

Virtual Screening Workflow

Schrödinger Suite 2012 Update 2

Virtual Screening Workflow Copyright © 2012 Schrödinger, LLC. All rights reserved.

While care has been taken in the preparation of this publication, Schrödinger assumes no responsibility for errors or omissions, or for damages resulting from the use of the information contained herein.

BioLuminate, Canvas, CombiGlide, ConfGen, Epik, Glide, Impact, Jaguar, Liaison, LigPrep, Maestro, Phase, Prime, PrimeX, QikProp, QikFit, QikSim, QSite, SiteMap, Strike, and WaterMap are trademarks of Schrödinger, LLC. Schrödinger and MacroModel are registered trademarks of Schrödinger, LLC. MCPRO is a trademark of William L. Jorgensen. DESMOND is a trademark of D. E. Shaw Research, LLC. Desmond is used with the permission of D. E. Shaw Research. All rights reserved. This publication may contain the trademarks of other companies.

Schrödinger software includes software and libraries provided by third parties. For details of the copyrights, and terms and conditions associated with such included third party software, see the [Legal Notices](#), or use your browser to open `$$SCHRODINGER/docs/html/third_party_legal.html` (Linux or Mac OS) or `%SCHRODINGER%\docs\html\third_party_legal.html` (Windows OS).

This publication may refer to other third party software not included in or with Schrödinger software ("such other third party software"), and provide links to third party Web sites ("linked sites"). References to such other third party software or linked sites do not constitute an endorsement by Schrödinger, LLC or its affiliates. Use of such other third party software and linked sites may be subject to third party license agreements and fees. Schrödinger, LLC and its affiliates have no responsibility or liability, directly or indirectly, for such other third party software and linked sites, or for damage resulting from the use thereof. Any warranties that we make regarding Schrödinger products and services do not apply to such other third party software or linked sites, or to the interaction between, or interoperability of, Schrödinger products and services and such other third party software.

Revision A, September 2012

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	<code>\$SCHRODINGER/maestro</code>	File names, directory names, commands, environment variables, command input and output
Italic	<i>filename</i>	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].

Virtual Screening Workflow

The Virtual Screening Workflow is designed to run an entire sequence of jobs for screening large collections of compounds against one or more targets. The workflow includes ligand preparation using LigPrep, filtering using `ligfilter` on QikProp properties or other structural properties, and Glide docking at the three accuracy levels, HTVS, SP, and XP. The design allows you to choose which of the stages to include in any run, and which selection of results from one stage are passed on to the next.

Before you run the workflow, you must ensure that receptors are properly prepared. See the [Protein Preparation Guide](#) for information on protein preparation. You can generate grids as part of the workflow, or you can generate them beforehand with Glide. See [Chapter 4](#) of the *Glide User Manual* for information on grid generation.

The ligand files for the workflow can consist of 2D structures or 3D structures. If you supply 2D structures, you must run the ligand preparation part of the workflow to convert the structures to 3D for docking. This stage runs a LigPrep job, with the standard options for most parts of the LigPrep process. You can control some of these options, but if you want to use other options, you should run LigPrep on the ligands independently. For more information, see the [LigPrep User Manual](#).

Prefiltering of the ligands can also be performed as part of the workflow. If you want to use the Lipinski Rule option, you must run QikProp to obtain the required properties. You can run QikProp as part of the workflow, regardless of the ligand source. If you have already run QikProp, you do not need to run it again. If you want to specify a custom filter, you can provide an input file for `ligfilter`, which will then filter the structures. For more information on `ligfilter`, see [Section 2.5](#) of the *General Utilities* manual.

The full workflow includes three docking stages. The first stage performs HTVS docking. The ligands that are retained are then passed to the next stage, which performs SP docking; the survivors of this stage are passed on to the third stage, which performs XP docking. At each stage you can decide how many ligands are kept, and whether all the ionization and tautomeric states of each ligand are kept, or only some.

The workflow is intended to be robust, so that if a subjob fails, the master job will attempt to rerun the job a few times before quitting. If the master job fails (for example, due to a system failure or network failure) you can restart the job, and it will pick up the calculation from the latest point for which it has results and can restart from them.

1 Running Schrödinger Software

Schrödinger applications can be started 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 Maestro. This directory then becomes Maestro's 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-2012 > 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 (Documents on Windows 7/Vista, My Documents on XP).

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 that 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-2012. 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 SchrodingerSuite2012 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. If you need to set any other variables, use the command

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

All Platforms:

To run remote jobs, you must have access to a hosts file, named `schrodinger.hosts`, that lists hosts on which Schrödinger software is installed for execution. Details of setting up this file can be found in [Chapter 7](#) of the *Installation Guide*.

2 The Virtual Screening Workflow Panel

The Virtual Screening Workflow panel sets up the input files for LigPrep, QikProp, and Glide ligand docking and submits them to the selected host in order.

To open the Virtual Screening Workflow panel, choose Workflows > Virtual Screening Workflow in the main window.

The Virtual Screening Workflow panel consists of five tabs for setting up virtual screening jobs. The features of these tabs and how to use them are described in separate sections.

- Input Tab
- Filtering Tab
- Preparation Tab
- Receptors Tab
- Docking Tab

Below the tabs is a row of buttons:

- **Start**—Write the input file and start the job. Opens the Start dialog box, in which you can set job parameters, including distributing jobs over multiple processors. When you distribute jobs, the driver is run on the local host, and submits subjobs to the specified hosts.
- **Write**—Write the input file but do not start the job. Opens a dialog box in which you can specify the input file name, without the extension. This is the equivalent of the job name in the Start dialog box, and is used to construct file names.
- **Reset**—Reset all panel settings to their defaults

3 Specifying the Ligands

In the Input tab you specify the source of the ligands, and ensure that each ligand is uniquely identified. Ligands can be taken from two sources: files (in Maestro, SD, or SMILES format), or a Phase database.

3.1 Using Ligands from Files

To use ligands from one or more files, select Files for the Source of ligands option.

- To specify ligand files, you can enter a comma-separated list of file names in the Files text box, or click **Browse** to navigate to and select the files.
- To specify multiple ligand files with related names, you can use the wild card characters * and ? in the file name. These characters have their usual Unix file-matching meanings: ? matches a single character, and * matches zero or more characters.
- To specify multiple ligand files with unrelated names, you can create a text file that contains a list of ligand file names, and specify this text file in the Input ligand file text box, or click **Browse** and navigate to it. This is useful if the files are in different directories.

If you type in the file name, you must press **ENTER** to ensure that the name is read and the Using existing property option menu is populated.

The Use input files directly for subjobs option allows you to run a separate subjob for each ligand file. Doing so is useful if your ligands are already split across files in a way that is suitable for running a distributed job efficiently. If you do, you cannot give ligands a unique identifier using the controls described in the next section. In addition, no checking is done for duplicate ligands.

