MedeA: Structure List Editor: Manage Your Structure Repository With Properties

Contents

1 Introduction

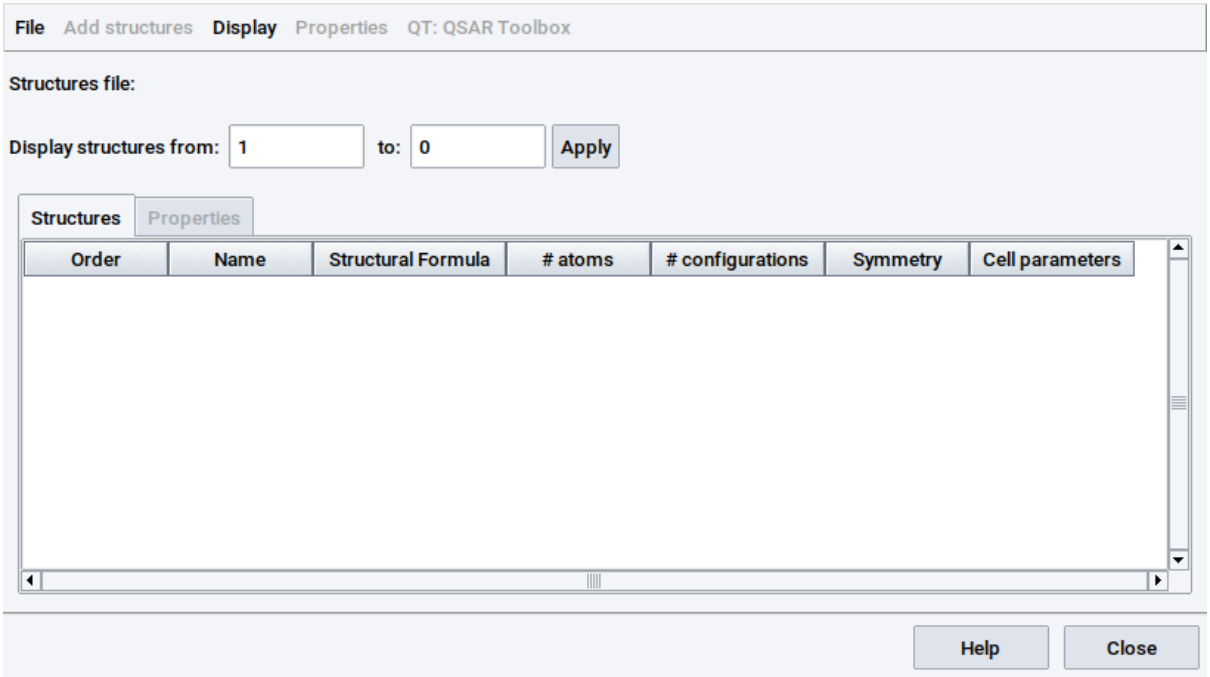
Define a list of structures, with arbitrary configurations, compositions and associated properties and descriptors. Structures can be simple molecules, a fluid (ensemble of molecules to form a gaseous or liquid state), or solids in a crystalline or an amorphous state. Structure lists are stored as a single file either in the database format SQLite or as clearly readable text in the ASCII format. With structure list you can archive and store your calculated results (properties, descriptors) together with structures that are used to calculate the properties. Also use structure lists to leverage the power of flowcharts to perform high-throughout (HT) screening and use HT applications). For HT calculations, use the *MedeA* module *HT-Launchpad* and the flowchart stage **Foreach Structure Loop** to apply the same compute workflow on each record of the structure list. With this you can obtain consistently calculated data. Another application of structure lists is to predict properties of compounds using correlations with the *MedeA* modules *HT-Descriptor* and *QT*. The latter is the *MedeA* QSPR Toolbox, that employs an interactive graphical user interface to allow you to explore and analyze the relationships between descriptors, structural properties, and materials properties.

For more information about structure lists in the context of high-throughput calculations read the Section MedeA HT.

In case you are interested to use structure lists in conjunction with *MedeA QT* then read Section MedeA QT.

The structure list editor is opened in the *MedeA GUI* with File >> Structure List Editor. Without any structure list loaded the invoked dialogue looks such as depicted in the below image.

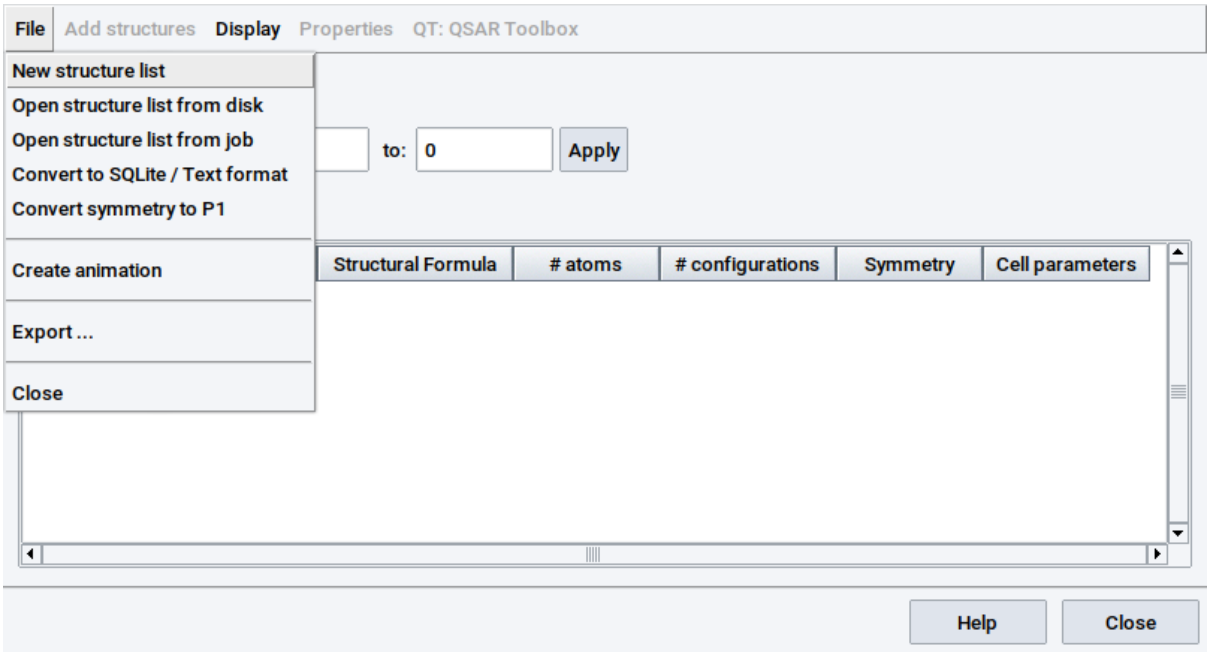
DOCUMENTATION



Note: Every structure list editor can handle only one structure list. However, you can open several structure list editors in parallel via File >> Structure List Editor.

