Covalent Docking

Schrödinger Software Release 2015-2

Schrödinger Press

May 2015

Document Conventions

In addition to the use of italics for names of documents, the font conventions that are used in this document are summarized in the table below.

Font	Example	Use
Sans serif	Project Table	Names of GUI features, such as panels, menus, menu items, buttons, and labels
Monospace	`$SCHRODINGER/maestro`	File names, directory names, commands, environment variables, command input and output
Italic	*filename*	Text that the user must replace with a value
Sans serif uppercase	CTRL+H	Keyboard keys

Links to other locations in the current document or to other PDF documents are colored like this: Document Conventions.

In descriptions of command syntax, the following UNIX conventions are used: braces { } enclose a choice of required items, square brackets [ ] enclose optional items, and the bar symbol | separates items in a list from which one item must be chosen. Lines of command syntax that wrap should be interpreted as a single command.

File name, path, and environment variable syntax is generally given with the UNIX conventions. To obtain the Windows conventions, replace the forward slash / with the backslash \ in path or directory names, and replace the $ at the beginning of an environment variable with a % at each end. For example, `$SCHRODINGER/maestro` becomes `%SCHRODINGER%\maestro`.

Keyboard references are given in the Windows convention by default, with Mac equivalents in parentheses, for example CTRL+H (⌘H). Where Mac equivalents are not given, COMMAND should be read in place of CTRL. The convention CTRL-H is not used.

In this document, to *type* text means to type the required text in the specified location, and to *enter* text means to type the required text, then press the ENTER key.

References to literature sources are given in square brackets, like this: [10].

Docking Covalently Bound Ligands

Docking of ligands that are bound to a receptor through hydrogen bonds or various other nonbonded interactions is relatively straightforward, and can be performed with high efficiency by Glide. When ligands bind covalently to the receptor, it is not as easy to screen ligands by docking, because the chemistry of the reaction of the ligand with the receptor must be accounted for. Covalent binding is usually considered to consist of an initial association, in which the ligand associates noncovalently with the protein in a pose that facilitates the reaction, followed by the reaction. The reaction is fast, so the binding is controlled by the access of the ligand to the binding site in a suitable pose.

The Covalent Docking protocol addresses the process as follows. First, it determines whether ligands can associate with the receptor in a suitable pose. This is done by mutating the reactive residue on the receptor to alanine, so that the pose of the side chain does not unduly influence the association of the ligand, then docking the ligands with Glide, with constraints between the reactive residue and the reactive group on the ligand. Once suitable poses are found, the receptor is restored, and the poses of the side chain on the reactive residue are explored in the presence of the associated ligand, to find the best poses for reaction. The covalent bond is formed, and the ligand and reactive residue are minimized to relieve strain. The poses for a given ligand are clustered, and a representative pose is chosen from each cluster. These poses undergo a full minimization, and the representatives are ranked by their Prime energy.

As the energy is not suitable for comparing different ligands, an affinity score is calculated, which measures the affinity of the ligand to the receptor for noncovalent binding prior to reaction. Ligands that do not score well are considered less likely to approach the receptor in a way that permits covalent bond formation. To calculate the score, the bond is broken again, the reactive receptor residue is mutated to alanine, and the bond to the ligand is capped with hydrogen. Scoring is then done in place with Glide, and the affinity is reported as the average of the pre-reacted and post-reacted GlideScore for a given pose.

Covalent docking involves a reaction, which must be accounted for by the protocol, so that the reactive functional group on the ligand and the reactive residue are identified, and the bond formed between the correct atoms on each. The necessary information for several common reaction types are included with the protocol.

The default Covalent Docking protocol is time-consuming, so it is not suitable for screening large numbers of ligands. Typically it takes 1–2 hours per ligand. A “virtual screening” protocol is also available, which skips time-consuming steps such as rotamer sampling and

minimization. This protocol is at least 10 times faster than the default ("pose prediction") protocol, and is suitable for screening thousands of ligands (with a limit of 10 000 in any run).

1 Running Schrödinger Software

Schrödinger applications can be run from a graphical interface or from the command line. The software writes input and output files to a directory (folder) which is termed the *working directory*. If you run applications from the command line, the directory from which you run the application is the working directory for the job.

Linux:

To run any Schrödinger program on a Linux platform, or start a Schrödinger job on a remote host from a Linux platform, you must first set the `SCHRODINGER` environment variable to the installation directory for your Schrödinger software. To set this variable, enter the following command at a shell prompt:

csh/tcsh: `setenv SCHRODINGER` *installation-directory*

bash/ksh: `export SCHRODINGER=`*installation-directory*

Once you have set the `SCHRODINGER` environment variable, you can run programs and utilities with the following commands:

`$SCHRODINGER/`*program* `&`
`$SCHRODINGER/utilities/`*utility* `&`

You can start the Maestro interface with the following command:

```
$SCHRODINGER/maestro &
```

It is usually a good idea to change to the desired working directory before starting the Maestro interface. This directory then becomes the working directory.

Windows:

The primary way of running Schrödinger applications on a Windows platform is from a graphical interface. To start the Maestro interface, double-click on the Maestro icon, on a Maestro project, or on a structure file; or choose Start → All Programs → Schrodinger-2015-2 → Maestro. You do not need to make any settings before starting Maestro or running programs. The default working directory is the Schrodinger folder in your Documents folder.

If you want to run applications from the command line, you can do so in one of the shells that are provided with the installation and have the Schrödinger environment set up:

- Schrödinger Command Prompt—DOS shell.
- Schrödinger Power Shell—Windows Power Shell (if available).

You can open these shells from Start → All Programs → Schrodinger-2015-2. You do not need to include the path to a program or utility when you type the command to run it. If you want access to Unix-style utilities (such as `awk`, `grep`, and `sed`), preface the commands with `sh`, or type `sh` in either of these shells to start a Unix-style shell.

Mac:

The primary way of running Schrödinger software on a Mac is from a graphical interface. To start the Maestro interface, click its icon on the dock. If there is no Maestro icon on the dock, you can put one there by dragging it from the SchrodingerSuite2015-2 folder in your Applications folder. This folder contains icons for all the available interfaces. The default working directory is the Schrodinger folder in your Documents folder (`$HOME/Documents/Schrodinger`).

