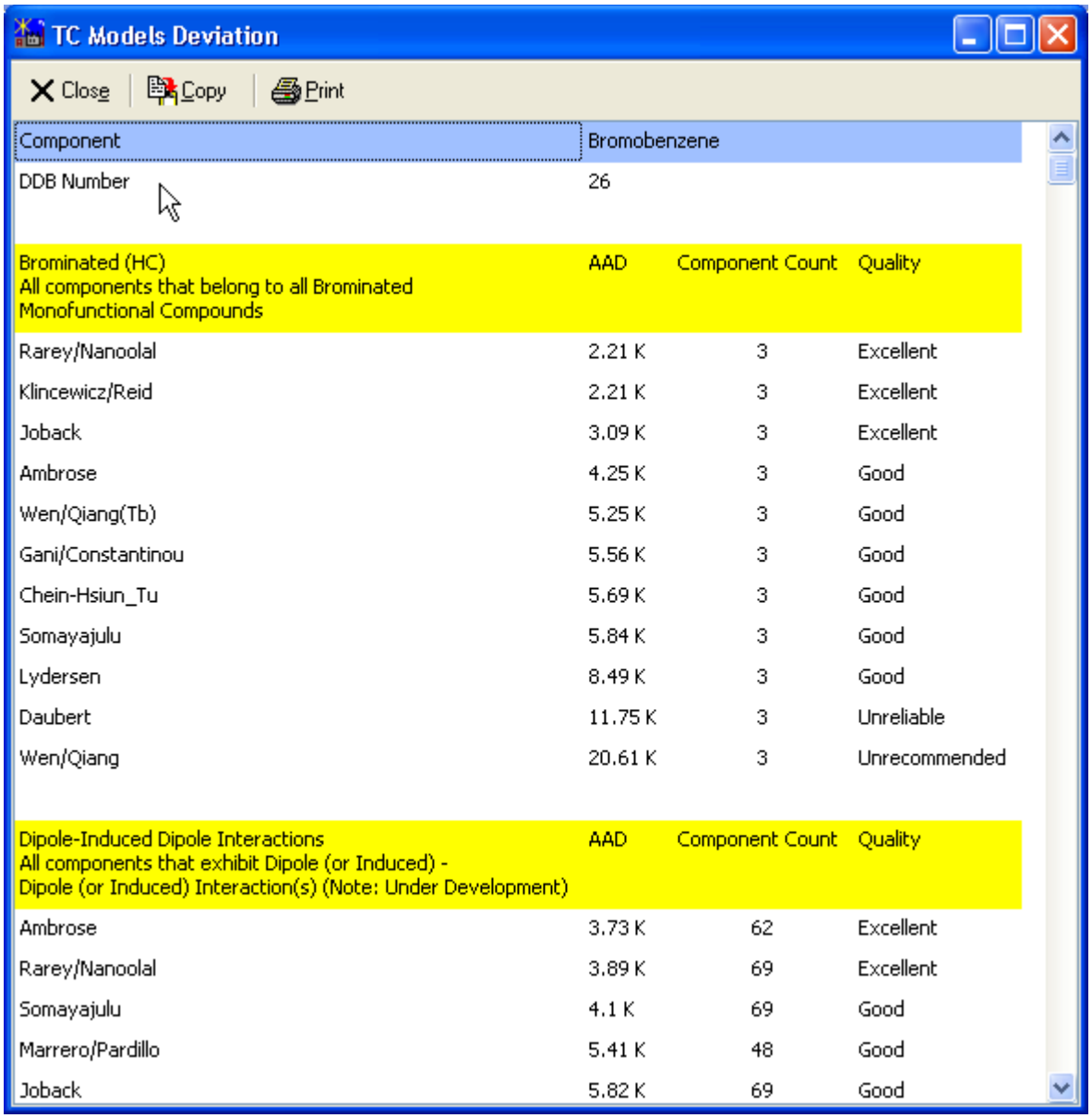


RasMol is a rather simple, but free, display program for chemical structures. Before first use the location of the program has to be specified.

Model Quality

Artist contains informations on mean deviation for different models estimating normal boiling points, critical temperature, critical pressure, and critical volume.

The quality dialogs can be called either by the Q button from the main dialog or by the Quality button from the calculation dialog.