The Combine input files and redistribute for subjobs option allows you to optimize the load balancing of the subjobs, and ensure that each compound has a unique identifier. This option reads all the input files, processes the structures to create a unique property and checks for

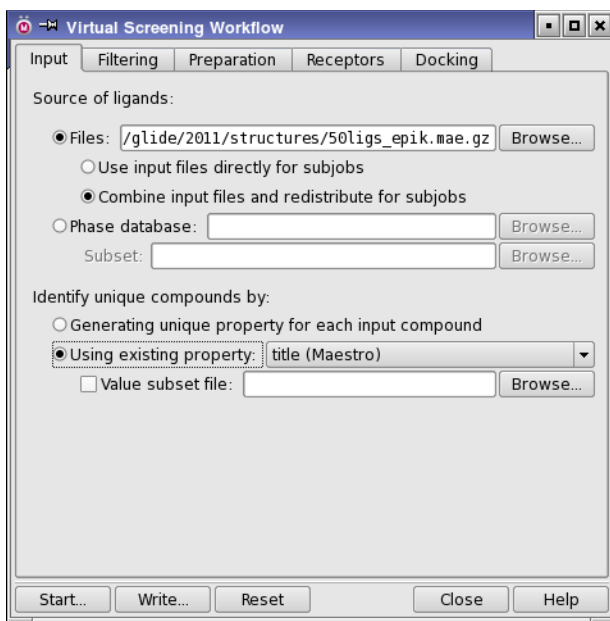


Figure 1. The Input tab of the Virtual Screening Workflow panel.

duplicates, then writes out new files for the subjobs. In doing so, it uses more disk space than using the files directly for subjobs.

3.2 Using Ligands from a Phase Database

To use ligands from a Phase database, select Phase database for the Source of ligands option.

You can type the name of the database into the text box, without the .phdb extension, or click Browse and navigate to the database. If you have a Phase database in an earlier format (with a _phasedb suffix), you must convert it first, using the convert task of phase_database.

If you have subsets defined for the Phase database, you can restrict the ligand set to the subset specified in the Subset text box. Type in the path to the subset, minus the _phase.inp suffix, or click Browse and navigate to the database subset.

3.3 Identifying Ligands and their States

In the docking sequence, you can choose to keep all the ionization and tautomeric states of a given compound for which one of these states docks well. The ionization and tautomeric states that originate from the same compound are identified by a property that is unique for the compound. The controls under Identify unique compounds by enable you to set or select the

unique property for the ligands. These controls are not available until you have specified the ligand source.

If the structures you have are all unique compounds, you can assign a unique property to each with the first option, **Generating unique property for each input compound**. The value assigned to this property is an integer.

If the structures contain different ionization states or tautomers of the same compound, you can use an existing property for the unique compound identifier by selecting **Using existing property** and choosing a property from the option menu. The property names are taken from the first structure in each file, and only those properties that exist in each file are presented. You should ensure that the property you choose exists for each structure in the file, not just the first. The option menu becomes available when a file with valid properties is specified.

The property that you choose can also be used to filter the structures, so that only those containing certain values for the selected property are docked. The values are specified in a text file, one value per line. To apply this filter, select **Value subset file**, and enter the file name in the text box, or click **Browse** and navigate to the file. This option is only available if you use a property to create ligand titles. However, you can define a more flexible filtering scheme using ligand properties in the **Filtering** tab, described below.

4 Filtering the Ligands

In the **Filtering** tab you can set options for ligand filtering. This tab allows you to run QikProp and filter the ligands by property.

The Virtual Screening Workflow offers three choices for prefiltering ligands: prefiltering using Lipinski's Rule of 5, removing ligands with reactive functional groups, and prefiltering with a custom filter. Ligands that do not meet the specified criteria are removed from the ligand list for docking. Both prefiltering choices use `ligfilter`, which can filter the structures in a Maestro file based on any property in the file.

When you have finished setting up the filter, you can choose whether to apply the filter before preparing the ligands, or afterwards, by selecting **Filter before preparation** or **Filter after preparation**. Filtering that depends only on 2D properties, such as functional group counts, can be done beforehand, and thus save the preparation time. Filtering that depends on 3D properties of the structures can only be done afterwards, if the input ligands are 2D.

4.1 Filtering by Lipinski's Rule

To prefilter the ligands using Lipinski's Rule of 5, select **Prefilter by Lipinski's Rule**. Ligands that do not satisfy this rule are not docked. This option requires QikProp properties. If the input structure files do not have QikProp properties, select **Run QikProp**.

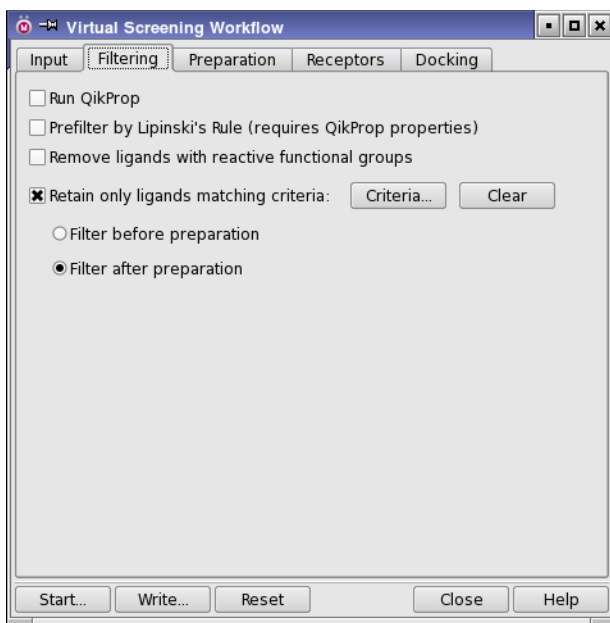


Figure 2. The Filtering tab of the Virtual Screening Workflow panel.

4.2 Filtering Out Reactive Functional Groups

To filter out ligands that have reactive functional groups, select Remove ligands with reactive functional groups. The functional groups that are considered reactive are:

- Acyl halides
- Sulfonyl halides
- Sulfinyl halides
- Sulfenyl halides
- Alkyl halides without fluorine
- Anhydrides
- Perhalomethylketones
- Aldehydes
- Formates
- Peroxides
- R-S-O-R
- Isothiocyanates
- Isocyanates
- Phosphinyl halides
- Phosphonyl halides
- Alkali metals
- Alkaline-earth metals
- Lanthanide series metals
- Actinide series metals
- Transition metals
- Other metals
- Toxic nonmetals
- Noble gases
- Carbodiimides
- Silyl enol ethers
- Nitroalkanes
- Phosphines
- Alkyl sulfonates
- Epoxides
- Azides
- Diazonium compounds
- Isonitriles
- Halopyrimidines
- 1,2-Dicarbonyls
- Michael acceptors
- Beta-heterosubstituted carbonyls
- Diazo compounds
- R-N-S-R
- Disulfides