Running software from the command line is similar to Linux—open a terminal window and run the program. You can also start Maestro from the command line in the same way as on Linux. The default working directory is then the directory from which you start Maestro. You do not need to set the `SCHRODINGER` environment variable, as this is set in your default environment on installation. To set other variables, on OS X 10.7 use the command

`defaults write ~/.MacOSX/environment` *variable* `"`*value*`"`

and on OS X 10.8, 10.9, and 10.14 use the command

`launchctl setenv` *variable* `"`*value*`"`

2 Preparing the Structures

The receptor structure and the ligand structures should be prepared for use with Glide and Prime. The recommended procedure is to use the Protein Preparation Wizard for preparing the receptor, and LigPrep for preparing the ligands. Details of these tools and their use can be found in the *Protein Preparation Guide* and the *LigPrep User Manual*. An overview of the preparation needed for Glide is given in Chapter 3 of the *Glide User Manual*.

The ligand structures must all contain the reactive functional group for the reaction by which the covalent bond is formed. This functional group is defined by a SMARTS pattern, which is displayed in the Covalent Docking panel. The default SMARTS patterns are listed in Table 1, below. You can filter the ligands in the Ligand Filtering panel, as follows.

1. Choose Tools → Ligand Filtering to open the Ligand Filtering panel.

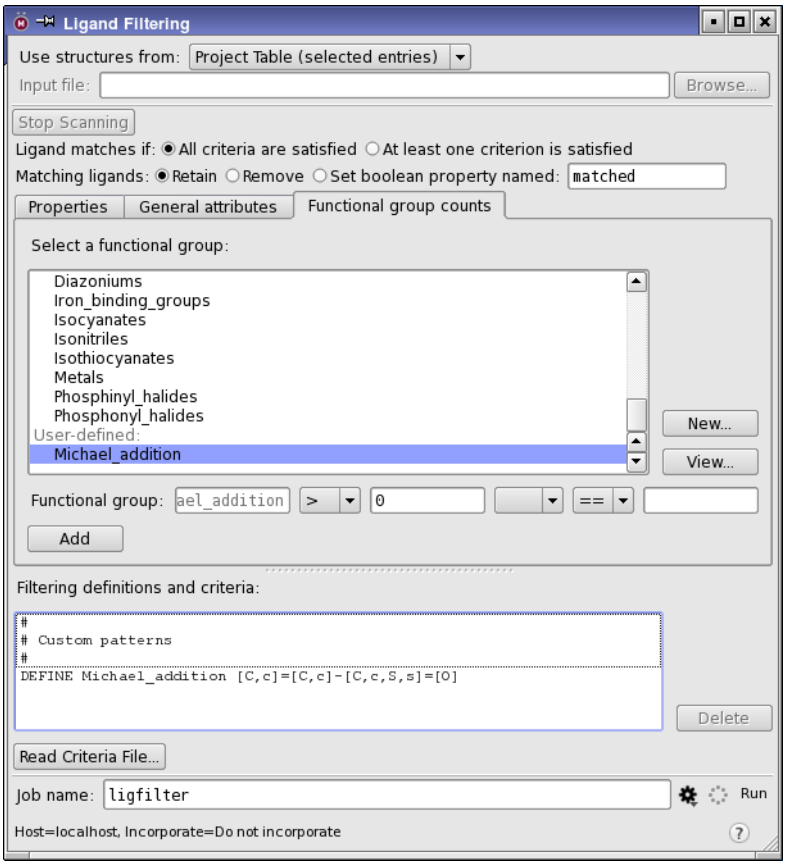

Figure 1. The Ligand Filtering panel.

2. Select the source of ligands from the Use structures from option menu.
   - If you choose File, click Browse and select the file in the file selector that opens.
   - If you choose Project Table, select the ligands in the Project Table.
3. For Matching ligands, ensure that Retain is selected.

   It doesn't matter what is selected for Ligand matches if, as you will only use one criterion.
4. In the Functional Group Counts tab, click New.

   The Ligand Filtering - Add Definition dialog box opens. As the functional group for the reaction is not likely to be included in the predefined list, you will have to add the definition for this functional group.

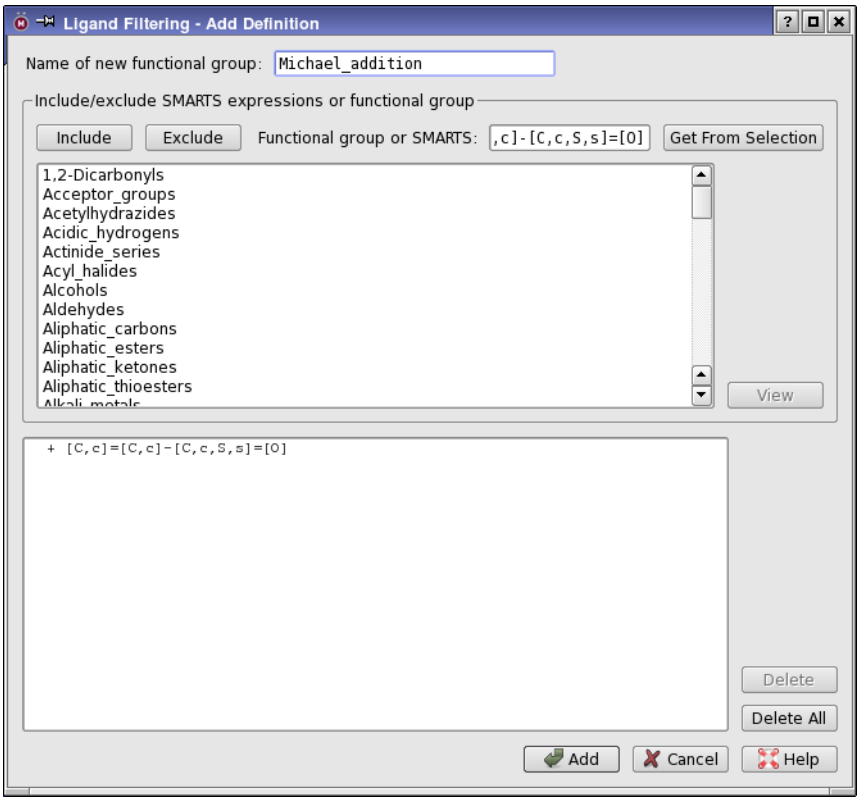

Figure 2. The Add Definition dialog box.