2 Handle Entire Structure Lists

An entire structure list can be handled via File menu item of the structure list editor. You can create a new and empty structure list, open existing structure lists, convert structure lists and their entire content, and export the properties of the structure list to a file on disk:



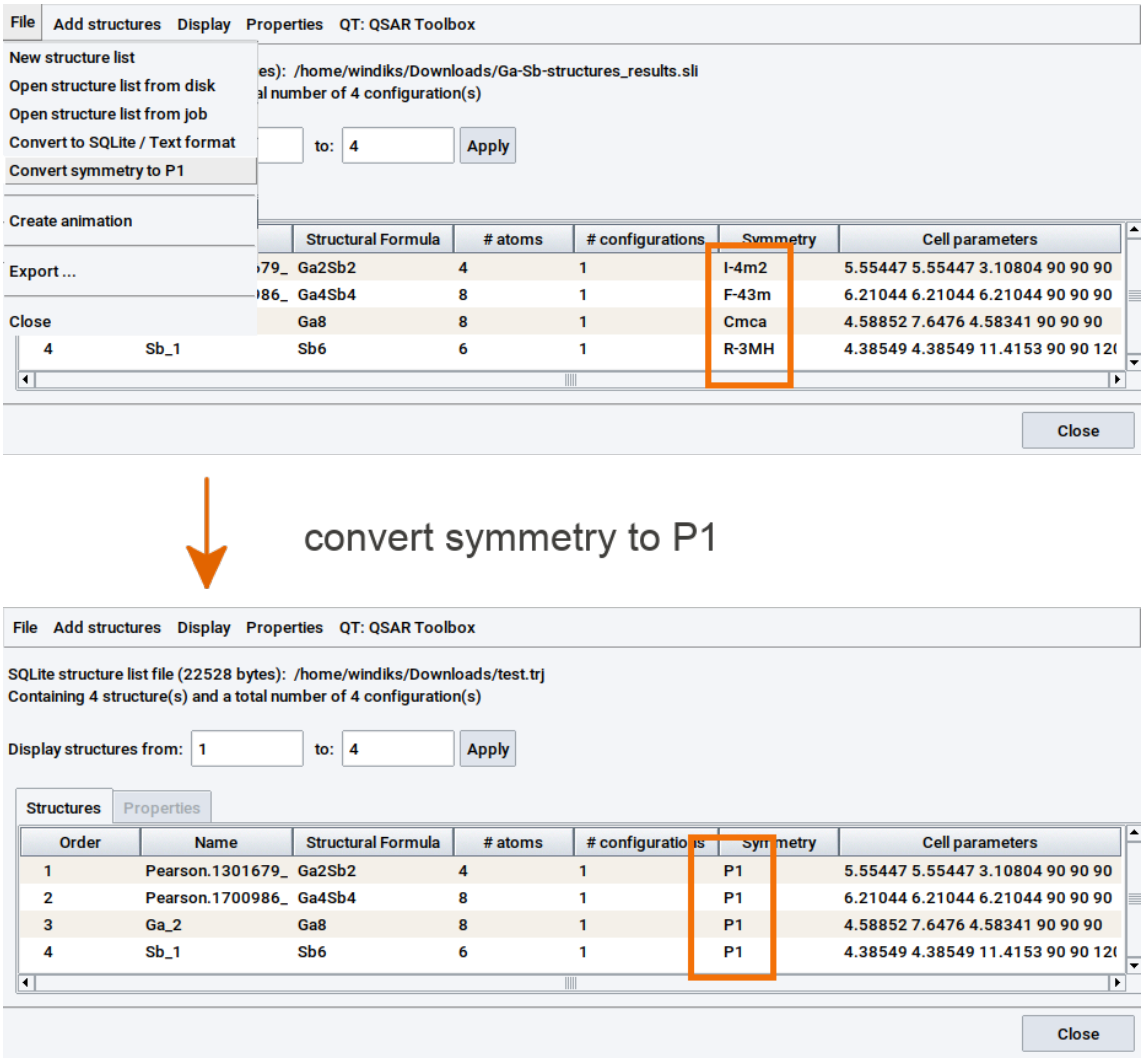
- New structure list: Create a new and empty structure list that can be filled with the features of the menu item Add structures; the file that stores the structure list should have one of the following extensions: *sli*, *slst*, or *list*

- Open structure list from disk : load a file that contains the data of a *MedeA* structure list
- Open structure list from job : load a structure list from the currently selected JobServer; the latter is defined via Jobs >> Select server
- Convert to SQLite / Text format : transform the format of a currently opened structure list from the SQLite into a clear text format and vice versa

---

Hint: By default *MedeA* structure list are stored in the database format SQLite which has the advantage of creating small files, especially, in case of structure lists with several thousand records of data and properties.

- Convert symmetry to P1 : Lower the symmetry of all structures of the structure list to the space group *P1*, i.e. remove all point group symmetry operations. With this feature the periodic boundary conditions (PBC) of structures with simulation cells are maintained.

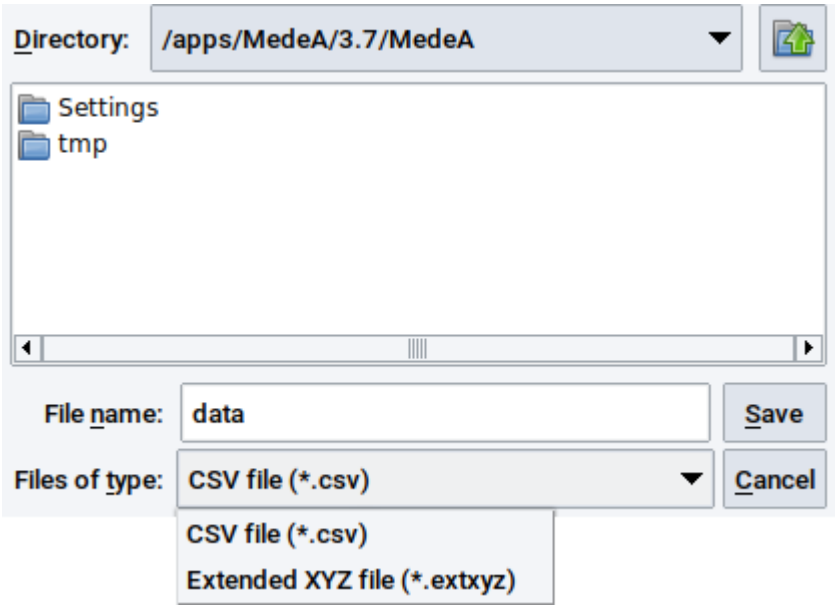


Note: This feature works if the structure list encompasses only the properties that are depicted in the above figure: Order , Name , Structural Formula , # atoms , # configurations , Symmetry , and Cell parameters

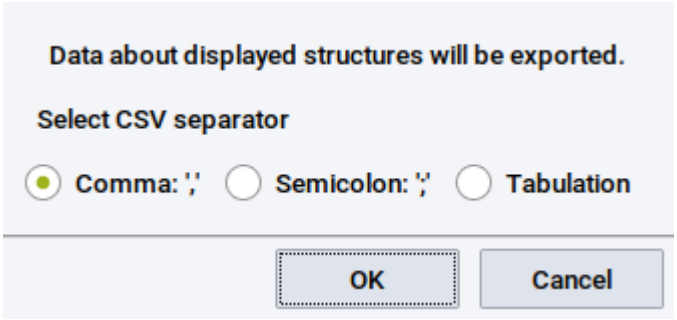
---

- Create Animation : Animates the structures of the currently open list in an extra animation windows (requires that a stucure list is opened and that the structure list contains at least two records)

- Export... : Write all the properties that are displayed in the table of the structure list to data file with the extension *csv* or the structure data in the *extended XYZ* format (extxyz).