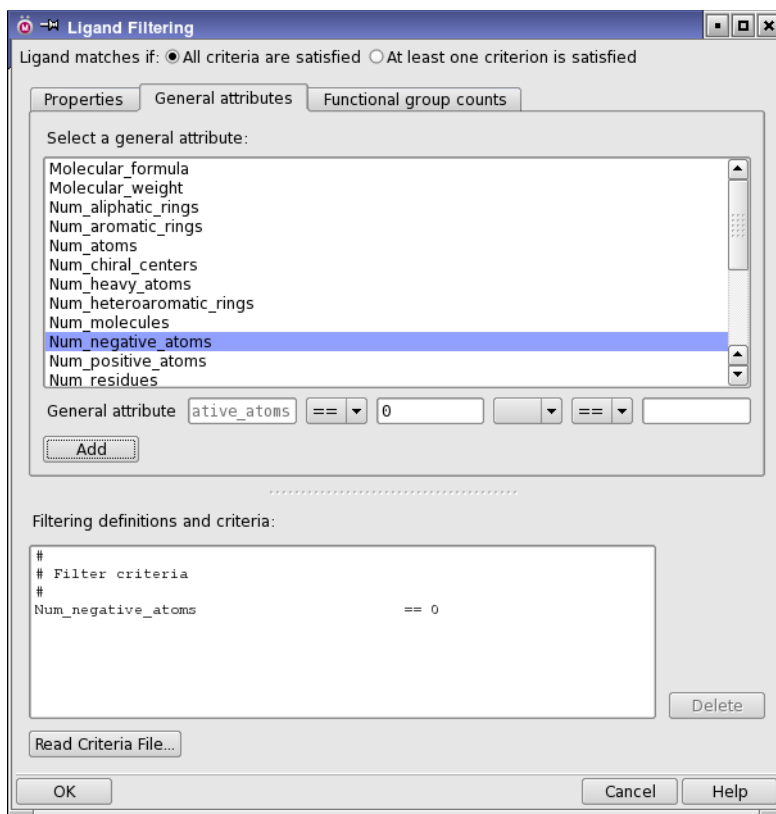


Figure 3. The Ligand Filtering dialog box, showing the General Attributes tab.

4.3 Filtering with Custom Filters

To set up your own filter, select Retain only ligands matching criteria, then click Criteria to set up or read in the filter. The Ligand Filtering dialog box opens, which allows you to define the filtering criteria or read a file containing filtering criteria.

Filters are constructed from a combination of criteria set for various Maestro properties, certain predefined feature counts, or counts of SMARTS patterns for functional groups. Each ligand must meet all of the criteria to pass through the filter.

The criteria are set by selecting an item from the list in one of the tabs, selecting a relational operator, and setting a value, then clicking Add. The criterion is added to the list at the bottom of the panel. You can delete a criterion by selecting it and clicking Delete. When you have added all the desired criteria, click OK to accept the filter.

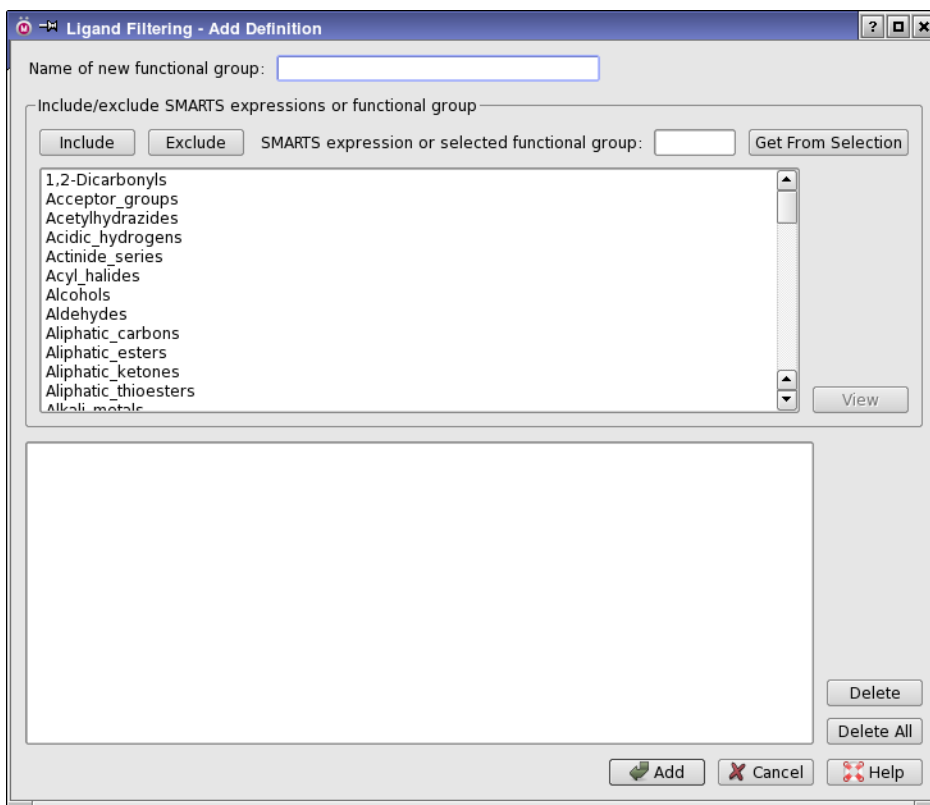


Figure 4. The Ligand Filtering - Add Definition dialog box

To read a filter, click Read criteria file, and navigate to the desired file in the file chooser that opens. The file must have the extension `.lff`. When you click OK, the file is read and the criteria are listed in the Filtering definitions and criteria text area. You can then add criteria to the list, or delete criteria from the list.

In the Properties tab, all properties that are defined in the input files for each ligand are listed. You can display properties by family (this includes product families, such as QikProp) to limit the range of properties available, by choosing the family from the Show family option menu.

In the Functional Group Counts tab, you can create a new pattern by clicking New and selecting or entering patterns in the Add definition dialog box. This dialog box provides tools for including or excluding patterns: ligands must match included patterns, but must not match excluded patterns. A complex pattern can thus be built up out of simple patterns. You can choose from the existing patterns, or enter your own SMARTS pattern either by typing it in the text box or by selecting atoms in the Workspace and clicking Get From Selection. Then click

Include or Exclude to add the pattern to the definition. The pattern is added to the list in the lower half of the dialog box. You can delete patterns from this list by selecting them and clicking Delete. To view the definition of an existing pattern, select the pattern and click View definition. To complete the pattern definition, enter a name in the Name of new functional group text box at the top of the dialog box, and click Add.

5 Preparing the Ligands

In the Preparation tab you set options for ligand preparation with LigPrep. If you only want to prepare and filter a set of ligands, you can run the job when you have made settings in this tab, and you can ignore the remaining two tabs, which specify the receptors and the docking options.

If your structures are already 3D all-atom structures with the appropriate ionization and tautomeric states, you do not need to prepare the ligands, and you can deselect the Prepare ligands option. This is usually the case if you are docking from a Phase database. Otherwise, ensure that this option is selected, and choose the desired options for ligand preparation.

You can choose to regularize the input geometries, so that ligands that are identical apart from small deviations in the atom positions are always treated the same. To do so, select Regularize input geometries. This process is run prior to LigPrep, and uses a conversion to unique SMILES to ensure the same atom numbering and geometry. All atom properties are lost when you perform this conversion, but entry properties, including the title, are retained.