5. Enter a name for the functional group in the Name of new functional group text box.

   For example, use the reaction type. There must be no spaces in the name.

6. Copy or type the SMARTS pattern for the ligand reactive group into the SMARTS expression or selected functional group text box.

   If you intend to use the default pattern, you can copy it from Table 1. If you intend to modify the default pattern, to make it more specific, you should instead use the Get From Selection button, as this is the same procedure as is used in the Covalent Docking panel. Display a ligand in the Workspace, select the atoms that you want to include in the SMARTS pattern, then click Get From Selection.

7. Click Include.

   The SMARTS pattern is listed in the lower text area. You can add more SMARTS patterns to the group by repeating the last two steps. You might want to do this if you intend to modify the default pattern.

8. Click Add.

   The functional group is added under the User-defined heading, at the bottom of the functional group list.

9. Select the functional group you just created.

   The name is displayed in the Functional group text box.

10. Choose > from the option menu to the right, then enter 0 in the text box.

11. Click Add.

    A new line is added in the Filtering definitions and criteria. This line shows that the number of occurrences of the functional group in any ligand must be greater than zero for the ligand to be retained in the output.

12. If you want to add the filtered ligand set to the Project Table, click the Job Settings (gear) button, and choose Append new entries from the option menu in the Output section.

13. Enter a job name in the Job name text box, and click Run.

    The job name is used to name the output file.

You can also run ligand filtering from the command line—see Section 2.4 of the *General Utilities* manual.

3 Running Covalent Docking from Maestro

Covalent docking calculations can be set up and run from the Covalent Docking panel. To open this panel, choose Tasks → Docking → Covalent Docking or Applications → Glide → Covalent Docking in the main window.

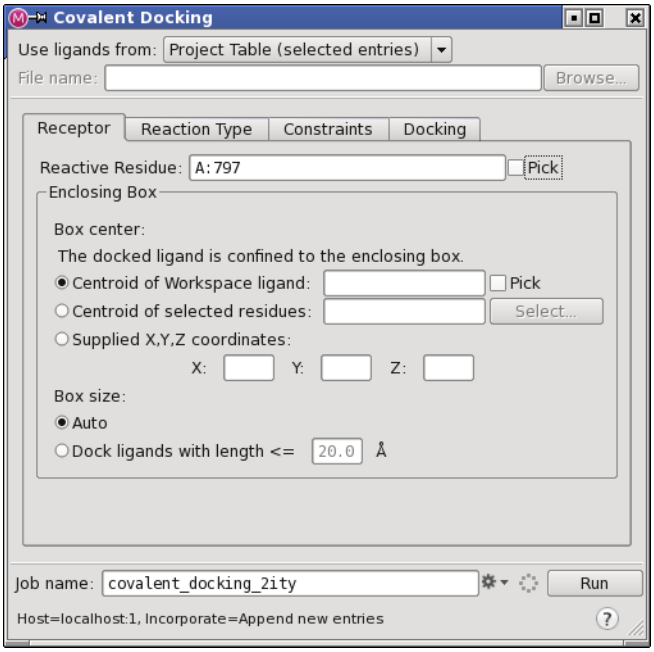

Figure 3. The Covalent Docking panel.

3.1 Selecting the Ligands

The ligands that you dock to the receptor can be taken from the Project Table or from a file.

- To use the structures that are selected in the Project Table, choose Project Table from the Use ligands from option menu.
- To read the ligands from a file, choose File from the Use ligands from option menu. You can then either enter the file name in the File text box, or click Browse and navigate to the file in the file selector that opens. The file must be a Maestro file, and can be compressed (`.maegz`, `.mae.gz`) or uncompressed (`.mae`).

The structures must be all-atom, 3D structures that are properly prepared, for example by using LigPrep. See the *LigPrep User Manual* for more information on ligand preparation.

The poses that are returned include both the ligand and the receptor, and the job can take hours per ligand, so you should consider carefully how many ligands to dock.

3.2 Specifying the Receptor

In the Receptor tab you specify the reactive residue (or group) in the receptor and set options for Glide docking. The receptor must be displayed in the Workspace, and is written to a file when you start the job.

To define the reactive residue, select Pick (to the right of the Reactive residue text box), and pick the residue in the Workspace. The residue type must be one of those listed in Table 1 or defined in a custom chemistry, otherwise a warning is posted and you will have to pick a different residue. When you have picked a residue, the chain name and residue number are shown in the Reactive residue text box, separated by a colon, and the residue is highlighted in the Workspace. You can also enter the chain name and residue number directly in the text box.

For the Glide docking stage, the specification of a grid box is required that is approximately centered on the active site. There are two options for setting the center of the docking grid box.

- Centroid of Workspace ligand—This option centers the grid at the centroid of the ligand molecule that is displayed in the Workspace. You must display a ligand in the Workspace as well as the receptor to use this option. Select Pick, then pick a ligand atom in the Workspace.
- Centroid of selected residues—Center the grid at the centroid of a set of selected residues. This allows you to define the active site (where grids should be centered) with only the receptor in the Workspace. When this option is chosen, the Select button becomes available. This button opens the Atom Selection dialog box, so you can select the residues. You can also enter an ASL expression directly in the text box.

  If you use this option, you should not simply select the reactive residue, as the reactive residue can be on the periphery of the binding site. Instead, you should select residues whose centroid is approximately at the center of the binding site. Alternatively, use the Centroid of Workspace ligand option.
- Supplied X, Y, Z coordinates—This option centers the grid at the Cartesian coordinates that you specify in the X, Y, and Z text boxes. These text boxes are only available when you choose this option.

You can also select options for the size of the grid box.

- Auto—Automatically determine the size of the grid box. If the Box center option is Centroid of Workspace ligand, the grid box size is calculated automatically from the size of the ligand. Otherwise, the grid box size is set to a cube with sides of length 26 Å.

- Dock ligands with length <= *N* Å—Select this option to set the size of the grid box. Enter the desired side length in the text box, in angstroms. The grid box has sides of equal length given by the value in the text box.