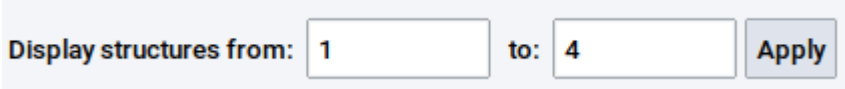


In case you want to save the properties of the table as a data, then first define the filename and then how the values are separated in the data file. The values in the data file can be either separated by commas, semicolons, or tabulators:



- Close : Close the structure list and also the structure list editor. A new instance of the structure list editor that is opened via File >> Structure List Editor does not contain any structure list.

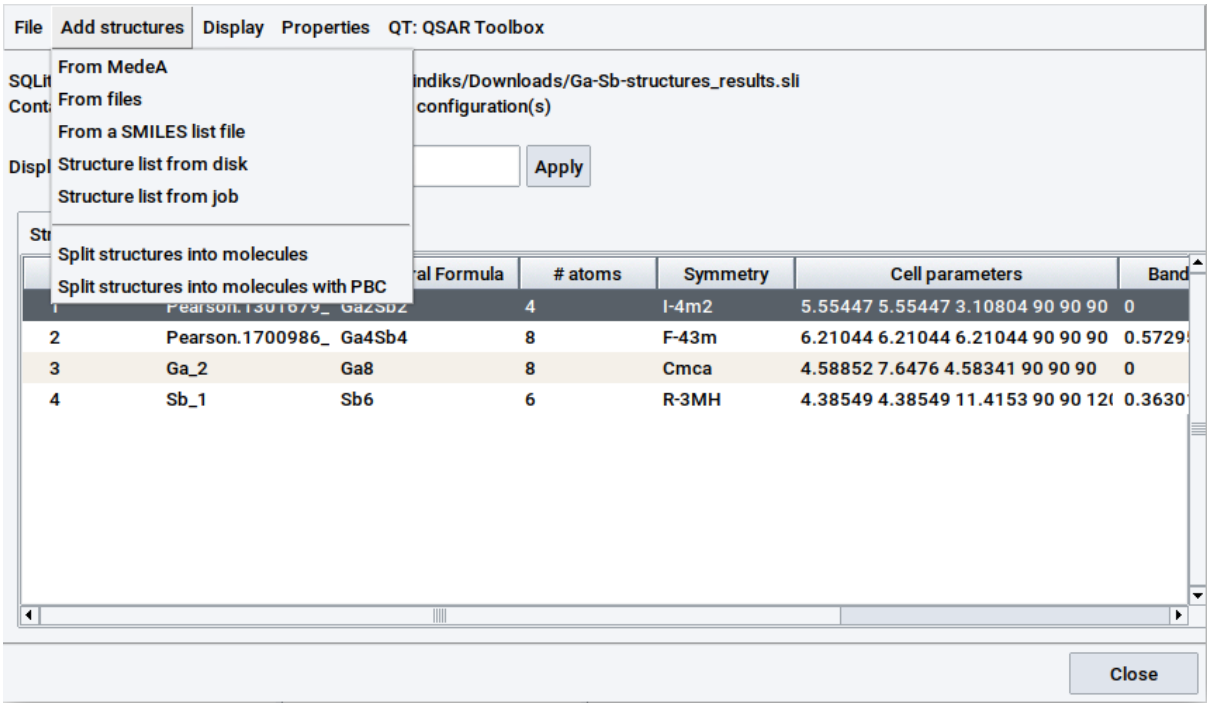
Note: The Close button in the lower right corner of the structure list editor has the same effect as the menu item.

With the options bar 

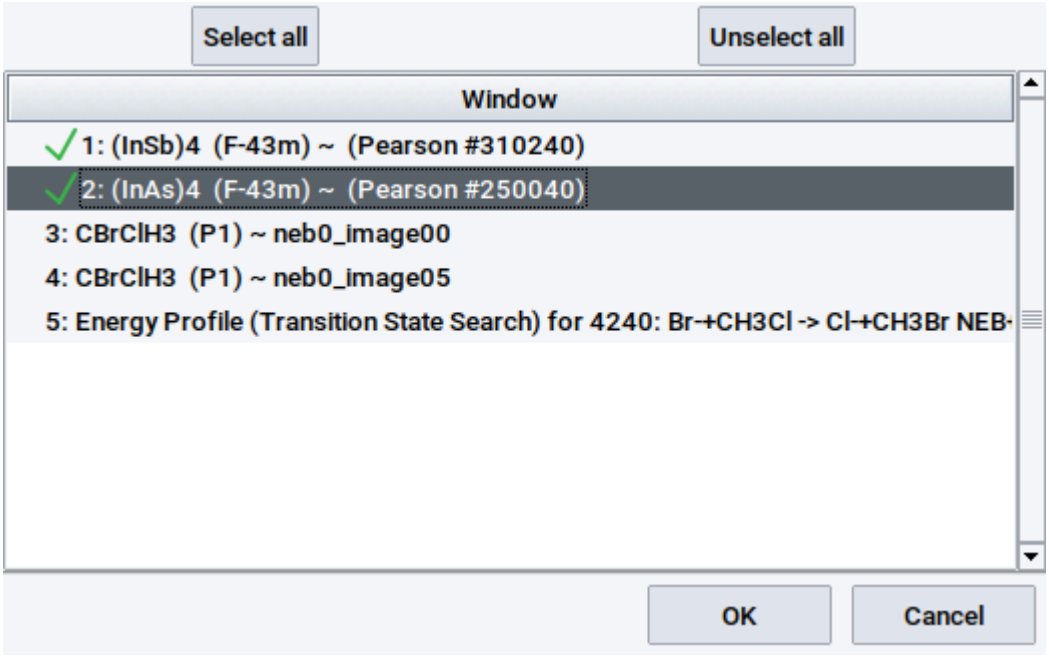
above the two tabs Structures and Properties you can define how many structure records of the list are displayed in the structure list editor. After you defined the range with values for from: and to: confirm the range with Apply .