As part of the regularization, you can remove duplicate structures, by selecting Remove duplicates. When testing for duplicates, you can remove any other small molecules such as counter ions, by selecting Desalt; you can neutralize the structures, by selecting Neutralize; and you can merge the properties from the two structures, by selecting Merge properties. If the properties have different values, the property is converted to a string that lists the values.

To generate ionization (protonation) states that are likely to exist in a given pH range, enter the target pH and the range in the Generate states at pH text boxes. You can generate these states with either the Ionizer or Epik, by selecting the appropriate option. Epik is licensed separately from LigPrep; the Ionizer is included with LigPrep. In addition to varying the ionization state, LigPrep generates tautomeric states that are likely in the given pH range. You can remove ionization and tautomeric states that have large penalties by selecting Remove high-energy ionization/tautomer states. These are states that are likely to have low populations under the chosen conditions. If you select Epik for state generation, you can also generate extra states that are appropriate for binding to metals in metalloproteins, by selecting Add Epik metal binding states. These states are not normally generated at physiological pH.

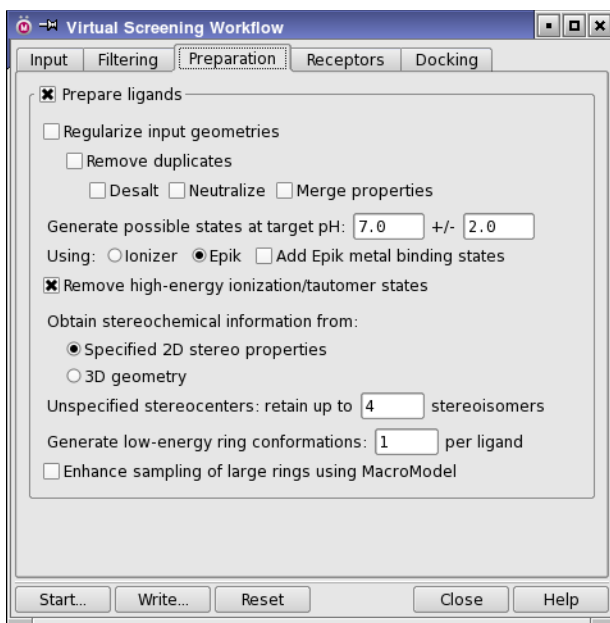


Figure 5. The *Preparation* tab of the *Virtual Screening Workflow* panel.

The LigPrep run varies the stereochemistry, if it is not predetermined. Any existing chirality information in the input file is preserved, and chemically reasonable chiralities are initially assigned for steroids, fused ring systems, and peptides. If you already have 3D structures, you can select 3D Geometry under Obtain stereochemical information from to determine the stereochemistry from the structure. Otherwise select Specified 2D stereo properties to make use of information in 2D structure files. For the unspecified stereocenters, up to 32 stereoisomers are generated. These are ranked by energy, and only the lowest-energy stereoisomers are returned. You can control how many are returned by entering the number to return in the Unspecified stereocenters: retain up to N stereoisomers text box.

Ring conformations are varied by LigPrep. You can limit the number of ring conformations returned by entering the maximum number in the Generate low-energy ring conformations: N per ligand text box. This sampling does not include 7-membered and larger rings, which can be sampled with MacroModel after ligand preparation by selecting Enhance sampling of large rings using MacroModel.

6 Selecting or Setting Up the Receptor Grids

If you want to dock the ligands that you have prepared and filtered, you must specify one or more grids for the receptor in the Receptors tab. The ligands are docked to each receptor that you specify. Grids can be added to the Receptors for docking table by clicking Add or Add Multiple. The maximum number of receptors you can add is 100. You can edit the grid information by clicking Edit. To remove grids from the table (and thus not use them), select the table rows and click Delete.

6.1 Adding a Single Grid

When you click Add, the VSW - Receptor dialog box opens. In this dialog box, you can select a pregenerated grid or generate a grid from the Workspace structure. You can also add constraints (described later), identify the receptor, and add an offset to the GlideScore.

Any receptor that you specify for grid generation must be properly prepared. To prepare a receptor, you can use the Protein Preparation Wizard panel, which you open from the Workflows menu. For information on protein preparation, see the [Protein Preparation Guide](#). If the receptor has not been marked as prepared with the Protein Preparation Wizard, a warning dialog box opens when you start the job. This dialog box offers you the choice of marking the protein as prepared (if, for example, it was prepared with a previous software release), ignoring the indication that the protein is not prepared, or canceling.

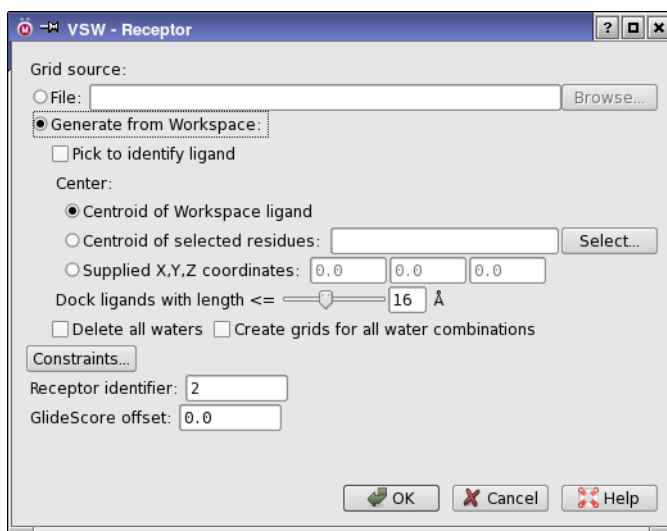


Figure 6. The VSW - Receptor dialog box.

To add an existing grid:

1. Select File under Grid source.
2. Enter the path in the text box, or click Browse and navigate to the grid.

To generate a grid for the Workspace structure:

1. Select Generate from Workspace under Grid source.
2. Display the receptor in the Workspace.

If you want to use a ligand to define the grid, the receptor entry must include a ligand. If it does not, create a new entry by merging the receptor and the ligand entry. You cannot use multiple entries to define the grid.

3. If the receptor entry contains a ligand, select Pick to identify ligand and click on a ligand atom.
4. Choose an option for defining the grid center:
 - Centroid of Workspace ligand—use the ligand centroid to define the grid center.
 - Centroid of selected residues—select the residues to define the grid center, by clicking Select and using the Atom Selection dialog box.
 - Supplied X, Y, Z coordinates—enter the coordinates of the grid center.
5. Enter the maximum ligand size in the Dock ligands with length <= text box.
6. Choose an option for handling waters.
 - If you do not want to include structural waters, select Delete all waters.
 - If you want to run calculations on the same protein structure with and without structural waters, select Create grids for all water combinations. This option generates grids for each number and combination of water molecules that are explicitly included with the receptor, so that you can dock the ligands with all possible combinations of structural waters. The number of grids increases exponentially with the number of water molecules, so you should take care to only include critical waters.

When you have selected the grid source and set up options, you can then choose to add constraints to the generated grid or to apply constraints for a preexisting grid. Constraints are described in [Section 6.3](#).

You must provide an identifier for the receptor, by entering text in the Receptor identifier text box. The default is to use a numeric index.