3.3 Defining the Reaction Chemistry

After selecting the ligands, you must choose the type of reaction by which the ligands bind to the receptor, which you do in the Reaction Type tab. The reaction defines the functional group in the ligand that reacts, the type of receptor residue that the ligand reacts with, and the atom in the ligand that becomes bonded to the receptor. The ligand functional group is represented by a SMARTS pattern, with an atom index that identifies the atom that forms the bond with the receptor. When the ligands are docked, all patterns are tried for a match to the ligand, and each match is docked, which could include multiple matches on a single ligand.

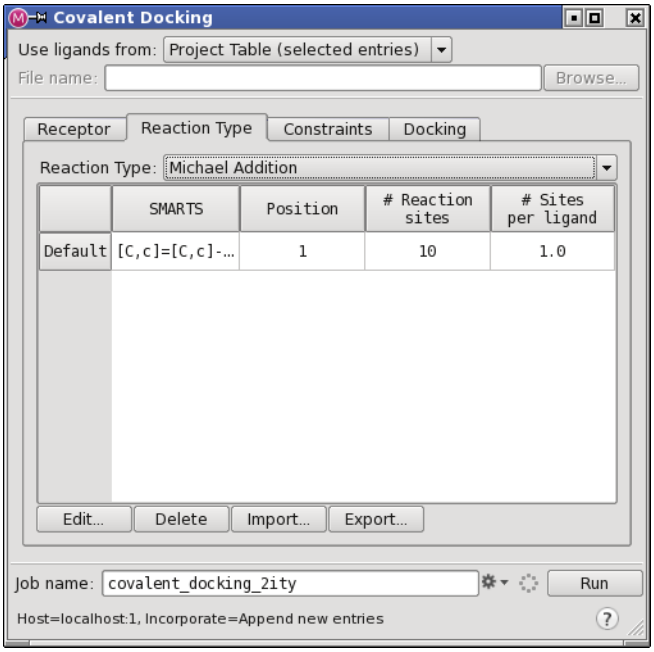

Figure 4. The Reaction Type tab of the Covalent Docking panel.

You can choose from several types of reaction on the Reaction type option menu. When you have selected a reaction type, the reactive functional groups for this reaction are located on the ligands you chose to dock, and a progress bar is displayed for the analysis of the ligands if necessary. The functional group is identified by SMARTS patterns; the default SMARTS pattern for the ligand functional group is shown in the table. If the ligands do not have the functional group, you might want to choose a different ligand set or a different reaction type.

The reaction types, the ligand functional groups, and the allowed residue types are listed in Table 1. The notation R-SH,OH means one of six residues with an SH or OH bond: CYS, SER, THR, ASP, GLU, TYR.

Table 1. Reaction types with ligand and receptor groups involved in each reaction.

Reaction type	Default Ligand SMARTS	Reactive Residues
Michael Addition	[C,c]=[C,c]-[C,c,S,s]=[O]	R-SH,OH
Nucleophilic Addition to a Double Bond	[C,c]=[O,S]	R-SH,OH
Nucleophilic Addition to a Triple Bond	[C]#[N]	R-SH,OH
Nucleophilic Substitution	[*][F,Cl,Br,I]	R-SH,OH
Boronic Acid Addition	[B]([O])[O]	R-SH,OH
Epoxide Opening	[C;r3][O;r3][C;r3]	HIS, R-SH,OH
Imine Condensation	[C](=[O])-[C]	LYS, ASN, GLN, ARG
Phosphonate Addition	[P]-[O;H1,-1]	R-SH,OH
Beta Lactam Addition	[O-0X1]=[C]1[C][C][N]1	CYS, SER
Conjugate Addition to Alkene (nitrile activated)	[C,c]=[C,c]-[C,c]#[N,n]	CYS, SER
Conjugate Addition to Alkyne (carbonyl activated)	[C-0X2]#[C-0X2][C-0X3]=[O-0X1]	CYS, SER
Conjugate Addition to Alkyne (aryl activated)	[C]#[C]-[c]	CYS, SER
Ion Pair to Covalent Bond: Lig(-1)/Rec(+1)	[-1]	LYS, ARG, HIP
Ion Pair to Covalent Bond: Lig(+1)/Rec(-1)	[+1]	ASP, GLU

If the reaction type that you are interested in is not supplied in the default set, you can create a custom reaction type with the instructions in Section 5 on page 21, and save it to a file. This includes defining receptor sites that are not standard residues. Once you have a custom reaction type, you can choose Custom from the Reaction type option menu, and import it using the Custom chemistry file text box and Browse button. The file must have a `.cdock` extension.

The default SMARTS pattern for the ligand is generic and not very specific. You might want to use more specific SMARTS patterns, to apply additional criteria beyond the actual reactive

functional group. Some tools are provided to replace the default pattern and add more patterns. You must edit the default SMARTS pattern first, as this is the most general pattern, and then you can add other patterns.

These patterns are best defined by selecting atoms in a typical ligand, displayed in the Workspace, and generating a SMARTS pattern from these atoms. You can also type in SMARTS patterns, without having to place a ligand in the Workspace.

To edit the default SMARTS pattern or add a pattern:

1. (Optional) Display a ligand in the Workspace, on which you can select atoms that define the functional group.
2. Click Edit or New, to open the Edit Reactive Group dialog box.

Figure 5. The Edit Reactive Group dialog box.

3. Do one of the following to change the SMARTS pattern:
   - Edit the pattern in the SMARTS text box.
   - Pick atoms in the Workspace that include the functional group and any other atoms you want to include, then click Get From Selection.

     The atoms you choose must at least match the default SMARTS pattern. The SMARTS pattern for the selection is displayed in the SMARTS text box.
4. Choose an atom in the SMARTS pattern from the Position option menu to define the atom that forms a bond with the receptor.

   This option menu only displays the indices of the atoms that are capable of bonding to the receptor for the chosen reaction chemistry, so it might only contain one item.
5. Click OK.

When you have edited the default pattern, you can add more patterns. The Edit button changes to New, and you can click it to add a new pattern, using same procedure above.

The SMARTS patterns are listed in the table in the Reaction Type tab. This table shows the position in the SMARTS pattern, the total number of reaction sites found in the set of ligands you chose to dock, and the average number of sites. If these values indicate that some ligands may have more that one reactive site, you might want to edit the SMARTS patterns to provide a more restrictive expression and eliminate unwanted sites.