3 Add Structures

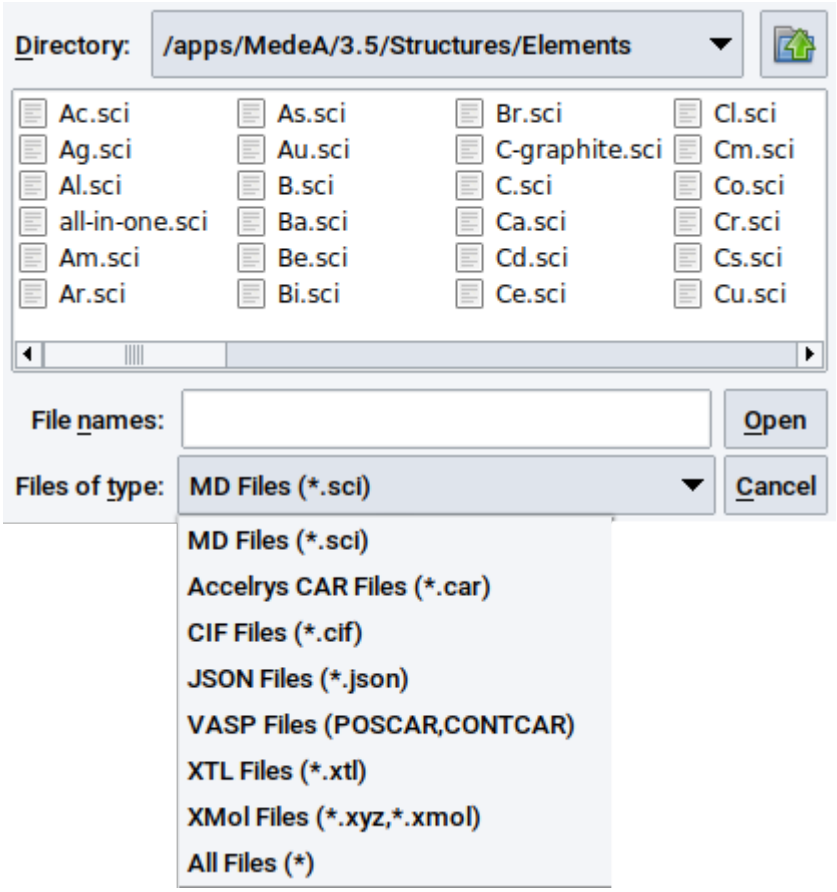
You can add structures via the menu item Add structures once a structure list was created or opened with the File menu.



- From MedeA : Add structures that are displayed in the structure windows of the *MedeA GUI*



- From files : Add structures that are stored in structure data files; supported structure file formats are

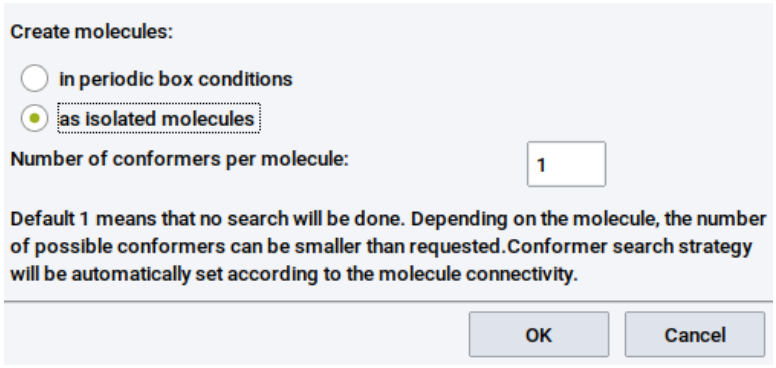


- From a SMILES list file :

  Use a list of SMILES strings of a file to add molecular structures. The file should contain on SMILES string per line, eventually preceded by a name, e.g.

```
`Alanine  CC(N)C(=O)O`
`Caffeine O=C1C2=C(N=CN2C)N(C(=O)N1C)C`
`Valine  CC(C)C(N)C(=O)O`
`Leucine  CC(C)CC(N)C(=O)O`
`Tyrosine NC(Cc1ccc(O)cc1)C(=O)O`
```

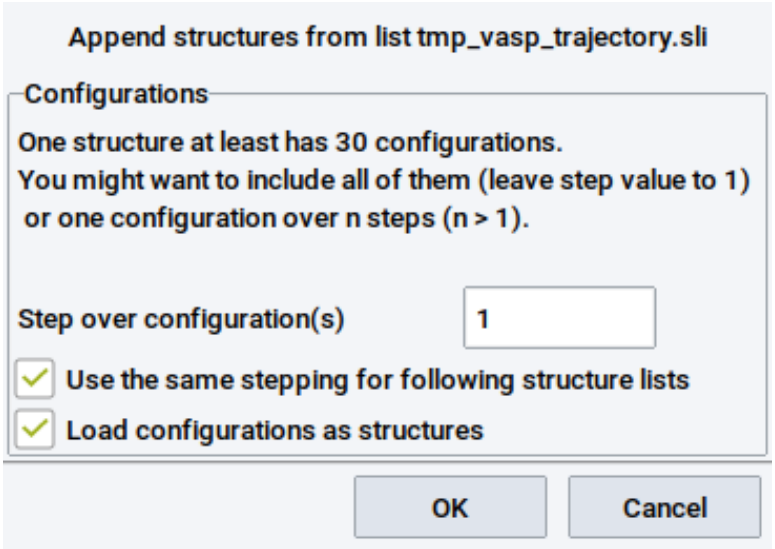
  You can later rename and delete structures from the list and also export all or a selection to individual supported structure data files (1 per structure).



  With the above dialogue define whether each structure should be an isolated structure (i.e. a molecule in the gas phase) or present in a simulation cell with periodic boundary conditions. Since each structure can have several conformers you can also define how many conformers should be created from each SMILES string.

- Structure list from disk : Add entire *MedeA* structure lists that are stored on disk

- Structure list from job : Add entire *MedeA* structure lists that are stored on a selected JobServer. ; the latter is defined via Jobs >> Select server

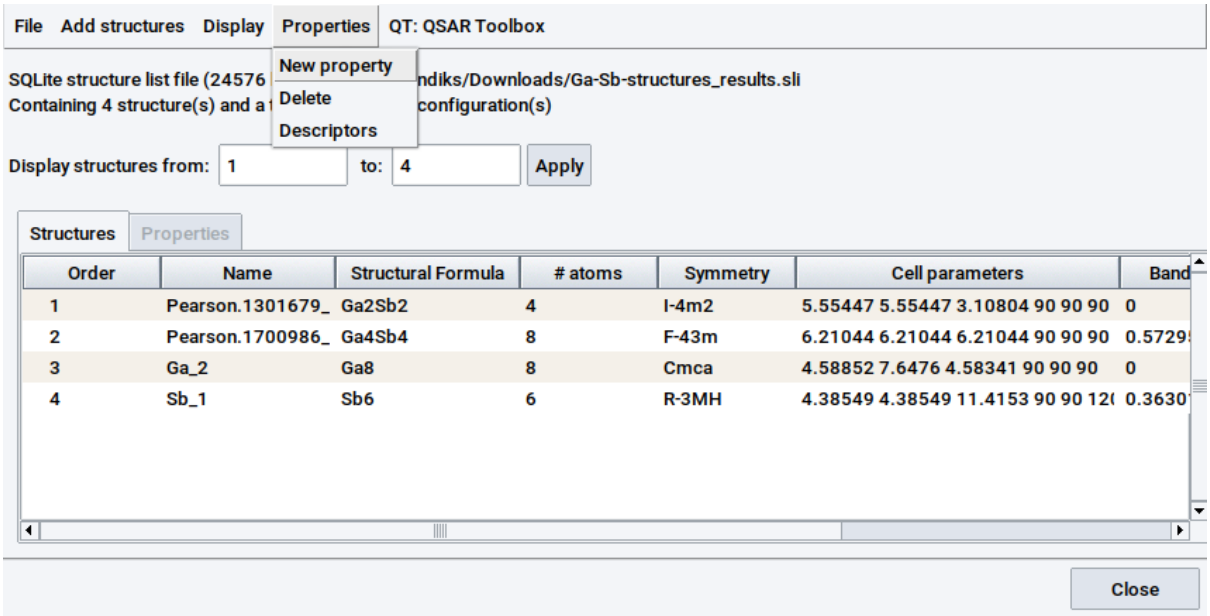


With the above dialogue define if you want to consider each record, every second record, every third record, etc. of the structure list (Step over configuration(s)) and whether this stepping should be applied to all imported structure lists, in case you want to import more than one structure list. In case a structure list is a trajectory containing several configurations you can decide whether for each configuration a separate record is created in the structure list.