Component	Bromobenzene		
DDB Number	26		
Brominated (HC) All components that belong to all Brominated Monofunctional Compounds	**AAD**	**Component Count**	**Quality**
Rarey/Nanoolal	2.21 K	3	Excellent
Klincewicz/Reid	2.21 K	3	Excellent
Joback	3.09 K	3	Excellent
Ambrose	4.25 K	3	Good
Wen/Qiang(Tb)	5.25 K	3	Good
Gani/Constantinou	5.56 K	3	Good
Chein-Hsiun_Tu	5.69 K	3	Good
Somayajulu	5.84 K	3	Good
Lydersen	8.49 K	3	Good
Daubert	11.75 K	3	Unreliable
Wen/Qiang	20.61 K	3	Unrecommended
Dipole-Induced Dipole Interactions All components that exhibit Dipole (or Induced) - Dipole (or Induced) Interaction(s) (Note: Under Development)	**AAD**	**Component Count**	**Quality**
Ambrose	3.73 K	62	Excellent
Rarey/Nanoolal	3.89 K	69	Excellent
Somayajulu	4.1 K	69	Good
Marrero/Pardillo	5.41 K	48	Good
Joback	5.82 K	69	Good

Figure 18: Quality Data Bank

The data bank contains quality information for component classes. The current atom is assigned to

these classes and the stored information for the appropriate component classes. The component classes are organized hierarchical. There are comprehensive classes like “Hydrocarbons” and more detailed classes like “Aromatic Hydrocarbons”.The AAD (average absolute deviation) is given in Kelvin, the “Component Count” column displays the number of tested components and the “Quality” column gives a (quite arbitrary) comment on the quality.

Appendix

File Formats

Artist allows to store four and read four file formats. Readable formats are

1. The proprietary DDBST format with the extension ".CTC".
2. The MOL format ".MOL" defined by MDL Information Systems (http://www.mdli.com/).
3. The Tinker ".XYZ" format.
4. The COSMO formats from Gaussian and Turbomole

Writable formats are

1. The proprietary DDBST format with the extension ".CTC".
2. The MOL format ".MOL" defined by MDL Information Systems (http://www.mdli.com/).
3. The MOPAC-Z format
4. The Gaussian ".gjf" format.

The format of the MOL format can be obtained from MDL (http://www.mdli.com/solutions/white_papers/ctfile_formats.jsp).

The CTC File Format

The CTC format is specially designed to match the requirements of the program Artist with respect to storing molecular structures. It is a simple tagged format in pure ASCII. It can be viewed and edited by simple editors such as Notepad or UltraEdit.

The program uses the following tags:

Tag	*Description*
#ATOMS	list and description of atoms in the molecule
#BONDS	list and description of bonds between the atoms
#CAS	CAS registry number of the component
#FORMULA	empirical formula of the component
#ENAME	English name of the component
#TIME	file creation time
#PROGRAM	program the file was created with
#COMMENT	comment

Tag	*Description*
#FILE	original filename

These tags start blocks within the file. Blocks are ended by another tag or the end of file. There is no special order in which the blocks must appear.

#ATOMS

This block contains a list and description of the atoms in the molecule. Hydrogen atoms can be included or may be omitted. The first line must contain the number of atoms in the molecule. The following lines, one for each atom, contain the following entries separated by at least one blank character:

- x-, y- and z-coordinate (Cartesian).
- atomic symbol. Only symbols from the standard PSE are allowed. Especially, functional groups like COOH or NO2 are not legal entries.
- Charge or radical. The following predefined numbers are used to encode the different types of charges or radicals:
    - 0 → no charge, no radical
    - 1 → charge +3
    - 2 → charge +2
    - 3 → charge +1
    - 4 → radical
    - 5 → charge -1
    - 6 → charge -2
    - 7 → charge -3

    The same way of encoding is used in mol files by MDL (Molecular Design). Additional number are

    - 13 → charge +4
    - 14 → charge +5
    - 11 → charge +6
    - 15 → charge -4
    - 16 → charge -5
    - 17 → charge -6
- mass differences to the most common isotope to specify different isotopes.
- chemical environment. The following environment are defined:
    - K → aliphatic chain
    - R → aliphatic ring
    - A → aromatic system
    - N → non-aromatic neighborhood
    - C → aromatic or ring (cyclic neighborhood)

    - * → not specified
- molecule number (a ctc structure can contain multiple structures)

#BONDS

This block contains a list and description of the bonds between the atoms. The first line must contain the number of different bonds, double and triple bonds are counted as one bond. The following lines, one for each bond, contain the following entries separated by at least one blank character:

- atom counter of the first atom.
- atom counter of the second atom.
- bond multiplicity (1 – single bond, 2 – double bond, 3 – triple bond)
- bond orientation allows simple coding of stereo chemistry.
    - 0 → not specified
    - 1 → in plane
    - 5 → in front of plane
    - 6 → behind plane
- chemical environment in format %c. The following environment are used:
    - K → aliphatic chain
    - R → aliphatic ring
    - A → aromatic system
    - N → non-aromatic neighborhood
    - C → aromatic or ring (cyclic neighborhood)
    - * → not specified

#CAS

This block contains the CAS registry number of the component.