If you want to remove a pattern from the table, select it and click Delete. If you delete the only pattern in the table, the default pattern is added back to the table.

If you want to save the reactive group definitions for later use, you can export the SMARTS patterns and atom indexes to a CSV file, by clicking Export. You can then import them for another run by clicking Import. A file selector opens when you click the buttons, so you can specify the location and the file.

3.4 Constraining the Docking to a Reference Position

It can be useful to constrain a particular part of the ligands to the same position as in a known binder, which functions as a reference ligand. This is done in the Constraints tab, by defining the "core" of the reference ligand and specifying the maximum acceptable RMSD between the reference core position and the ligand core position.

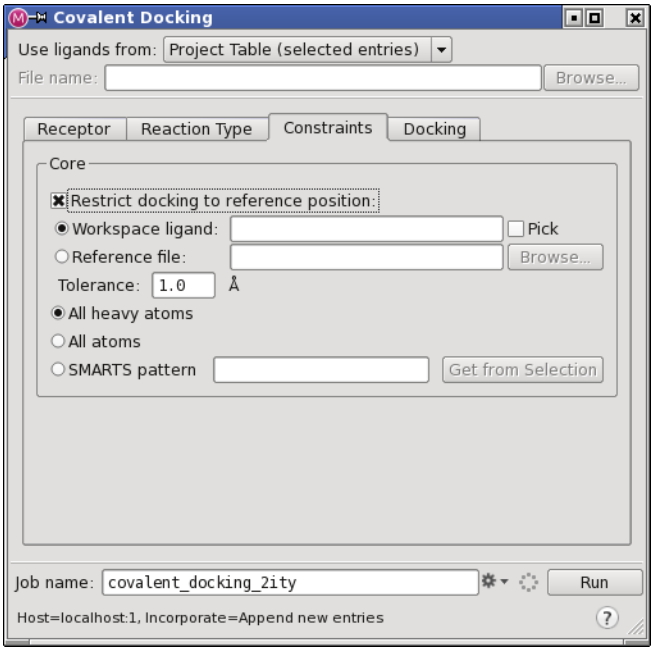

Figure 6. The Constraints tab of the Covalent Docking panel.

To apply a core constraint, select Restrict docking to reference position. The core is defined in terms of a set of atoms or a SMARTS pattern; if the ligand does not contain these atoms, it is skipped. This option must be selected to apply the core constraint, and to make settings for the core constraint.

There are two choices for the reference ligand.

- Workspace ligand—Use the current pose of the Workspace ligand as the reference ligand. If you used a ligand to define the grid center, this ligand is chosen by default.
- Reference file—Use the first structure in a specified file as the reference ligand. Click Browse to navigate to the file and select it. The file name is entered in the text box, and can be edited. You must ensure that the structure in the file is positioned properly with respect to the receptor.

There are three options for defining the core atoms. The first two, All heavy atoms and All atoms, select these atoms in the specified reference ligand. The third option, SMARTS pattern, allows you to define the core atoms in terms of a SMARTS pattern. You can pick atoms in the Workspace and click Get From Selection to define the SMARTS pattern, or you can type a SMARTS pattern into the text box. The atoms in the core-containing molecule that match the pattern are marked in the Workspace with green markers.

You can enter the tolerance for the maximum RMSD between the reference core position and the ligand core position in the Tolerance text box.

For further information on core constraints and how they work, see Section 5.5 of the *Glide User Manual*.

3.5 Choosing the Docking Mode

Covalent docking has two modes: Pose prediction, which performs the full protocol for accurate pose prediction, and Virtual screening, which skips time-consuming steps like rotamer sampling and minimization, to screen larger numbers of ligands. Virtual screening is at least 10 times faster than pose prediction. If you want to find possible covalent binders from a large number of ligands, choose Virtual screening. There is a limit of 10 000 ligands, so if you have more, you will have to run the jobs in batches. If you are concerned with finding good poses for a few ligands, choose Pose prediction. Pose prediction can take 1–2 hours per ligand. The choices are made from the Docking mode option menu in the Docking tab.

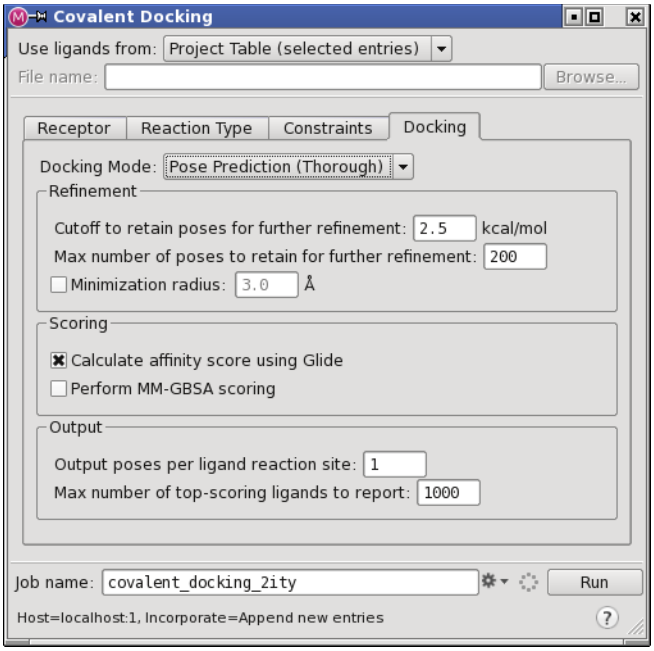

Figure 7. The Docking tab of the Covalent Docking panel.

3.6 Setting Refinement Options

After docking, the poses are filtered, then refined before clustering.You can change the filters for the poses by setting a cutoff on the GlideScore in the Cutoff to retain poses for further refinement text box, and by setting the maximum number of poses in the Max number of poses to retain for further refinement text box. Poses whose GlideScore is higher than the best GlideScore by the specified cutoff are discarded. The cutoff on the number of poses selects the best poses by GlideScore. The default values are 2.5 kcal/mol for the GlideScore cutoff, and a maximum of 200 poses.

If you want to refine more than just the ligand and the reactive residue before clustering the covalent poses, you can select Minimization radius and enter a distance in the text box. Residues that have any atoms within the specified distance of any atom in the ligand or reactive residue are included in the minimization.