If you want to shift the GlideScores for this receptor, to align results for different receptors, enter a value in the GlideScore offset text box.

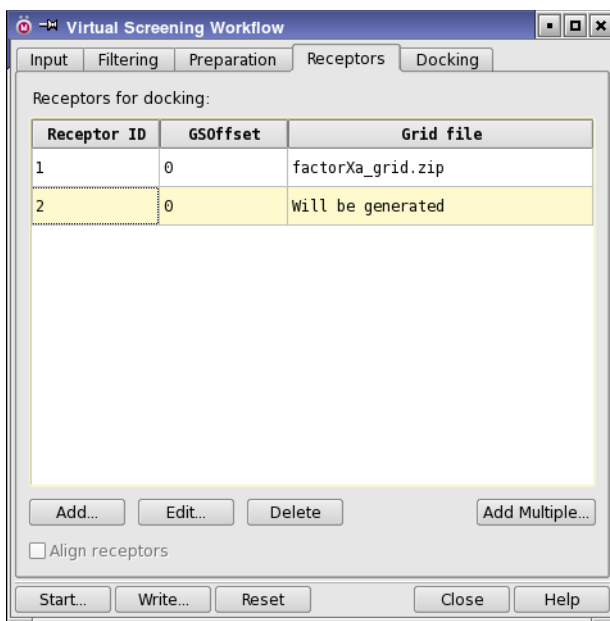


Figure 7. The Receptors tab of the Virtual Screening Workflow panel.

When you click OK, the dialog box closes and the receptor is added to the Receptors for docking table in the main panel.

If you want to change any data associated with a receptor, click Edit. The VSW - Receptor dialog box opens with the grid information loaded. You can do anything to edit the receptor that is described in this section, including completely replacing it. When you have finished making changes, click OK.

6.2 Adding Multiple Pregenerated Grids

If you want to add multiple receptors to the table in a single action, click Add Multiple. A file selector opens, in which you can browse to the location for the grid files and select multiple files. When you click OK, the grids are added to the Receptors for docking table, with the default Receptor ID values.

Once you have added the grids, you can edit them individually, by selecting a grid in the table and clicking Edit. The VSW - Receptor dialog box opens with the grid information loaded. You can do anything to edit the receptor that is described in [Section 6.1](#), including completely replacing it.

6.3 Setting Constraints

You can set up constraints for grids that are to be generated, you can request the use of constraints for both pregenerated and new grids, and you can set constraints on the ligand core position. All of these tasks are performed in the VSW - Receptor Constraints dialog box, which you open by clicking Constraints in the VSW - Receptor dialog box. You can define and use positional, NOE, H-bond, and metal constraints. Hydrophobic constraints cannot be defined or used here. You can choose whether or not to apply the constraints you have selected when you set up the docking options. When you have finished setting constraints, click OK.

To set up an H-bond or metal constraint for a new grid:

1. Select Pick atom to add H-bond or metal constraint.
2. Pick the receptor atom for the constraint in the Workspace.

The constraint is added to the table and marked in the Workspace. The type of constraint is determined from the type of the atom that you picked.

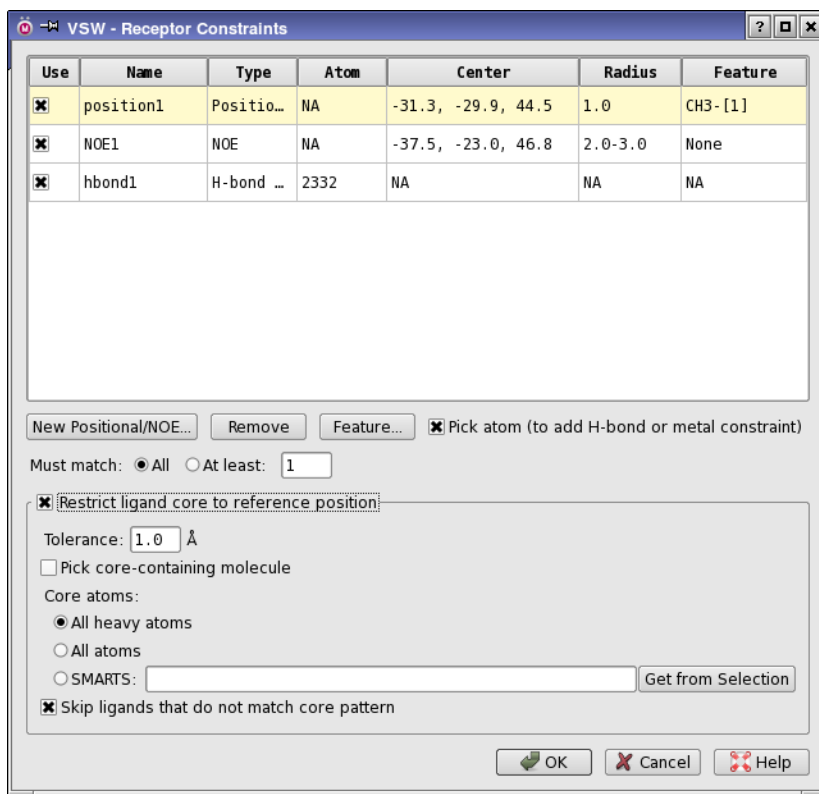


Figure 8. The VSW - Receptor Constraints dialog box.

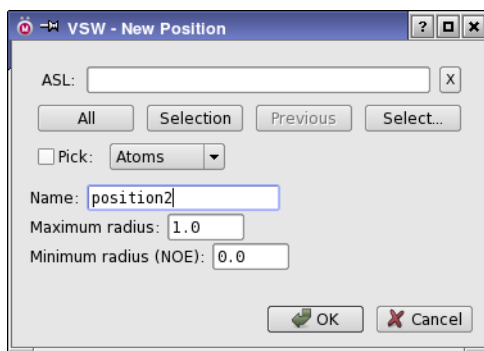


Figure 9. The New Position dialog box.

To set up a positional or NOE constraint for a new grid:

1. Click New Positional/NOE.
2. Use the picking controls in the New Position dialog box to pick the receptor atoms that define the constraint center.
3. Change the name from the default, if desired.
4. Enter a maximum radius for the constraint.
5. If defining an NOE constraint, enter a minimum radius.
6. Click OK.

The constraint is added to the table, with the appropriate type. You must also define a ligand feature for the constraint. See [Section 4.4.1](#) of the *Glide User Manual* for more information.

To define a ligand feature that must match a positional or NOE constraint:

1. Select the constraint in the table.
2. Click Feature to open the Edit Feature dialog box.
3. Click New to open the Edit Feature Pattern dialog box.

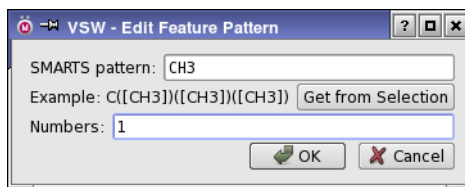


Figure 10. The Edit Feature Pattern dialog box.