#FORMULA

This block contains the empirical formula of the component.

#ENAME

This block contains the English name of the component.

#DATE

This block contains the file creation date. The format is (day. Month. year).

#TIME

This block contains the file creation time as the only entry. The format is (hour: minute: second).

#PROGRAM

This block contains the name of the program the file was created with as the only entry.

#COMMENT

This block contains comments in free format (single line)

#FILE

This block contains the original filename.

Examples:

The following examples illustrate the use of the ctc format.

Ethanol (hydrogen atoms omitted)

```
#DDB
11
#DATE
20.6.2000
#TIME
19:50:59
#ATOMS
3
    87.97         35          0  C  0  0 *  1
       35      42.93      14.03  O  0  0 *  1
    65.88      48.36      23.63  C  0  0 *  1
#BONDS
2
 1  3  1  1 *
 2  3  1  1 *
```

Ethyl benzene (hydrogen atoms included)

```
#DDB
25
#DATE
20.6.2000
#TIME
19:53:30
#ATOMS
18
    149.4      59.29      41.23  C  0  0 *  1
    68.11      116.3      31.62  H  0  0 *  1
    171.1      113.6      93.87  C  0  0 *  1
    224.4      82.79      44.89  C  0  0 *  1
    238.7      131.1      93.24  H  0  0 *  1
       35      126.7      77.62  H  0  0 *  1
    92.86      98.73      125.6  H  0  0 *  1
    68.21      124.8      65.88  C  0  0 *  1
    200.3         35          0  H  0  0 *  1
    257.4      80.93      33.24  H  0  0 *  1
    192.2      56.86      26.11  C  0  0 *  1
      163      135.4        120  H  0  0 *  1
    124.6      39.28      26.76  H  0  0 *  1
       77      58.97      87.07  H  0  0 *  1
    138.9      87.63       75.1  C  0  0 *  1
    213.9      111.1      78.76  C  0  0 *  1
    84.01        156       70.2  H  0  0 *  1
     92.8      90.24      91.39  C  0  0 *  1
#BONDS
```

```
18
 4 10   1   1 *
 1 13   1   1 *
11  9   1   1 *
 3 12   1   1 *
 4 11   1   1 *
 8 17   1   1 *
 8  2   1   1 *
 8  6   1   1 *
 4 16   2   1 *
 1 15   1   1 *
16  5   1   1 *
11  1   2   1 *
 3 15   2   1 *
 8 18   1   1 *
15 18   1   1 *
16  3   1   1 *
18  7   1   1 *
18 14   1   1 *
```