This setting only affects the preliminary minimization before the poses are clustered and a representative pose is selected. Setting it should only be important if there is significant strain on the ligand due to nearby receptor residues. The default value of 3 Å should be sufficient; if the value is too large then the time for docking could increase significantly. The final poses

undergo a full minimization, so the ranking of poses for a given ligand should not be too dependent on the relaxation of the rest of the protein.

3.7 Setting Scoring Options

There are two options for scoring the poses. You can select either or both of these options.

- Calculate affinity score using Glide—Calculate a noncovalent binding affinity score using the results of Glide docking and a score-in-place calculation using the final docked pose. The average of the two GlideScore values is used for the affinity. This option is selected by default.
- Perform MM-GBSA scoring—Score the final docked pose using Prime MM-GBSA, to produce a binding affinity. Note that this binding affinity does not include the covalent binding: it is the noncovalent binding of the capped final pose to the mutated receptor. This option is not selected by default.

3.8 Setting the Number of Output Poses

By default, covalent docking only returns a single pose per ligand, which is the lowest-energy pose. If you want multiple poses to be returned, you can enter the number of poses you want in the Output poses per ligand reaction site text box. This could be important if there are multiple matches of the ligand SMARTS pattern or the receptor site, as the ranking of poses by Prime energy can only be done on the basis of a particular binding site.

For virtual screening, you can limit the number of ligands whose poses are returned, by setting a value in the Max number of top-scoring ligands to report text box. The poses are sorted by score and the cutoff is applied to the best scoring ligands. This limit can be applied to pose prediction as well, but is probably more useful for virtual screening, to limit the output.

3.9 Running the Job

When you have finished making settings for the docking process, you might want to make job settings before starting the job. In particular, you might want to distribute the calculation over multiple processors, as the docking process for a single ligand could take hours. Click the Settings button (gear icon) to open the Job Settings dialog box. In this dialog box you can make job settings, and also start the job by clicking Run. Once you have made job settings for one job, you can click Run in the Covalent Docking panel to run a job with those job settings.

3.10 Output

The structural output from a covalent docking job consists of the covalently bound poses, in a Maestro file. Each pose has three main properties associated with it: the affinity score, the Prime energy of the pose, and an index that identifies the reactive site on the ligand if the ligand has more than one such site. Since comparing poses is only meaningful for poses bonded at the same site, you can use the index to select and compare poses for a particular ligand site. If you selected the option to calculate the Prime MM-GBSA binding affinity, this is also included as a property.

4 Running Covalent Docking from the Command Line

It is also possible to run covalent docking jobs from the command line. The entire job can be set up with command options, or with an input file, or both. The syntax of the command is:

`covalent_docking` [*input-file*] [*options*] *receptor-file ligand-file attachment-residue*

where *input-file* is an optional input file with extension `.inp`, *receptor-file* is a Maestro file that contains the receptor structure, *ligand-file* is a Maestro file that contains the ligand structures, and *attachment-residue* identifies the residue on the protein to which the ligands are attached, in the format *chain*`:`*residue-number*, where *chain* is the letter for the chain name, and *residue-number* is the number of the residue in that chain, e.g `A:123`.

For information on the command options, run the command `covalent_docking -h`. Standard Job Control options that specify the host and other resources are also accepted. These options are described in Section 2.3 of the *Job Control Guide*.

The input file uses exactly the same syntax as the command options, except that the initial dash is not present when you convert the command option into an input keyword. Keywords are case-insensitive, and are separated from their value by spaces. Any options given on the command line supersede the corresponding options in the input file.

As the jobs can take several hours per ligand in pose prediction mode, you should consider distributing them over multiple processors. You can do this with the `-NJOBS` option. To protect against failure of the subjobs (e.g. due to network issues), you can set the `-RETRIES` option to specify the number of times the master job reruns a failed subjob before it is considered to have failed. If the job finishes with failed subjobs, or if the job itself fails for any reason, you can restart the job with the `-RESTART` option, and it will resume at the first subjob that did not run successfully, keeping the results of the successful subjobs.

5 Custom Chemistry Definitions

If you want to use a reaction other than the predefined reactions, you can define your own custom chemistry, by using keywords that you add to a file, one keyword per line. You can also use the corresponding command options to define a custom chemistry. However, storing the keywords in a file is more useful, as it allows reuse of the definitions.

5.1 Defining the Reactive Groups

The first task is to define the ligand reactive group and the reactive residue on the receptor. This task is done with the `LIGAND_SMARTS_PATTERN` and `RECEPTOR_SMARTS_PATTERN` keywords. You can only have one instance of the `RECEPTOR_SMARTS_PATTERN` keyword, but you can have multiple instances of the `LIGAND_SMARTS_PATTERN` keyword, and all matches to the ligand SMARTS patterns are included in the docking run.

The reaction sites are defined in terms of SMARTS patterns, and the atom that forms the covalent bond on each side is defined by the index of an atom in the SMARTS pattern. The general syntax is:

keyword atom-number , SMARTS-pattern

where *keyword* is one of the two keywords above, and *atom-number* is the atom number in the SMARTS pattern defined by *SMARTS-pattern*.

For example, if the receptor SMARTS pattern is:

```
[C]-[S,O;H1,-1]
```

and the second atom in the pattern is where the reaction will occur, the line to use is:

```
RECEPTOR_SMARTS_PATTERN 2,[C]-[S,O;H1,-1]
```

where the number 2 defines the location of the reactive atom in the SMARTS pattern.

If the reaction type is an epoxide opening on the ligand, for example, the ligand SMARTS pattern could be:

```
[C;r3][O;r3][C;r3]
```

and the reaction site of the ligand is the first carbon in the pattern. The line to use in the file is:

```
LIGAND_SMARTS_PATTERN 1,[C;r3][O;r3][C;r3]
```

Note that this pattern could match either of the carbon atoms in an epoxide—see Figure 8. The attachment atom could actually be either of the atoms numbered 1 and 3 in the figure. As a consequence, both of the possible matches to the ligand will be docked.

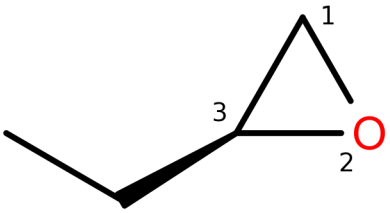

Figure 8. Numbering of the atoms in the SMARTS pattern for an epoxide.

