Induced Fit Docking

Schrödinger Suite 2012 Update 2

Schrödinger Press

Revision A, September 2012

Contents

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].

Introduction

This document provides information about the Schrödinger Induced Fit Docking (IFD) protocol, which uses Glide and Prime to induce adjustments in receptor structures, and the Python script that has been developed to automate the process.

Following this introduction, succeeding chapters provide:

- An Induced Fit Docking tutorial in Chapter 2.
- A description of structure preparation tasks in Chapter 3.
- A description of the Induced Fit Docking panel in Chapter 4.
- Information on using the IFD protocol from the command line in Chapter 5.

For the most up-to-date information, check our web site, http://www.schrodinger.com.

1.1 Induced Fit Docking Applications

In standard virtual docking studies, ligands are docked into the binding site of a receptor where the receptor is held rigid and the ligand is free to move. However, the assumption of a rigid receptor can give misleading results, since in reality many proteins undergo side-chain or backbone movements, or both, upon ligand binding. These changes allow the receptor to alter its binding site so that it more closely conforms to the shape and binding mode of the ligand. This is often referred to as “induced fit” and is one of the main complicating factors in structure-based drug design.

The ability to model induced-fit docking has two main applications:

- Generation of an accurate complex structure for a ligand known to be active but that cannot be docked in an existing (rigid) structure of the receptor.
- Rescue of false negatives (poorly scored true binders) in virtual screening experiments, where instead of screening against a single conformation of the receptor, additional conformations obtained with the induced fit protocol are used.

1.2 The Induced Fit Docking Protocol

Schrödinger has developed and validated an Induced Fit Docking protocol based on Glide and the Refinement module in Prime that accurately predicts ligand binding modes and concomitant structural changes in the receptor.

The Schrödinger IFD protocol models induced fit docking of one or more ligands using the following steps:

1. An optional constrained minimization of the receptor (protein preparation, refinement only) with an RMSD cutoff of 0.18 Å. Normally this is done when preparing the protein with the Protein Preparation Wizard.
2. Initial Glide docking of each ligand using a softened potential (van der Waals radii scaling), and optional removal of side chains and application of constraints. By default, a maximum 20 poses per ligand are retained, and by default poses to be retained must have a Coulomb-vdW score less than 100 and an H-bond score less than –0.05.
3. Prime side-chain prediction for each protein/ligand complex, on residues within a given distance of any ligand pose (default 5 Å), with optional inclusion or exclusion of other residues, and an optional implicit membrane model.
4. Prime minimization of the same set of residues and the ligand for each protein/ligand complex pose. The receptor structure in each pose now reflects an induced fit to the ligand structure and conformation.
5. Glide redocking of each protein/ligand complex structure within a specified energy of the lowest-energy structure (default 30 kcal/mol). The ligand is now rigorously docked, using default Glide settings, into the induced-fit receptor structure.
6. Estimation of the binding energy (IFDScore) for each output pose.

The induced fit docking process is run with a Python script, which runs the specified stages of the protocol. You can specify the structures and enter settings for various options, and then start the job running from Maestro. The script then completes the protocol without further intervention. The tutorial in Chapter 2 will guide you through the process of entering settings, launching the job, and examining the results. You can also customize the process by editing the input file.

The structures you use for induced-fit docking must be prepared in the same manner as for Glide. The protein and ligand preparation must precede the use of the protocol outlined above. For details on protein and ligand preparation, see Chapter 3 of the *Glide User Manual* and the *Protein Preparation Guide* and *LigPrep User Manual*.

The Induced Fit Docking protocol can also be run from the command line, and you can customize the protocol to perform the Glide and Prime steps of your choice. Chapter 5 describes the input file and how to use it for customization of the protocol.

1.3 Sample Results

In studies of 14 ligand-receptor pairs that required induced fit docking, the Schrödinger Induced Fit Docking protocol yielded an average heavy-atom RMSD of 1.2 Å for the top-ranked output ligand pose to the native ligand. In contrast, rigid-receptor docking with the same ligand-receptor pairs yielded eight cases in which a pose could not be found and an average RMSD of 6.1 Å for the remaining six pairs. Targets included aldose reductase, CDK2 (2), estrogen receptor, HIV protease, protein kinase B, PPAR-gamma, LXR-beta, and thymidine kinase.

1.4 Installation

To run the Schrödinger Suite 2012 Induced Fit Docking protocol, you must install Prime 3.1 and Glide 5.8. To use the automated protocol from Maestro, you must also install Maestro 9.3. Induced Fit Docking using Glide 5.8 and Prime 3.1 is supported on Linux, Windows and Mac platforms. For installation instructions and information on platform support and hardware and software requirements, see the *Installation Guide*.