List of References

Critical Properties T_c, P_c, V_c

1. Wen X., Qiang Y., "A New Group Contribution Method for Estimating Critical Properties of Organic Compounds", Ind. Eng. Chem. Res. 40, 6245 - 6250 (2001)
2. Ambrose D. in Reid R.C., Prausnitz J.M., Poling B.E., "The Properties of Gases & Liquids", Monograph, McGraw-Hill, 4th Ed., 12 - 22 (1987)
3. Ambrose D., "Correlation and Estimation of Vapour-Liquid Critical Properties. I. Critical Temperatures of Organic Compounds", National Physical Laboratory, Teddington, NPL Rep. Chem. 92, September 1978, corrected March 1980
4. Ambrose D., "Correlation and Estimation of Vapour-Liquid Critical Properties. II. Critical Pressures and Volumes of Organic Compounds", National Physical Laboratory, Teddington, NPL Rep. Chem. 98 (1979)
5. Tu Ch.-Hs., "Group-Contribution Estimation of Critical Temperature With Only Chemical Structure", Chem. Eng. Sci. 50(22), 3515 - 3520, 1995
6. Daubert T.E, "State of the Art Property Predictions", Hydrocarb. Proc. & Petr. Ref., 107 - 112 (1980)
7. Lydersen A. L. in Reid R.C., Prausnitz J.M., Poling B.E., "The Properties of Gases & Liquids", Monograph, McGraw-Hill, 3th Ed., 12 - 22 (1977)
8. Lydersen A. L. , "Estimation of Critical Properties of Organic Compounds", Univ. Wisconsin Coll. Eng., Eng. Exp. Stn. rept. 3, Madison, WI, April 1955
9. Klincewicz K.M., Reid R.C., "Estimation of Critical Properties with Group Contribution Methods", AIChE J. 30(1), 137 - 142 (1984)
10. Somayajulu G.R., "New Vapour Pressure Equations from Triple Point to Critical Point and a Predictive Procedure for Vapour Pressure", J. Chem. Eng. Data 31, 438 - 447 (1986)
11. Joback K.G., Reid R.C., "Estimation of Pure-Component Properties from Group-Contributions", Chem. Eng. Commun. 57, 233 - 243 (1987)
12. Gani R., Constantinou L., "Molecular Structure Based Estimation of Properties for Process Design", Fluid Phase Equilib. 116, 75 – 86 (1996)
13. Marrero-Morejon J., Pardillo-Fontdevila E., "Estimation of Pure Compound Properties Using Group-Interaction Contributions", AIChE J., 45(3), 615 – 621 (1999)
14. Wen X., Wenying W., "Group Vector Space Method for Estimating the Critical Properties of Organic Compounds", Ind.Eng.Chem.Res., 42(24), p6258-6262, 2003

Normal Boiling Point T_b

1. Cordes W., Rarey J., "A New Method for the Estimation of the Normal Boiling Point of Non-electrolyte organic compounds", Fluid Phase Equilib. 201, 409 - 433 (2002)
2. Nannoolal Y., Rarey J., Ramjugernath D., Cordes W., " Estimation of Pure Component Properties, Part 1: Estimation of the Normal Boiling Point of Non-Electrolyte Organic Compounds via Group Contributions and Group Interactions (Preprint 2004)

3. Lydersen A. L. in Reid R.C., Prausnitz J.M., Poling B.E., "The Properties of Gases & Liquids", Monograph, McGraw-Hill, 3th Ed., *12 - 22* (1977)
4. Stein S.E., Brown R.L., Estimation of Normal Boiling Points from Group Contributions", J. Chem. Inf. Comput. Sci. 34, 581 - 587 (1994)
5. Devotta S., Rao Pendyala V., Modified Joback Group Contribution Method for Normal Boiling Point of Aliphatic Halogenated Compounds", Ind. Eng. Chem. Res. 31, 2042 -2046 (1992)
6. Joback K.G., Reid R.C., "Estimation of Pure-Component Properties from Group-Contributions", Chem. Eng. Commun. 57, 233 - 243 (1987)
7. Gani R., Constantinou L., "Molecular Structure Based Estimation of Properties for Process Design", Fluid Phase Equilib. 116, 75–86 (1996)
8. Marrero-Morejon J., Pardillo-Fontdevila E., "Estimation of Pure Compound Properties Using Group-Interaction Contributions", AIChE J., 45(3), 615 – 621 (1999)
9. Wen X., Qiang Y., "Group Vector Space (GVS) Method for Estimating Boiling and Melting Points of Hydrocarbons", J.Chem.Eng.Data, 47(2), p286-288, 2002

Freezing Point T_m

1. Joback K.G., Reid R.C., "Estimation of Pure-Component Properties from Group-Contributions", Chem. Eng. Commun. 57, 233-243 (1987)
2. Gani R., Constantinou L., "Molecular Structure Based Estimation of Properties for Process Design", Fluid Phase Equilib. 116, 75-86 (1996)
3. Wen X., Qiang Y., "Group Vector Space (GVS) Method for Estimating Boiling and Melting Points of Hydrocarbons", J.Chem.Eng.Data, 47(2), p286-288, 2002

Heat Capacity of Ideal Gases

1. Harrison B.K., Seaton W.H., "Solution to Missing Group Problem for Estimation of Ideal Gas Heat Capacities", Ind.Eng.Chem.Res., 27(8), p1536-1540, 1988
2. Benson (implemented from "The Properties of Gases and Liquids", 4th edition).
3. Benson S.W., Cruickshank F.R., Golden D.M., Haugen G.R., O'Neal H.E., Rodgers A.S., Shaw R., Walsh R., "Additivity Rules for the Estimation of Thermochemical Properties", Chem.Rev., 69, p279-324, 1969
4. Joback K.G., Reid R.C., "Estimation of Pure-Component Properties from Group-Contributions", Chem. Eng. Commun. 57, p233-243 (1987)