5.2 Defining the Reaction

With the receptor and ligand reactive groups defined, the next stage is to define the reaction. This involves specifying the changes in the bonding and charges that have to be made to create the covalently bound complex from the receptor and the ligand, and can be viewed as a sequence of steps.

Each step is defined by a `CUSTOM_CHEMISTRY` keyword, and the steps are combined in the order in which they are specified to produce a reaction. The steps do not have to match the actual reaction pathway, they only need to produce the correct product. There is no need to adjust hydrogens, as they will be added and removed as necessary.

The formation of the covalent bond is done with one or more of the following operations:

- set the formal charge of any atom
- add, remove or change the bond order of any bond
- remove atoms (leaving groups)
- change the chirality of any chiral center

Atoms are selected with SMARTS patterns with a few additions:

- `<1>` matches the receptor attachment atom. This is the atom specified by its index in the `RECEPTOR_SMARTS_PATTERN` keyword. For example, `<1>-[O]` matches the receptor attachment atom bound to an oxygen.
- `<2>` matches the ligand attachment atom. This is the atom specified by its index in the `LIGAND_SMARTS_PATTERN` keyword. For example, `<2>(C)O` matches the ligand attachment atom bound to a carbon and an oxygen.
- `|` means that the pattern spans two molecules. For example, `<1>|<2>` matches the ligand attachment atom and the receptor attachment atom when they are not connected.

The general syntax of the `CUSTOM_CHEMISTRY` keyword is

`CUSTOM_CHEMISTRY (`*pattern*`,("`*keyword*`",[`*value*`,]`*indices*`))`

where *pattern* is an augmented SMARTS pattern as described above, *keyword* is one of those listed below, *value* is the desired value if a value is required, and *indices* is the atom index or list of atom indices in the SMARTS pattern that define which atoms to apply the operation to. If the syntax requires more than one index, the indices should be enclosed in parentheses and separated by commas.

There are four keywords: `charge`, `bond`, `delete`, and `chiral`. Each of these is described below with examples.

To set the formal charge on an atom:

`CUSTOM_CHEMISTRY (`*pattern*`,("charge",`*formal-charge*`,`*indices*`))`

For example, to set the formal charge on the receptor attachment atom to zero:

```
CUSTOM_CHEMISTRY ("<1>",("charge",0,1))
```

The receptor atom has the index 1 in the pattern `"<1>"`, and the charge is set to 0.

To set the bond order between two atoms:

`CUSTOM_CHEMISTRY (`*pattern*`,("bond",`*bond-order*`,`*indices*`))`

Here, bond-order is a positive integer. For example, to set the bond order between the receptor attachment atom and the ligand attachment atom to 1, i.e. to create a bond:

```
CUSTOM_CHEMISTRY ("<1>|<2>",("bond",1,(1,2)))
```

To delete a bond:

`CUSTOM_CHEMISTRY (`*pattern*`,("bond",-1,`*indices*`))`

Setting the bond order to a negative value means “delete the bond”. For example, to delete a C–O bond in an epoxide opening on the ligand:

```
CUSTOM_CHEMISTRY ("<2>[O;r3][C;r3]",("bond",-1,(1,2)))
```

Here, the bond is broken between atoms 1 and 2 in the SMARTS pattern, which are the ligand attachment atom, `<2>`, and the ring oxygen, `[O;r3]`.

To delete heavy atoms:

`CUSTOM_CHEMISTRY (`*pattern*`,("delete",`*indices*`))`

No value is needed here as the atoms are simply removed. For example, to delete a hydroxyl group adjacent to the ligand attachment atom:

```
CUSTOM_CHEMISTRY ("<2>-[O;H1,-1]",("delete",2))
```

To set the stereochemistry of chiral centers:

`CUSTOM_CHEMISTRY (`*pattern*`,("chiral",R|S,`*indices*`))`

For example, change the receptor attachment atom chirality to R:

```
CUSTOM_CHEMISTRY ("<1>",("chiral","R",1))
```

5.2.1 Saving and Using the Definition

You can save the custom chemistry definition in a file. The file must contain the complete definition, and can also contain comment lines, which begin with a # symbol. The file name must have a `.cdock` extension.

Below is an example of the complete chemistry definition for an epoxide opening from the examples above:

```
LIGAND_SMARTS_PATTERN    1,[C;r3][O;r3][C;r3]
RECEPTOR_SMARTS_PATTERN  2,[C]-[S,O;H1,-1]
CUSTOM_CHEMISTRY         ("<1>",("charge",0,1))
CUSTOM_CHEMISTRY         ("<1>|<2>",("bond",1,(1,2)))
CUSTOM_CHEMISTRY         ("<2>[O;r3][C;r3]",("bond",-1,(1,2)))
```

You could, for example, save this definition in a file named `epoxide_opening.cdock`.

To use the saved definition, you can import it for use in the Covalent Docking panel (see Section 3.2 on page 12), or you can use the `-reaction_file` option to select it when running from the command line.

You can also add a custom chemistry definition to your input file and run the job from the command line. Saving it as a separate file makes it easier to use for more than one job.

Getting Help

Information about Schrödinger software is available in two main places:

- The `docs` folder (directory) of your software installation, which contains HTML and PDF documentation. Index pages are available in this folder.
- The Schrödinger web site, http://www.schrodinger.com/, In particular, you can use the Knowledge Base, http://www.schrodinger.com/kb, to find current information on a range of topics, and the Known Issues page, http://www.schrodinger.com/knownissues, to find information on software issues.

Finding Information in Maestro

Maestro provides access to nearly all the information available on Schrödinger software.

To get information:

- Pause the pointer over a GUI feature (button, menu item, menu, ...). In the main window, information is displayed in the Auto-Help text box, which is located at the foot of the main window, or in a tooltip. In other panels, information is displayed in a tooltip.

  If the tooltip does not appear within a second, check that Show tooltips is selected under General → Appearance in the Preferences panel, which you can open with CTRL+, (⌘,). Not all features have tooltips.