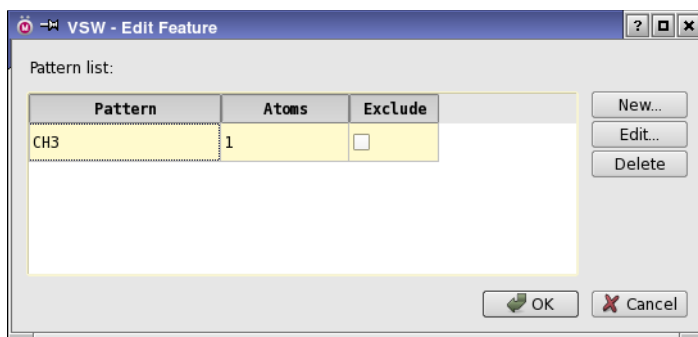


Figure 11. The Edit Feature dialog box.

4. Enter a SMARTS pattern in the SMARTS pattern text box, or click Get from Selection to create a SMARTS pattern from the selected atoms in the Workspace.
5. Enter the atom numbers of the atoms in the SMARTS pattern that you want to match the constraint, and click OK.
6. If you want to add more patterns, click New in the Edit Feature dialog box again, and repeat the above two steps to define the pattern. Patterns are matched in the order given in the pattern list.
7. If you want to exclude a particular pattern from matching the constraint, select the check box in the Exclude column. Excluded patterns are processed before other patterns.
8. Click OK.

The feature is added to the constraints table. You can edit the feature to add, edit, or delete patterns by selecting the constraint and clicking Feature. Features can be defined for both new and existing grids. See [Section 5.5.2](#) of the *Glide User Manual* for more information (but note also that there are differences between VSW and Glide).

To delete a constraint for a new grid:

1. Select the constraint in the table.
2. Click Remove.

To request use of constraints in docking:

1. Select the desired constraints in the Use column of the table.
2. Select a Must match option.
3. If you selected At least, enter the number of constraints that must be met in the text box.

To restrict the ligand core to a reference position:

1. Select Restrict ligand core to reference position.
2. Enter a tolerance for the RMSD of the ligand from the reference position in the Tolerance text box.
3. Select Pick core-containing molecule, and pick a ligand atom.

You must have the receptor and ligand displayed to pick the ligand atom.

4. Choose an option to define the core atoms in the ligand.

The choices are All (all atoms in the ligand), All heavy (all nonhydrogen atoms in the ligand), and SMARTS.

If you chose SMARTS, enter the SMARTS pattern in the text box, or select the desired core atoms in the Workspace, and click Get from selection.

5. If you want only the ligands that match the core pattern to be docked, select Skip ligands that do not match core pattern.

6.4 Aligning Grids

If you are generating grids, you can align the receptors before the grids are generated, using the Protein Structure Alignment tool. To do so, select **Align receptors**. This option does not align receptors for pregenerated grids, only those that are generated by the current VSW run. The receptors are aligned to the frame of reference of the first receptor in the table.

7 Setting Docking Options

If you specified grids in the Receptors tab, you can choose which of the three docking accuracy levels to include and set various options for docking in the Docking tab. The docking options are described in detail in [Chapter 5](#) of the *Glide User Manual*.

7.1 Setting Common Options

At the top of the Docking tab are several options that apply to all docking calculations. You can choose the force field from the Force Field option menu, from OPLS2001 and OPLS2005. The default is OPLS2005, because it supports a wider range of atom types, including boron, and it supports bonding to metals in a seamless fashion. The parameters in the Glide program and the Glide scoring functions were optimized for OPLS_2001, so use of another force field can result in a degradation of the results. OPLS_2005 produces very similar results to OPLS_2001.

Glide can also add ionization and tautomerization penalties from Epik to the docking score, including those for metal binding if they were calculated. These penalties are added by default. To exclude them, deselect Use Epik state penalties for docking.

You can soften the nonpolar part of the ligand potential by scaling the van der Waals radii of ligand atoms with small partial charges. To do so, enter the scaling factor and the partial charge cutoff in the text boxes under Scaling of ligand van der Waals radii for nonpolar atoms. See [Section 5.3.4](#) of the *Glide User Manual* for more information.

At the end of any docking run, you can run a Prime MM-GBSA calculation on the final poses to obtain ligand binding energies, by selecting Postprocess with Prime MM-GBSA. This option requires a Prime license.

7.2 Setting Up the Docking Stages and Options

The Docking tab allows you to choose which Glide docking runs to include in the workflow, and contains options for HTVS, SP, and XP docking runs. To include a docking run in the workflow, select the corresponding option. Next to each option is an Option button, which opens the VSW - Docking Options dialog box so that you can set options for the docking stage. A common set of controls is provided for each option, which are described below. In addition, there are mode-specific options, which are described after the common controls.

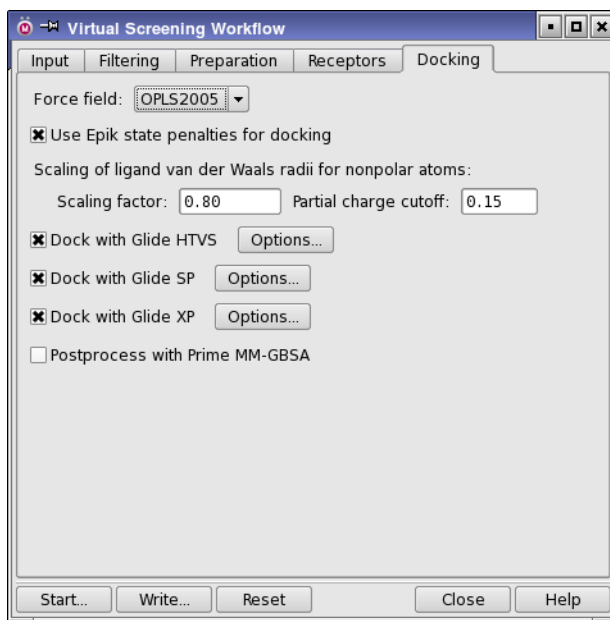


Figure 12. The Docking tab of the Virtual Screening Workflow panel.

7.2.1 Common Options

There are two common options for the docking method for all stages: Dock flexibly and Dock rigidly. For SP and XP docking, two more options are available: Refine and Score in place.

Under Dock flexibly, the Amide bonds option menu enables you to specify how to treat non-planar amide bonds. The options to choose from are:

- Vary amide bond conformation—Allow non-planar conformations, without a particular penalty.
- Allow trans conformers only—Only return results for conformations that are trans.
- Retain original amide bond conformation—No variation of the conformation is permitted.
- Penalize non-planar conformation—Apply a penalty to nonplanar amide bonds.