Heat Capacity of Liquids

1. Ruzicka V., Domalski E.S., "Estimation of the Heat Capacities of Organic Liquids as a Function of Temperature Using Group Additivity. II. Compounds of Carbon, Hydrogen, Halogens, Nitrogen, Oxygen, and Sulfur", J.Phys.Chem.Ref.Data, 22(3), p619-657, 1993
2. Missenard F.-A., "Méthode additive pour la détermination de la chaleur molaire de liquides.", Compt.Rend.Acad.Sci.Paris, 260, p5521-5523, 1965
3. Garvin J., "Calculate Heats of Combustion for Organics", Chem.Eng.Prog.(CEP), 94(5), p43-

45, 1998
4. Luria M., Benson S.W., "Heat Capacities of Liquid Hydrocarbons. Estimation of Heat Capacities at Constant Pressure as a Temperature Function, Using Additivity Rules" , J.Chem.Eng.Data, 22(1), p90-100, 1977

Heat Capacity of Liquids at 25 °C

1. Shaw R., "Heat Capacities of Liquids Estimation of Heat Capacity at Constant Pressure and 25°C., Using Additivity Rules", J.Chem.Eng.Data, 14(4), p461-465, 1969
2. Chueh C.F., Swanson A.C., "Estimation of Liquid Heat Capacity", Can.J.Chem.Eng., 51, p596-600, 1973

Energy of Vaporization

1. Horvath A.L., "Molecular design : chemical structure generation from the properties of pure organic compounds", Reihe / Serie, Elsevier, p1-1489, 1992

Entropy of Vaporization

1. Hoshino D., Nagahama K., Hirata M., "Prediction of the Entropy of Vaporization at the Normal Boiling Point by the Group Contribution Method", Ind.Eng.Chem. Fundam., 22(4), p430-433, 1983
2. Ma P., Zhao X., "Modified Group Contribution Method for Predicting the Entropy of Vaporization at the Normal Boiling Point", Ind.Eng.Chem.Res., 32, p3180-3183, 1993

Enthalpy of Fusion

1. Joback K.G., Reid R.C., "Estimation of pure-component properties from group-contributions", Chem.Eng.Commun., 57, p233-243, 1987

Gibbs Energy of Formation

1. Mavrovouniotis M.L., "Estimation of Standard Gibbs Energy Changes of Biotranformations", J.Biol.Chem., 266(22), p14440-14445, 1991
2. Joback K.G., Reid R.C., "Estimation of pure-component properties from group-contributions", Chem.Eng.Commun., 57, p233-243, 1987
3. Gani R., Constantinou L., "Molecular Structure Based Estimation of Properties for Process Design", Fluid Phase Equilib., 116, p75-86, 1996

Heat of Formation

1. Benson (implemented from "The Properties of Gases and Liquids", 4th edition).
2. Benson S.W., Cruickshank F.R., Golden D.M., Haugen G.R., O'Neal H.E., Rodgers A.S., Shaw R., Walsh R., "Additivity Rules for the Estimation of Thermochemical Properties", Chem.Rev., 69, p279-324, 1969
3. Domalski E.S., Hearing E.D., "Estimation of the Thermodynamic Properties of C-H-N-O-S-

Halogen Compounds at 298.15 K", J.Phys.Chem.Ref.Data, 22(4), p805-1159, 1993
4. Joback K.G., Reid R.C., "Estimation of pure-component properties from group-contributions", Chem.Eng.Commun., 57, p233-243, 1987
5. Gani R., Constantinou L., "Molecular Structure Based Estimation of Properties for Process Design", Fluid Phase Equilib., 116, p75-86, 1996

Heat of Vaporization

1. Basarova P., Svoboda V., "Prediction of the Enthalpy of Vaporization by the Group Contribution Method", Fluid Phase Equilib., 105, p27-47, 1995
2. Tu C.-H., Liu C.-P., "Group-contribution estimation of the enthalpy of vaporization of organic compounds", Fluid Phase Equilib., 121, p45-65, 1996