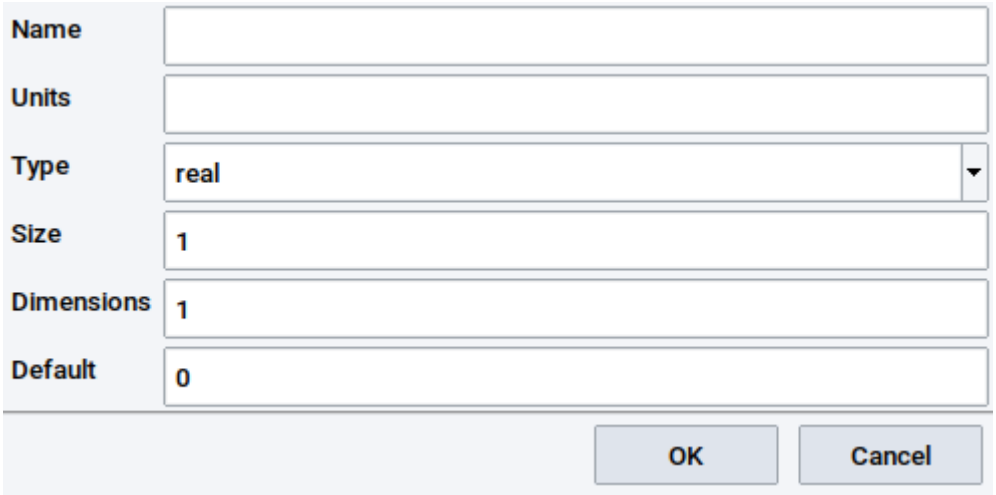
- Split structures into molecules : Add structures that are displayed in the structure windows of the *MedeA GUI*, but split them up in their individual constituents as isolated atoms or molecules, i.e. without periodic boundary conditions (PBC)
- Split structures into molecules with PBC : Add structures that are displayed in the structure windows of the *MedeA GUI*, but split them up into their individual constituents in simulation cells with periodic boundary conditions (PBC)

4 Add and Display Properties

Beside the structural information you can also add properties and structural descriptors to each record of a structure list. One possibility for you is to add properties and their values to an open structure list via the Properties menu item:



- New property : Define a new column in the structure list that you can fill with values. With the menu item you bring up the below dialogue:

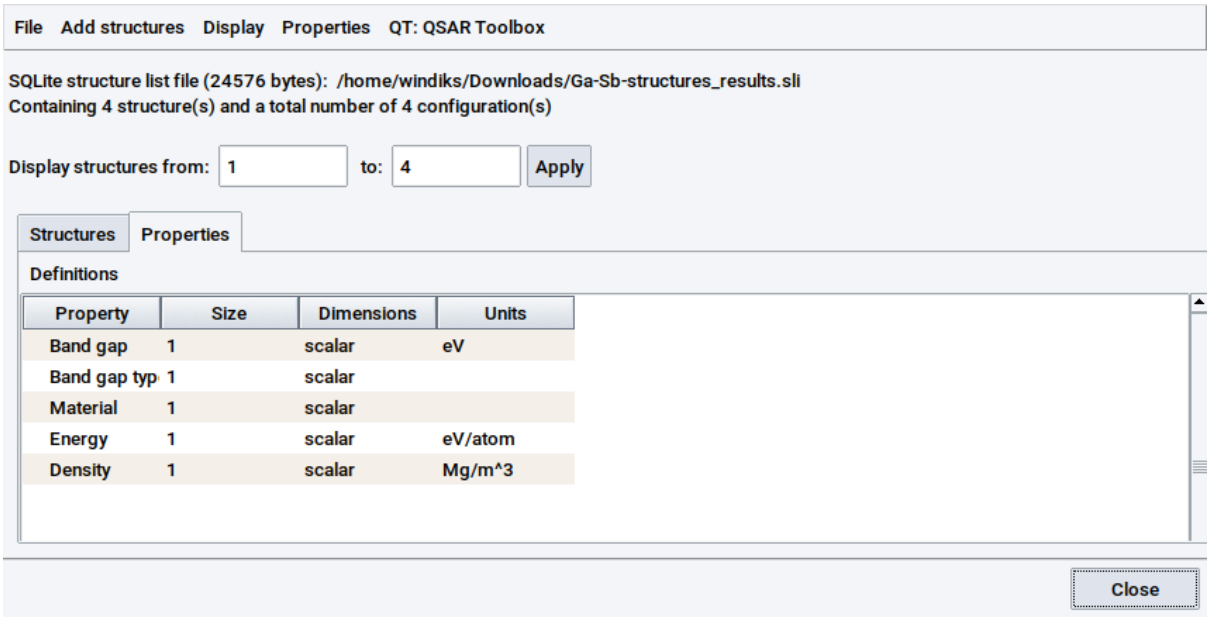


- Name : Enter a name to appear in the header cell of the column
- Units (optional) : The unit of the property
- Type : Select from the selection bar real for real numbers (default), integer for integer numbers, and string for text
- Size : Define the amount of values for this property; can be 1 (default) or larger
- Dimensions the values dimension, e.g. 3 3 means a 3×3 matrix
- Default (optional): The default value to initialize the newly created property column in the structure list; 0 is a reasonable initial value and can be modified afterwards by a double-click in a cell of the table