- Click the Help button in the lower right corner of a panel or press F1, for information about a panel or the tab that is displayed in a panel. The help topic is displayed in the Help panel. The button may have text or an icon:

  

- Choose Help → Online Help or press CTRL+H (⌘H) to open the default help topic.
- When help is displayed in the Help panel, use the navigation links in the help topic or search the help.
- Choose Help → Documentation Index, to open a page that has links to all the documents. Click a link to open the document.

- Choose Help → Search Manuals to search the manuals. The search tab in Adobe Reader opens, and you can search across all the PDF documents. You must have Adobe Reader installed to use this feature.

For information on:

- Problems and solutions: choose Help → Knowledge Base or Help → Known Issues → *product*.
- New software features: choose Help → New Features.
- Python scripting: choose Help → Python Module Overview.
- Utility programs: choose Help → About Utilities.
- Keyboard shortcuts: choose Help → Keyboard Shortcuts.
- Installation and licensing: see the *Installation Guide*.
- Running and managing jobs: see the *Job Control Guide*.
- Using Maestro: see the *Maestro User Manual*.
- Maestro commands: see the *Maestro Command Reference Manual*.

Contacting Technical Support

If you have questions that are not answered from any of the above sources, contact Schrödinger using the information below.

Web:	http://www.schrodinger.com/supportcenter
E-mail:	help@schrodinger.com
Mail:	Schrödinger, 101 SW Main Street, Suite 1300, Portland, OR 97204
Phone:	+1 888 891-4701 (USA, 8am – 8pm Eastern Time)
	+49 621 438-55173 (Europe, 9am – 5pm Central European Time)
Fax:	+1 503 299-4532 (USA, Portland office)
FTP:	ftp://ftp.schrodinger.com

Generally, using the web form is best because you can add machine output and upload files, if necessary. You will need to include the following information:

- All relevant user input and machine output
- Covalent Docking purchaser (company, research institution, or individual)
- Primary Covalent Docking user
- Installation, licensing, and machine information as described below.

Gathering Information for Technical Support

The instructions below describe how to gather the required machine, licensing, and installation information, and any other job-related or failure-related information, to send to technical support. Where the instructions depend on the profile used for Maestro, the profile is indicated.

For general enquiries or problems:

1. Open the Diagnostics panel.
   - **Maestro:** Help → Diagnostics
   - **Windows:** Start → All Programs → Schrodinger-2015-2 → Diagnostics
   - **Mac:** Applications → Schrodinger2015-2 → Diagnostics
   - **Command line:** `$SCHRODINGER/diagnostics`
2. When the diagnostics have run, click Technical Support.

   A dialog box opens, with instructions. You can highlight and copy the name of the file.
3. Upload the file specified in the dialog box to the support web form.

   If you have already submitted a support request, use the upload link in the email response from Schrödinger to upload the file. If you need to submit a new request, you can upload the file when you fill in the form.

If your job failed:

1. Open the Monitor panel, using the instructions for your profile as given below:
   - **Maestro/Jaguar/Elements:** Tasks → Monitor Jobs
   - **BioLuminate/MaterialsScience:** Tasks → Job Monitor
2. Select the failed job in the table, and click Postmortem.

   The Postmortem panel opens.
3. If your data is not sensitive and you can send it, select Include structures and deselect Automatically obfuscate path names.
4. Click Create.

   An archive file is created, and an information dialog box with the name and location of the file opens. You can highlight and copy the name of the file.
5. Upload the file specified in the dialog box to the support web form.

   If you have already submitted a support request, use the upload link in the email response from Schrödinger to upload the file. If you need to submit a new request, you can upload the file when you fill in the form.

6. Copy and paste any log messages from the window used to start the interface or the job into the web form (or an e-mail message), or attach them as a file.
   - **Windows:** Right-click in the window and choose Select All, then press ENTER to copy the text.
   - **Mac:** Start the Console application (Applications → Utilities), filter on the application that you used to start the job (Maestro, BioLuminate, Elements), copy the text.

If Maestro failed:

1. Open the Diagnostics panel.
   - **Windows:** Start → All Programs → Schrodinger-2015-2 → Diagnostics
   - **Mac:** Applications → SchrodingerSuite2015-2 → Diagnostics
   - **Linux/command line:** `$SCHRODINGER/diagnostics`
2. When the diagnostics have run, click Technical Support.

   A dialog box opens, with instructions. You can highlight and copy the name of the file.
3. Upload the file specified in the dialog box to the support web form.

   If you have already submitted a support request, use the upload link in the email response from Schrödinger to upload the file. If you need to submit a new request, you can upload the file when you fill in the form.
4. Upload the error files to the support web form.

   The files should be in the following location:

- **Windows:** `%LOCALAPPDATA%\Schrodinger\appcrash`
  (Choose Start → Run and paste this location into the Open text box.)
  Attach `maestro_error_`*`pid`*`.txt` and `maestro.exe_`*`pid_timestamp`*`.dmp`.
- **Mac:** `$HOME/Library/Logs/CrashReporter`
  (Go → Home → Library → Logs → CrashReporter)
  Attach `maestro_error_`*`pid`*`.txt` and `maestro_`*`timestamp_machinename`*`.crash`.
- **Linux:** `$HOME/.schrodinger/appcrash`
  Attach `maestro_error_`*`pid`*`.txt` and `crash_report_`*`timestamp_pid`*`.txt`.

If a Maestro panel failed to open:

1. Copy the text in the dialog box that opens.
2. Paste the text into the support web form.

120 West 45th Street
17th Floor
New York, NY 10036

101 SW Main Street
Suite 1300
Portland, OR 97204

245 First Street
Riverview II, 18th Floor
Cambridge, MA 02142

8910 University Center Lane
Suite 270
San Diego, CA 92122

155 Gibbs St
Suite 430
Rockville, MD 20850-0353

Dynamostraße 13
D-68165 Mannheim
Germany

Zeppelinstraße 73
D-81669 München
Germany

Potsdamer Platz 11
D-10785 Berlin
Germany

Quatro House
Frimley Road
Camberley GU16 7ER
United Kingdom

8F Pacific Century Place
1-11-1 Marunouchi
Chiyoda-ku, Tokyo 100-6208
Japan

No. 102, 4th Block
3rd Main Road, 3rd Stage
Sharada Colony
Basaveshwaranagar
Bangalore 560079, India

SCHRÖDINGER®