Standard Heat of Vaporization

1. Basarova P., Svoboda V., "Prediction of the Enthalpy of Vaporization by the Group Contribution Method", Fluid Phase Equilib., 105, p27-47, 1995
2. Tu C.-H., Liu C.-P., "Group-contribution estimation of the enthalpy of vaporization of organic compounds", Fluid Phase Equilib., 121, p45-65, 1996
3. Hoshino D., Nagahama K., Hirata M., "Prediction of the Entropy of Vaporization at the Normal Boiling Point by the Group Contribution Method", Ind.Eng.Chem. Fundam., 22(4), p430-433, 1983
4. Joback K.G., Reid R.C., "Estimation of pure-component properties from group-contributions", Chem.Eng.Commun., 57, p233-243, 1987
5. Gani R., Constantinou L., "Molecular Structure Based Estimation of Properties for Process Design", Fluid Phase Equilib., 116, p75-86, 1996

Liquid Viscosity

1. Orrick, Erbar (see "The Properties of Gases and Liquids, 4th edition, cited as "Private Communication")
2. Tatevskii V.M., Abramenkov A.V., Grikina O.E., "Calculation of the Temperature Dependence of the Viscosity of Liquids", Zh.Fiz.Khim., 61(10), p2770-2774, 1987
3. Van Velzen D., Cardozo R.L., Langenkamp H., "A liquid viscosity-temperature-chemical constitution relation for organic compounds", Ind.Eng.Chem.,Fundam., 11(1), p20-25, 1972
4. Thomas L.H., "The Dependence of the Viscosities of Liquids on Reduced Temperature, and a Relation between Viscosity, Density, and Chemical Constitution", J.Chem.Soc.London, p573-579, 1946
5. Souders, Jr. M., "Viscosity and chemical constitution", J.Am.Chem.Soc., 60, p154-158, 1938
6. Sastri S. R. S., Rao K. K., "A new method for predicting saturated liquid viscosity at temperatures above the normal boiling point", Fluid Phase Equilib., 175(1-2), p311-323, 2000
7. Skubla P., "Prediction of Viscosity of Organic Liquids", Collect.Czech.Chem.Commun., 50, p1907-1916, 1985
8. Joback K.G., Reid R.C., "Estimation of pure-component properties from group-contributions",

Chem.Eng.Commun., 57, p233-243, 1987

Vapor Viscosity

1. Reichenberg D., "New Methods for the Estimation of the Viscosity Coefficients of Pure Gases at Moderate Pressures (with Particular Reference to Organic Vapors)", AIChE J., 21(1), p181-183, 1975

Vapor Pressure

1. Gomez-Nieto M., Thodos G., "Generalized Vapor Pressure Behavior of Substances Between Their Triple Points and Critical Points", AIChE J., 23(6), p904-913, 1977
2. Riedel L., "Eine neue universelle Dampfdruckformel. Untersuchungen über eine Erweiterung des Theorems der übereinstimmenden Zustände. Teil I.", Chem.Ing.Tech., 26(2), p83-89, 1954
3. Miller D.G., "Estimating Vapor Pressures - A Comparision of Equations", Ind.Eng.Chem., 56(3), p46-57, 1954

Acentric Factor

1. Han B., Peng D.-Y., "A Group-Contribution Correlation for Predicting the Acentric Factors of Organic Compounds", Can.J.Chem.Eng., 71, p332-334, 1993
2. Hoshino D., Nagahama K., Hirata M., "Prediction of Acentric Factor of Alkanes by the Group Contribution Method", J.Chem.Eng.Jpn., 15(2), p153-155, 1982
3. Constantinou L., Gani R., O'Connell J.P., "Estimation of the Acentric Factor and the Liquid Molar Volume at 298 K using a New Group Contribution Method", Fluid Phase Equilib., 103, p11-22, 1995

Liquid Density

1. Tsibanogiannis I.N., Kalospiros N.S., Tassios D.P., "Extension of the GCVOL Method and Application to Some Complex Compounds", Ind.Eng.Chem.Res., 33, p1641-1643, 1994
2. Campbell S. W., Thodos G., "Prediction of Saturated-Liquid Densities and Critical Volumes fo Polar and Nonpolar Substances", J.Chem.Eng.Data, 30(1), p102-111, 1985
3. GCVOL
   Elbro H.S., Fredenslund A., Rasmussen P., "Group Contribution Method for the Prediction of Liquid Densities as a Function of Temperature for Solvents, Oligomers, and Polymers", Ind.Eng.Chem.Res., 30(12), p2576-2582, 1991
4. GCVOL60 (compl. new)
   Ihmels E.C., Gmehling J., "Extension and Revision of the Group Contribution Method GCVOL for the Prediction of Pure Compound Liquid Densities", Ind.Eng.Chem.Res., 42(2), p408-412, 2003
5. GCVOL60(Orig. plus new)
   Ihmels E.C., Gmehling J., "Extension and Revision of the Group Contribution Method GCVOL for the Prediction of Pure Compound Liquid Densities", Ind.Eng.Chem.Res., 42(2), p408-412, 2003