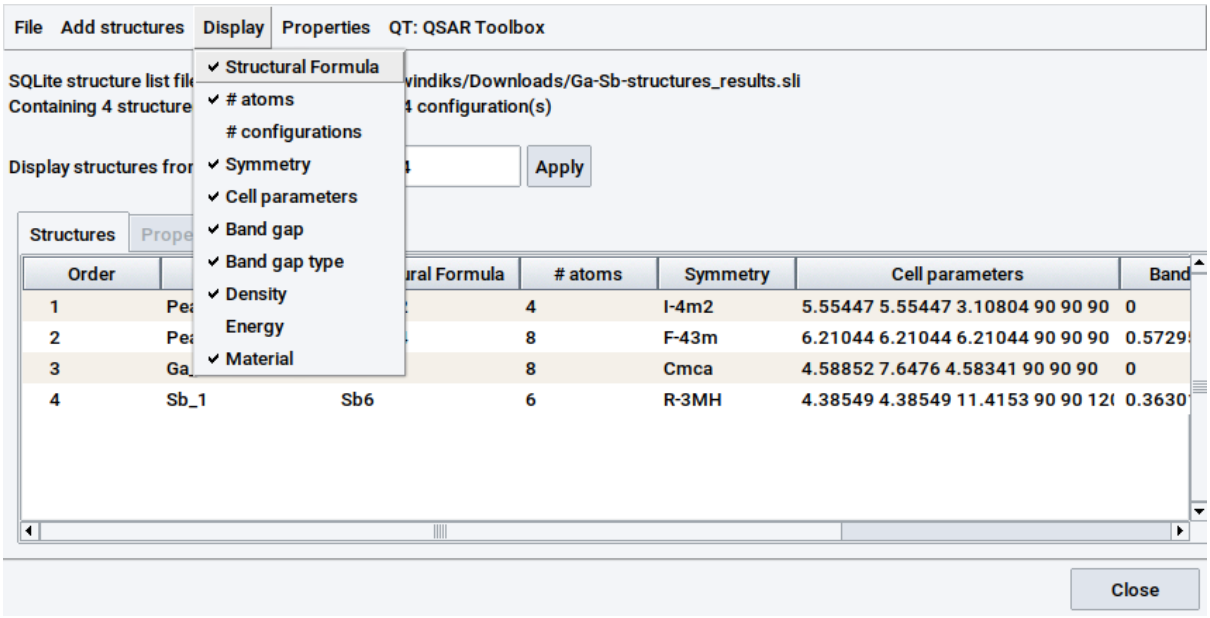
With the menu sequence Properties >> Descriptors you add descriptors that are calculated with other properties and/or structural descriptors based on a user-defined equation or correlation. For more information, read the Section HT-Descriptors in which is part of the description of MedeA HT: High Value from High Throughput Simulations.

Hint: This feature requires a separate license for the *MedeA* module *HT-Descriptor*.

Once you have added properties then the Properties tab gets accessible. This tab summarizes the name, size, dimension, and unit of each property.

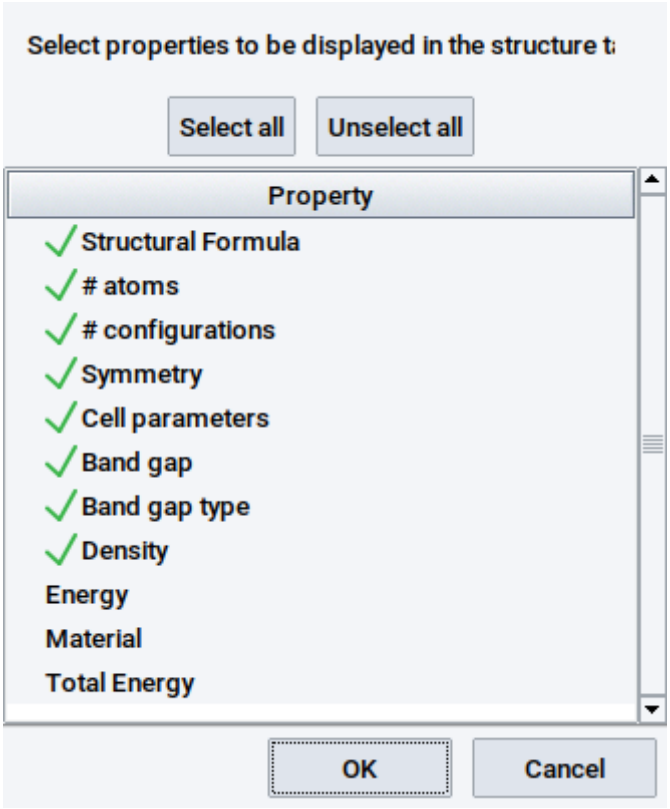


To determine which of the columns of a structure list should be shown use the Display menu item:



A click on a menu item has immediate effect, i.e. either display or un-display a column with a property.

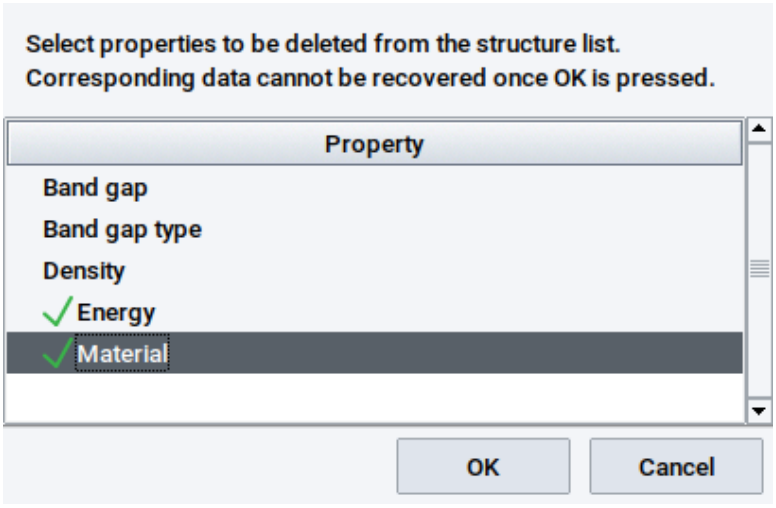
In case a structure list contains more than 10 columns then you can use the menu sequence Display >> Select displayed properties. That brings up the below dialogue to decide which properties should be shown in the structure list editor.



Click on a row to either add or remove a green check mark; only marked properties are displayed. Confirm your selection with OK. Click on Select all or Unselect all to show all or none, respectively, of the columns.

5 Remove Properties

To delete an entire column with property cells then use the menu sequence Properties >> Delete. In the below dialogue click on a row to either add or remove a green check mark; only marked properties are deleted.



Confirm your selection with OK.