If you want to minimize the poses after docking, select Perform post-docking minimizations. This can result in better poses. When you perform this minimization, you can also select Apply

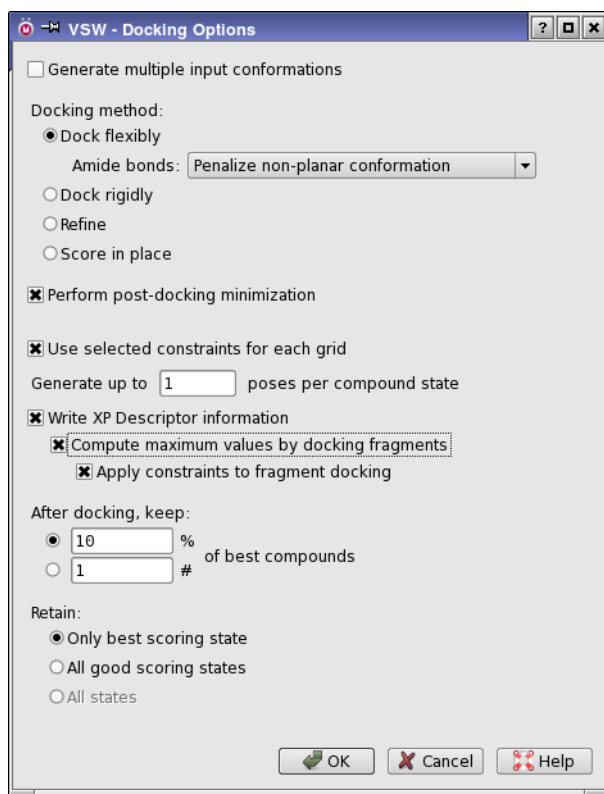


Figure 13. The VSW - Docking Options dialog box.

strain correction terms to apply strain corrections in the final scoring. These terms are evaluated by optimizing each ligand pose as a free ligand, first with constraints on all torsions, then without these constraints. The difference is used to compute a penalty term for unreasonably high strain.

To apply constraints for the docking stage, select **Use selected constraints for each grid**.

You can set a limit on the number of poses kept per compound state, by entering the number in the **Generate up to *N* poses per compound state** text box. This limit restricts the final results if you choose to retain all good scoring poses.

The two **After docking, keep** options allow you to specify the percentage (upper text box) or number (lower text box) of the best compounds to keep. A “compound” may consist of several ionization or tautomeric states, as generated by LigPrep. For each compound, you can choose how many ionization or tautomeric states to keep, by selecting an option under **Retain**:

- Only best scoring state (default for XP)
- All good scoring states (default for SP)
- All states (default for HTVS)

For XP docking, the last of these options is not available.

The rationale for keeping all ionization and tautomeric states is that the actual state that scores best can vary with the accuracy level. Keeping all states of a particular compound in the early stages ensures that the structures that will score best in later stages are not discarded.

7.2.2 Specific Options for XP Docking

For XP docking there are several additional options:

- **Generate multiple input conformations**—Run a MacroModel conformational search job on the input structures corresponding to the best poses from the previous stage to locate the lowest-energy conformer, using two different force fields. This option generates two extra input structures for XP docking, and therefore takes several times as long as an XP docking run with a single structure (including the conformational search). The variations in the input structures often produce better final results.
- **Write XP descriptor information**—This option writes the descriptor information needed for the XP Visualizer, and requires a license for this feature. See [Section 6.2](#) of the *Glide User Manual* for information on the XP Visualizer.
- **Compute maximum values by docking fragments**—This option is only available if you select **Write XP descriptor information**. It performs an additional XP docking calculation of a set of 50 fragments that were chosen to maximize the various XP descriptors. These fragments can be used to gain insight into the nature of the binding energetics. You can

view results for these fragments in the XP Visualizer, and use the results to evaluate how close the XP terms for ligands are to the maximum score. If you want to apply constraints to the fragments, select Apply constraints to fragment docking.

8 Running Jobs

When the general setup and the docking setup has been completed, click Start to open the Start dialog box, in which you can set job options and start the job. The run consists of a master (or “driver”) job and a set of subjobs. The master job starts all the subjobs for the various stages of the workflow, and collects the results. It is run on the local host. The master job must have access to the input files and the current directory, but the host for the subjobs does not need access to these files and directory.

Each stage of the workflow can be divided into a set of subjobs, which can be run concurrently. You can specify the number of subjobs in the Separate into N subjobs text box. You should consider all stages of the workflow when deciding on the number of subjobs. The master job may adjust the number of subjobs to better balance the load for each part of the workflow, so the number you enter is a target, not a requirement. If you are docking to more than one receptor, the workflow for each receptor can be distributed over multiple processors.

If you run multiple subjobs you can select a multiprocessor host such as a cluster. Each application can be run on a separate host, and has its own option menu to select the host. You can also choose to limit the number of processors allocated to the subjobs by entering the maximum number in the Use n processors text box for each application. The number of processors should not be more than the number of subjobs. For effective load balancing, the number of subjobs should be several times the number of processors. You can request that the number of subjobs is automatically adjusted to optimize the load balance and the time taken by each subjob by selecting the Adjust option (which is on by default).

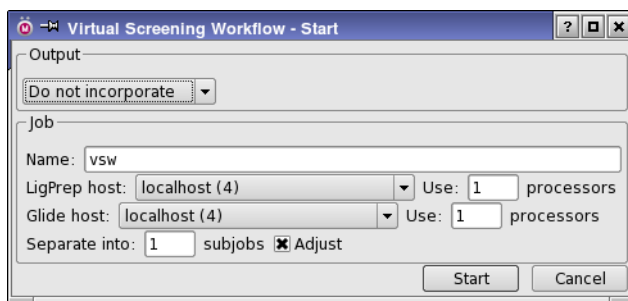


Figure 14. The Start dialog box.

The output of the job can be incorporated into Maestro by choosing Append new entries from the option menu in the Output section. The default is Do not incorporate, which means that the results will not be automatically incorporated into the Maestro project.

Jobs can be monitored and controlled from the Monitor panel—see [Section 3.2](#) of the *Job Control Guide*.

You can also run jobs from the command line by clicking Write to write the input file, then running the following command:

```
$SCHRODINGER/vsw [options] jobname.inp
```

The options are listed in [Table 1](#). The standard Job Control options are supported—see [Table 2.1](#) of the *Job Control Guide*. This includes the `-HOST` option, which is used to specify the hosts for the job. The `-WAIT` option, described in [Table 2.2](#) of the *Job Control Guide*, is also supported. The input file is described in detail in [Chapter 2](#) of the *Python Pipeline Manual*. This description covers all pipeline stages, not just those used by VSW.

Table 1. Options for the vsw command.

Option	Description
<code>-adjust</code>	Adjust the number of subjobs if the estimated job length is less than 10 minutes or more than 20 hours.
<code>-DRIVERHOST host</code>	Run the driver job on the specified host. By default, the driver (master) job runs on the local host.
<code>-h[elp]</code>	Print usage message and exit.
<code>-host_glide hosts</code>	Run Glide jobs on the specified hosts. Default: run on hosts specified by <code>-HOST</code> .
<code>-host_ligprep hosts</code>	Run LigPrep jobs on the specified hosts. Default: run on hosts specified by <code>-HOST</code> .
<code>-host_mmod hosts</code>	Run MacroModel jobs on the specified hosts. Default: run on hosts specified by <code>-HOST</code> .
<code>-host_prime hosts</code>	Run Prime jobs on the specified hosts. Default: run on hosts specified by <code>-HOST</code> .
<code>-host_qikprop hosts</code>	Run QikProp jobs on the specified hosts. Default: run on hosts specified by <code>-HOST</code> .
<code>-keywords</code>	List the keywords available for all stages of the workflow and exit.
<code>-LOCAL</code>	Run the driver job in local directory (default if the driver is run on the local host).
<code>-local</code>	Do not create a temporary directory for each subjob.