Liquid Density at 20 °C

1. Kolbasov V.I., Emel'yanov V.I., Motsarev G.V., "Method for Calculating the Density of the Halogen and Methyl Derivatives of Benzene", Zh.Prikl.Khim. (Leningrad) 6, 56(6), p1350-1353, 1983

Liquid Density at Normal Boiling Point

1. Schroeder (siehe "The Properties of Gases and Liquids, 4th edition)
2. Tyn M.T., Calus W.F., "", Processing, 21(4), p16, 1975
3. Sastri S.R.S., Mohanty S., Rao K., "Liquid Volume at Normal Boiling Point", Can.J.Chem.Eng., 74, p170-172, 1996
4. Le Bas G., "The Molecular Volumes of Liquid Chemical Compounds", Monograph, Longmans, Green, 1915 (siehe "The Properties of Gases and Liquids, 4th Edition)

Liquid Molar Volume at 25 °C

1. Constantinou, L., Gani. R., O'Connell, J.P., "Estimation of the acentric factor and the liquid molar volume at 298 K using a new group contribution method", Fluid Phase Equilib. 0, 103(0), p11-22, 1995

Thermal Conductivity

1. Nagvekar M., Daubert T.E., "A Group Contribution Method for Liquid Thermal Conductivity", Ind.Eng.Chem.Res., 26, p1362-1365, 1987
2. Sastri S.R.S., Rao K.K., "Quick Estimating for Thermal Conductivity", Chem.Eng.(New York), 8, p107-107, 1993
3. Lakshmi D.S., Prasad D.H.L., "A Rapid Estimation Method for Thermal Conductivity of Pure Liquids", Chem.Eng.J., 48, p211-214, 1992

Upper Flammability Limit

1. High M.S., Danner R.P., "Prediction of Upper Flammability Limit by a Group Contribution Method.", Ind.Eng.Chem.Res., 26, p1395-1399, 1987

Surface Tension

1. Brock J.R., Bird R.B., "Suface Tension and the Principle of Corresponding States", AIChE J., 1(2), p174-177, 1955

Water Solubility

1. Kühne R., Ebert R.-U., Kleint F., Schmidt G., Schüürmann G., "Group Contribution Methods to Estimate Water Solubility of Organic Chemicals", Chemosphere 11, 30(11), p2061-2077, 1995

Octanol Water Partiton Coefficents

1. Wienke G., Gmehling J., "Prediction of octanol-water partition coefficients, Henry coefficients

and water solubilities using UNIFAC", Toxicol.Environ.Chem. 0, 65(0), p57-86, 1998

Second Virial Coefficients

1. Reid R.C., Prausnitz J.M., Poling B.E., "The Properties of Gases & Liquids", Monograph, McGraw-Hill, 4 Ed., p1-742, 1987
2. McCann D.W., Daubert R.P., "Prediction of Second Virial Coefficients of OrganicCompounds by a Group Contribution Method", Ind.Eng.Chem.,Process Des.Dev., 23, p529-533, 1984

Domalski/Hearing

1. Domalski E.S., Hearing E.D., "Estimation of the Thermodynamic Properties of C-H-N-O-S-Halogen Compounds at 298.15 K", J.Phys.Chem.Ref.Data, 22(4), p805-1159, 1993

UNIQUAC by UNIFAC Groups

1. UNIFAC Consortium, Carl-von-Ossietzky-University Oldenburg, Prof. Jürgen Gmehling, http://www.unifac.org

Topological Indices

1. Platt. See "The Properties of Liquids and Gases", 4th edition, Ambrose-Model.

Parachor

1. Knotts T. A., Wilding W. V., Oscarson J. L. , Rowley R. L., "Use of the DIPPR Database of Development of QSPR Correlations: Surface Tensions", J.Chem.Eng.Data, 46(5), p1007-1012, 2001