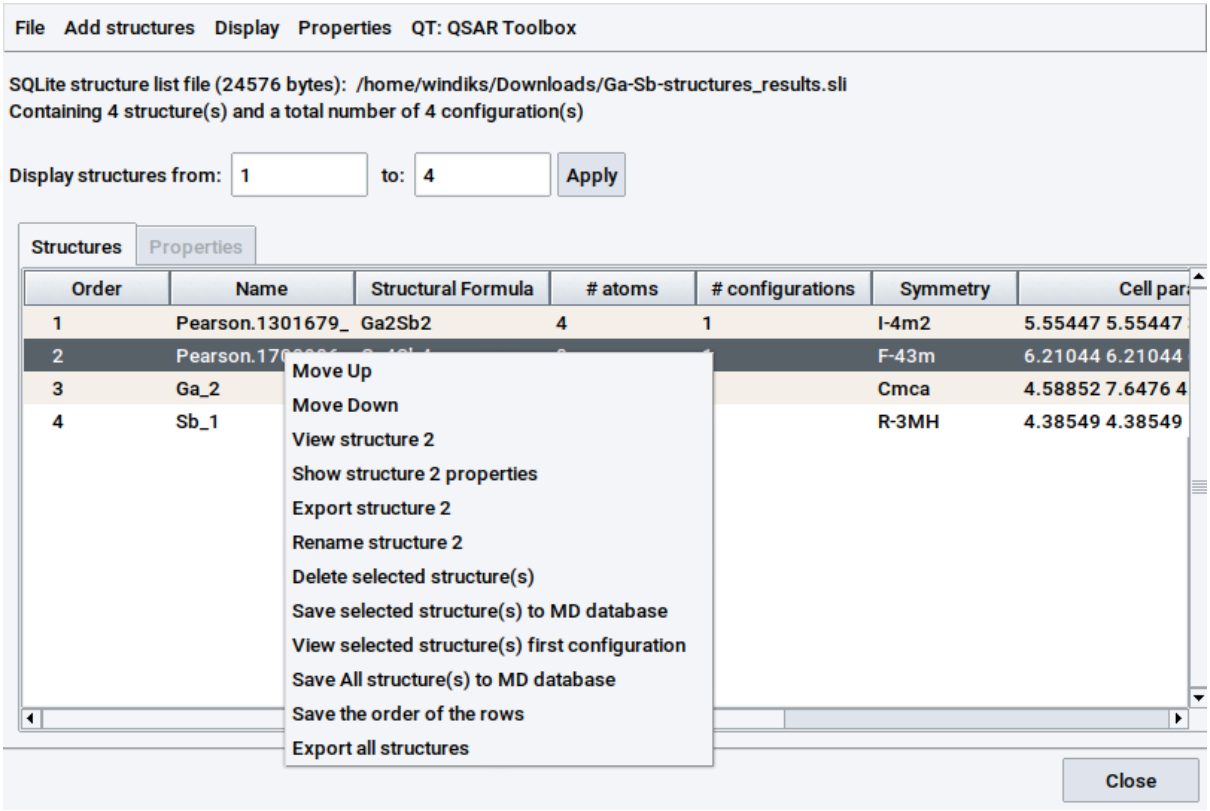
6 Modify List Records

With the context menu of the structure list editor you can modify and manage individual list records or several selected list records. The context menu appears upon right-clicking on a structure list record:

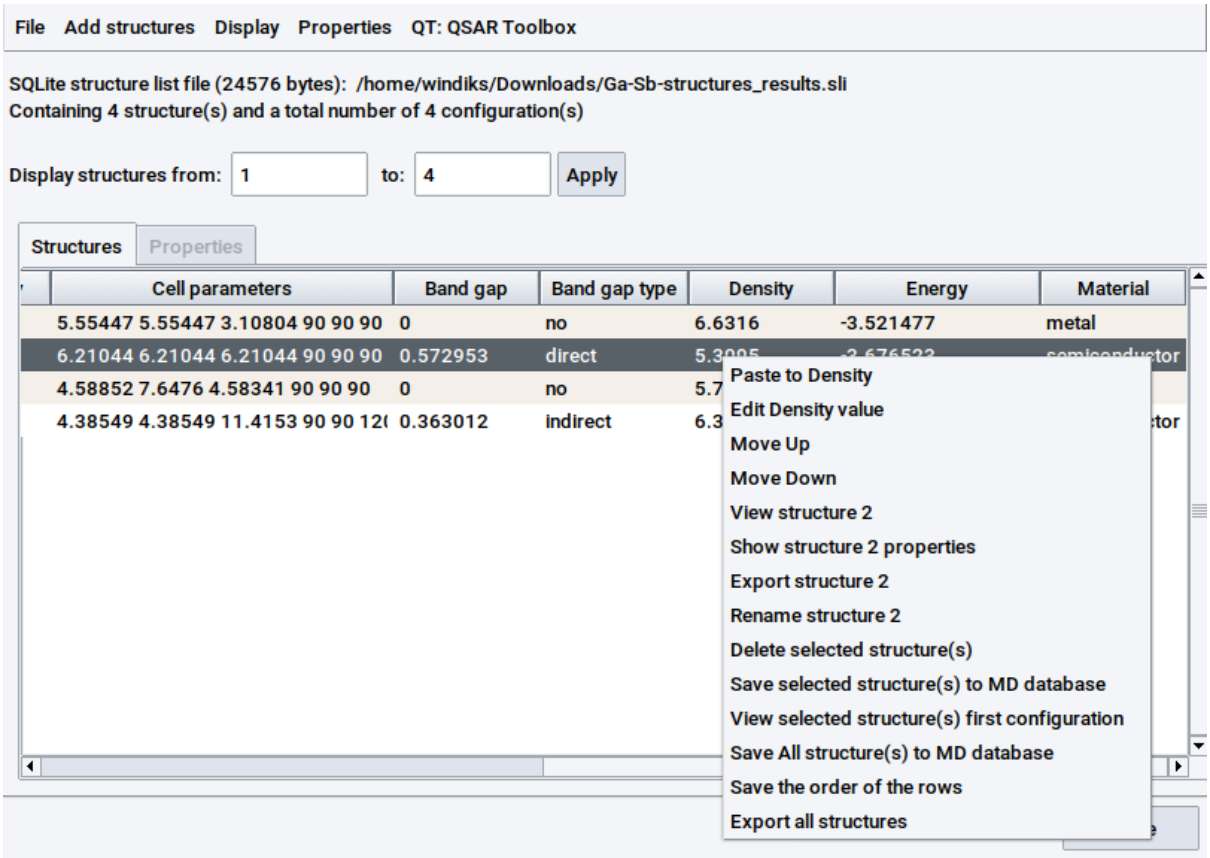
A right-click on one of the data cells (not the header cell!)

- ID
- Name
- Structural Formula
- # atoms
- # configurations
- Symmetry
- Cell parameters

brings up the context menu:

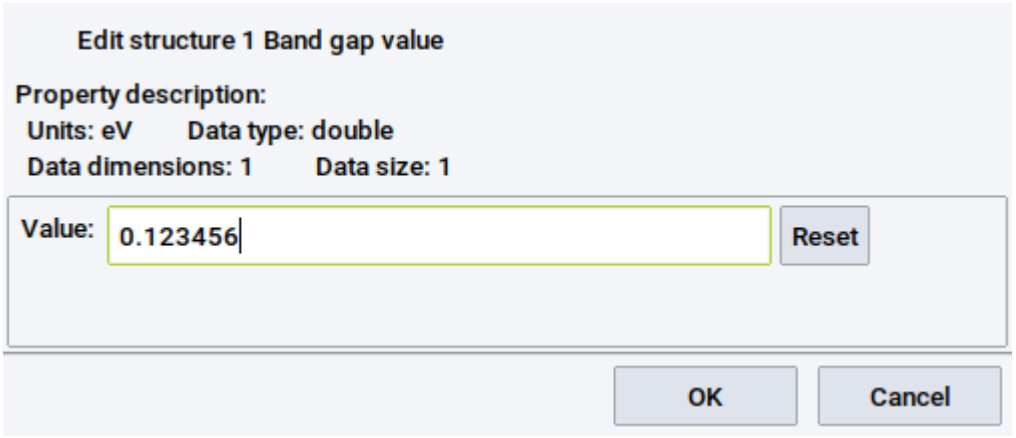


A right-click on one of the other cells brings up a context menu with two more items:



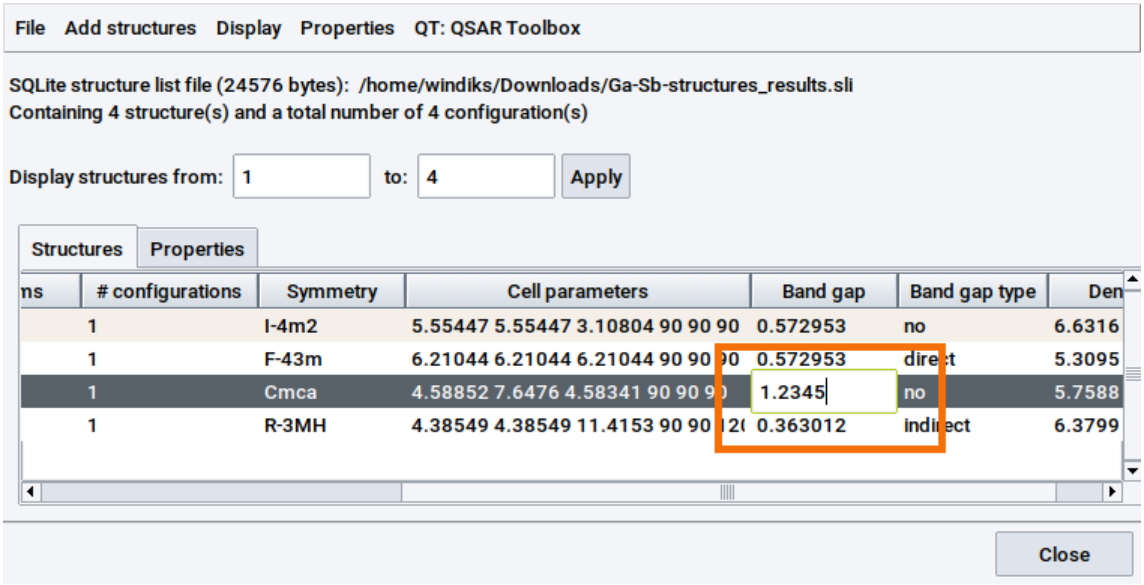
The individual items of the context menus have the following functions:

- Paste to property : either enter or replace a value of a property with value stored in the clipboard
- Edit property value : modify the value of a cell using the dialogue



Confirm the modification with OK

Note: Another possibility to modify a value of a property is to double-click in a data cell which makes the content of the cell editable:



To confirm the change simply press the Enter key of your keyboard.

- Move Down : move a row one position **down** (to make this modification permanent use the menu item Save the order of the rows of this context menu)
- Move Up : move a row one position **up** (to make this modification permanent use the menu item Save the order of the rows of this context menu)
- View structure : visualize the structure in the *MedeA GUI*
- Show structure properties : open the Properties tab
- Export structure : save this structure to a file on disk in one of the supported formats
- Rename structure : change the name of the structure
- Delete selected structure(s) : remove the record from the structure list

Hint: Once you click OK in the confirmation window then a structure is permanently removed from the list, together with the associated properties. You do not have any possibility to undo this option!

- Save selected structure(s) to MD database : create new entries in the *Materials Design Database*, or briefly *MD database* for computed results; this database is searchable with *InfoMaticA*
- View selcted structure(s) first configuration : visualize only the very first configuration of all selected records in the *MedeA GUI*; this works also if a structure list record comprises one configuration
- Save all structure(s) to MD database : create new entries in the *Materials Design Database*, or briefly *MD database* with all structures of the current structure list; the *Materials Design Database* is for computed results and searchable with *InfoMaticA*
- Save the order of the rows : make the order of re-arranged rows permanent
- Export all structures : save each structure of the current list to separate files on disk in one of the supported formats

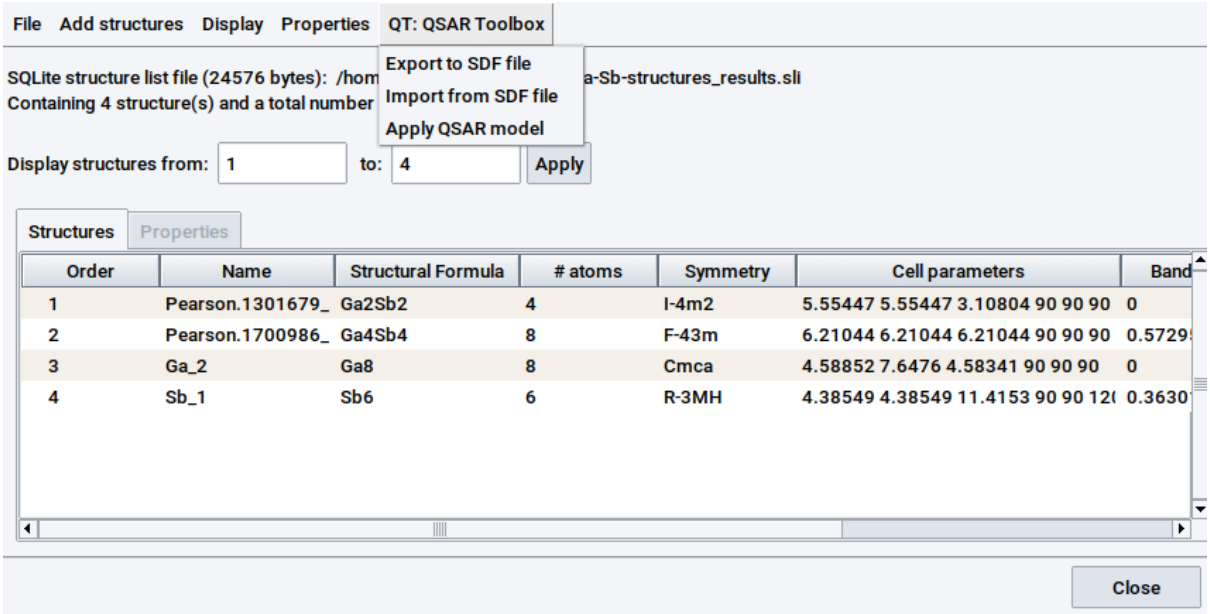
7 Order List Records

Structures are stored in a given order in the list, which appears in the Order column in the structures table. It is possible to change the order in the table by sorting rows (structures) according to a column from the right-click context menu commands Sort Ascending or Sort Descending, on a column title. Also, a given structure can be moved up or down by a right-click on its row with the corresponding command in the context menu.

If the initial order of the rows in the table is changed by one of these methods, the actual structure order in the structure list file remains unchanged. It is possible to apply the new displayed order to reorder the list internal ordering with the right-click context menu command Save the order of the rows. This can be useful, for example, to change the order of how structures should be processed in a *MedeA* job.

8 Predict Properties From Correlations

The structure list editor can be used in conjunction with *MedeA QT* to explore and analyze the relationships between descriptors and system properties. In case you have a license for *MedeA QT* then you also see the menu item QT: QSAR Toolbox. A click on this menu item opens a pull-down menu that has the items



- Export to SDF file: Writes to the properties and descriptors to a data file in the *sdf* format, which is readable by *MedeA QT*
- Import from SDF file: Reads properties and descriptors from a data file in the *sdf* format
- Apply QSAR model: Reads properties and descriptors from a data file in the *xml* format that was created with *MedeA QT*