Table 1. Options for the vsw command. (Continued)

Option	Description
-max_retries <i>n</i>	Maximum number of times to restart subjobs if they fail. If not specified, the value specified by SCHRODINGER_MAX_RETRIES value is used, if defined, otherwise the default is 2.
-NJOBS <i>n</i>	Number of subjobs to generate without adjusting. If not specified, the number of subjobs is set to the number of processors and the -adjust option is set.
-no_cleanup	Do not remove intermediate files.
-NOLOCAL	Run the driver job in the scratch directory (default if the driver is run on a remote host). Jobs run with this option cannot be restarted.
-OVERWRITE	Overwrite any existing files when running the job.
-REMOTEDRIVER	Run the driver job on the first host in the list specified by -HOST.
-RESTART	Restart the job. Restarting runs any subjobs that did not finish in the previous execution of the job.
-v	Display the version number and exit.

9 Restarting Jobs

To restart a job, run it from the command line with the -RESTART option:

```
$SCHRODINGER/vsw -RESTART [options] jobname.inp
```

If you omit the -RESTART option, you will be prompted to indicate whether to resume the job where it left off, to restart it from the beginning, or to exit without doing anything. You can also reset the host and CPU information by setting the appropriate options.

If you want to rerun a job and overwrite the output from the previous execution of the same job, use the -OVERWRITE option:

```
$SCHRODINGER/vsw -OVERWRITE [options] jobname.inp
```

10 Output Files

Output files are created in subdirectories of the working directory. The path to the output files is given at the end of the log file. For each receptor, a pose viewer file (`_pv.mae`) is created, containing the receptor and the docked ligands. In addition, the docked ligands are merged into a single file, with properties resulting from the merge, including the GlideScore offset, the receptor used, and the adjusted score. The merge is done by the `glide_ensemble_merge` utility, which is described below. The receptors are also placed in a single file.

11 Merging Results of Multiple Jobs

Although you can include multiple receptors in a single run of the workflow, you may want to merge the results of several runs of the workflow into a single Glide pose viewer file. You can do this with the utility `glide_ensemble_merge`. In the process, you can specify a scoring offset for calibration of different runs. The results need not be from a single receptor: you can merge results for multiple receptors. The script combines sorted Glide pose viewer (`*_pv.mae`) files into an output file, sorted by default by GlideScore.

The syntax is as follows:

```
$SCHRODINGER/utilities/gl意思ide_ensemble_merge [options]
    job1_pv.mae[:offset1] job2_pv.mae[:offset2]
    [job3_pv.mae[:offset3] ... ]
```

The options are listed in [Table 2](#). The list of files to be merged is a blank-separated list of pose-viewer file names. Each file name can be followed by a colon and an offset for the GlideScore. The default offset is zero.

Table 2. Options for the glide_ensemble_merge utility.

Option	Description
<code>--version</code>	Display program version number and exit
<code>-h, --help</code>	Display usage message and exit
<code>-n maxlig,</code> <code>--nreport= maxlig</code>	Maximum number of best ligands to keep after merging. Default: 10000.
<code>-j jobname,</code> <code>--jobname= jobname</code>	Job name. The output file is named <code>jobname_pv.mae</code> . Default name is <code>glide_ensemble_merge</code> .
<code>-m mpl,</code> <code>--max_per_lig=mpl</code>	Maximum poses per ligand (default 1). Specify 0 (zero) to save all poses.
<code>-f sortfield,</code> <code>--field=sortfield</code>	Field (property value) by which the input files are sorted. Default is the GlideScore (<code>r_i_glide_gscore</code>).

12 Citing the Virtual Screening Workflow in Publications

Schrödinger Suite 2012 Virtual Screening Workflow; Glide version 5.7, Schrödinger, LLC, New York, NY, 2012; LigPrep version 2.5, Schrödinger, LLC, New York, NY, 2012; QikProp version 3.4, Schrödinger, LLC, New York, NY, 2012.

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/>, particularly the Support Center, <http://www.schrodinger.com/supportcenter>, and the Knowledge Base, <http://www.schrodinger.com/kb>.

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 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 your browser.
- Choose **Help → Online Help** or press CTRL+H (⌘H) to open the default help topic in your browser.
- When help is displayed in your browser, use the navigation links or search the help in the side bar.
- Choose **Help → Manuals Index**, to open a PDF file that has links to all the PDF 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*.
- Software updates: choose Maestro → Check for Updates.
- New software features: choose Help → New Features.
- Scripts available for download: choose Scripts → Update.
- 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.

E-mail: help@schrodinger.com

USPS: Schrödinger, 101 SW Main Street, Suite 1300, Portland, OR 97204

Phone: (503) 299-1150

Fax: (503) 299-4532

WWW: <http://www.schrodinger.com>

FTP: <ftp://ftp.schrodinger.com>

Generally, e-mail correspondence is best because you can send machine output, if necessary. When sending e-mail messages, please include the following information:

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

Gathering Information for Technical Support

This section describes how to gather the required machine, licensing, and installation information, and any other job-related or failure-related information, to send to technical support.

For general enquiries or problems:

1. Open the Diagnostics panel.
 - **Maestro:** Help → Diagnostics
 - **Windows:** Start → All Programs → Schrodinger-2012 → Diagnostics
 - **Mac:** Applications → Schrodinger2012 → 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. Attach the file specified in the dialog box to your e-mail message.

If your job failed:

1. Open the Monitor panel in Maestro.

Use Applications → Monitor Jobs or Tasks → Monitor Jobs.
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 in your working directory, and an information dialog box with the name of the file opens. You can highlight and copy the name of the file.
5. Attach the file specified in the dialog box to your e-mail message.
6. Copy and paste any log messages from the window used to start Maestro (or the job) into the email 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-2012 → Diagnostics
- **Mac:** Applications → Schrodinger2012 → 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. Attach the file specified in the dialog box to your e-mail message.

4. Attach the file `maestro_error.txt` to your e-mail message.

This file should be in the following location:

- **Windows:** %LOCALAPPDATA%\Schrodinger\appcrash
(Choose Start → Run and paste this location into the Open text box.)
- **Mac:** Documents/Schrodinger
- **Linux:** Maestro's working directory specified in the dialog box (the location is given in the terminal window).

5. On Windows, also attach the file `maestro.EXE.dmp`, which is in the same location as `maestro_error.txt`.

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

155 Gibbs St
Suite 430
Rockville, MD 20850-0353

Quatro House
Frimley Road
Camberley GU16 7ER
United Kingdom

101 SW Main Street
Suite 1300
Portland, OR 97204

Dynamostraße 13
D-68165 Mannheim
Germany

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

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

Zeppelinstraße 73
D-81669 München
Germany

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

8910 University Center Lane
Suite 270
San Diego, CA 92122

Potsdamer Platz 11
D-10785 Berlin
Germany

SCHRÖDINGER®