1.5 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"
```

1.6 Citing Induced Fit Docking in Publications

The use of this protocol should be acknowledged in publications as:

Schrödinger Suite 2012 Induced Fit Docking protocol; Glide version 5.8, Schrödinger, LLC, New York, NY, 2012; Prime version 3.1, Schrödinger, LLC, New York, NY, 2012.

Please also cite the following publications:

- Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R., "Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects," *J. Med. Chem.*, **2006**, *49*, 534.
- Sherman, W.; Beard, H. S.; Farid, R., "Use of an Induced Fit Receptor Structure in Virtual Screening," *Chem. Biol. Drug Des.*, **2006**, *67*, 83.

Induced Fit Docking Tutorial

The tutorial in this chapter demonstrates the use of the Schrödinger Induced Fit Docking protocol. Starting with the receptor structure of a protein complexed with a ligand, you will dock a different known active ligand to the active site. The Induced Fit Docking protocol generates multiple poses of the ligand complex, each including unique structural modifications of the receptor to fit the ligand pose, and ranks these poses by GlideScore to find the best structure of the docked complex.

The protein used is human cyclin-dependent kinase 2 (CDK2). The structure of the receptor is derived from the PDB entry 1dm2. The native ligand in the 1dm2 structure is the inhibitor hymenialdisine (HMD); the new ligand that will be docked to that receptor is staurosporine.

This example was chosen as an introduction to the mechanics of using the Induced Fit Docking protocol and is not intended as a research study. The receptor structure provided with the tutorial has been truncated to reduce the time taken by the calculations.

For the purposes of this tutorial, the protein and ligand structures that are provided have already been prepared for Induced Fit Docking. In real applications, you must prepare the protein and the ligands to ensure that they are all-atom structures with correct bond orders and formal charges.

The parameters used in the tutorial have been selected so that the tutorial runs in a relatively short time. In real applications, the default parameters give good results in a very large majority of cases. If you use the default parameters with this tutorial, the Induced Fit Docking job can take approximately 9 CPU hours on a 2 GHz Pentium 4 processor. With the parameters in the tutorial, the total time is about 25 minutes on the same processor.

It is assumed that you already have access to an installation of Maestro 9.3, Glide 5.8, Prime 3.1, and supporting third-party programs and databases (PDB, BLAST, HMMER/Pfam). For installation instructions, see the *Installation Guide*.

2.1 Preparing for the Exercises

To run the exercises, you need a working directory in which to store the input and output, and you need to copy the input files from the installation into your working directory. This is done automatically in the Tutorials panel, as described below. To copy the input files manually, just unzip the `ifd` zip file from the `tutorials` directory of your installation into your working directory.

On Linux, you should first set the SCHRODINGER environment variable to the Schrödinger software installation directory, if it is not already set:

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

sh/bash/ksh: `export SCHRODINGER=`*installation-path*

If Maestro is not running, start it as follows:

- **Linux:** Enter the following command:

  ```
  $SCHRODINGER/maestro -profile Maestro &
  ```

- **Windows:** Double-click the Maestro icon on the desktop.

  You can also use Start → All Programs → Schrodinger-2012 → Maestro.

- **Mac:** Click the Maestro icon on the dock.

  If it is not on the dock, drag it there from the SchrodingerSuite2012 folder in your Applications folder, or start Maestro from that folder.

Now that Maestro is running, you can start the setup.

1. Choose Help → Tutorials.

   The Tutorials panel opens.

2. Ensure that the Show tutorials by option menu is set to Product, and the option menu below is labeled Product and set to All.

3. Select Induced Fit Docking Tutorial in the table.

4. Enter the directory that you want to use for the tutorial in the Copy to text box, or click Browse and navigate to the directory.

   If the directory does not exist, it will be created for you, on confirmation. The default is your current working directory.

5. Click Copy.

   The tutorial files are copied to the specified directory, and a progress dialog box is displayed briefly.

If you used the default directory, the files are now in your current working directory, and you can skip the next two steps. Otherwise, you should set the working directory to the place that your tutorial files were copied to.

6. Choose Project → Change Directory.

7. Navigate to the directory you specified for the tutorial files, and click OK.

You can close the Tutorials panel now, and proceed with the exercises.

2.2 Importing the Receptor

1. Click the Import button on the Project toolbar.

2. In the Import panel, select the file `IFD_receptor.mae`.

   Navigate to your working directory if necessary.

3. Click Options.

   The import options are displayed in the panel.

4. Ensure that the Include in Workspace option selected is First Imported Structure.

5. Click Open.

   The receptor-ligand complex appears in the Workspace, as shown in Figure 2.1. The ligand is a nonstandard residue and is therefore colored differently from the receptor. You must have the receptor displayed in the Workspace to set up the job.

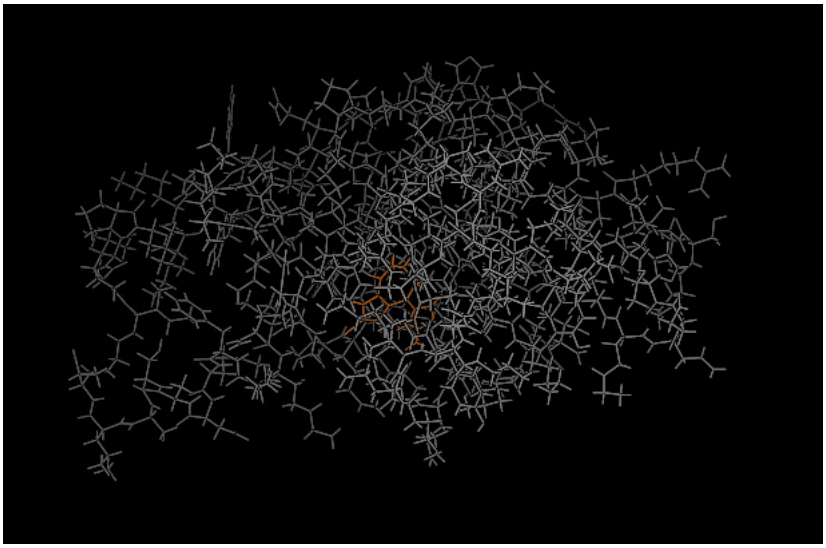

Figure 2.1. The receptor complex

2.3 Setting Up the Induced Fit Docking Job

- Choose Workflows → Induced Fit Docking or Tasks → Docking → Induced Fit Docking in the main window.

  The Induced Fit Docking panel opens with the Receptor tab displayed, as shown in Figure 2.2.

2.3.1 Specifying the Ligand To Be Docked

1. From the Ligands to be docked option menu, choose File.

   The Input file text box and Browse button are now available.

2. Click Browse.

   A file selector opens, showing the contents of your working directory.

3. Open the file containing the ligand, `IFD_ligand.mae`.

2.3.2 Defining the Receptor and Enclosing Box

The receptor must be distinguished from the complexed ligand in order for the Glide grid generation portion of the protocol to run correctly, and the parameters of the grid box defined.

1. Ensure that the Box center option selected is Centroid of Workspace ligand.

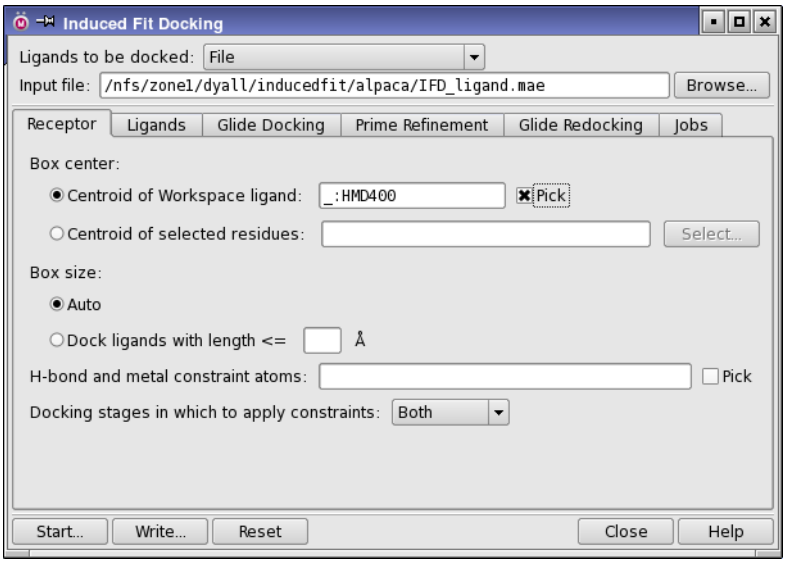

Figure 2.2. The Receptor tab of the Induced Fit Docking panel.

2. Select Pick to the right of the Centroid of Workspace ligand text box.
3. Click on a ligand atom in the Workspace.

   The ligand is colored orange. When you pick the ligand, it is listed in the text box, and the grid center, the ligand center box (green) and the grid box (purple) are displayed in the Workspace. The ligand is marked with green markers.
4. Ensure that the Box size option selected is Auto (the default).

   The position and size of the grid box (or *enclosing* box) are defined automatically, based on the selected ligand.

You can also specify hydrogen-bond and metal constraints, and choose the docking stage to which they apply. In this tutorial, constraints are not used.

2.3.3 Specifying Initial Glide Docking Options

In the Glide Docking tab, you specify the refinement phase of protein preparation, choose whether to temporarily remove active-site residue side chains, and select options for the first round of Glide ligand docking. This preliminary docking is typically performed with both the receptor and the ligand "softened" by van der Waals radii scaling. By default, the scaling factor is 0.50 for the receptor and 0.50 for the ligand. The Ligands tab has settings for the ligand conformations, but in this exercise, the defaults are used.

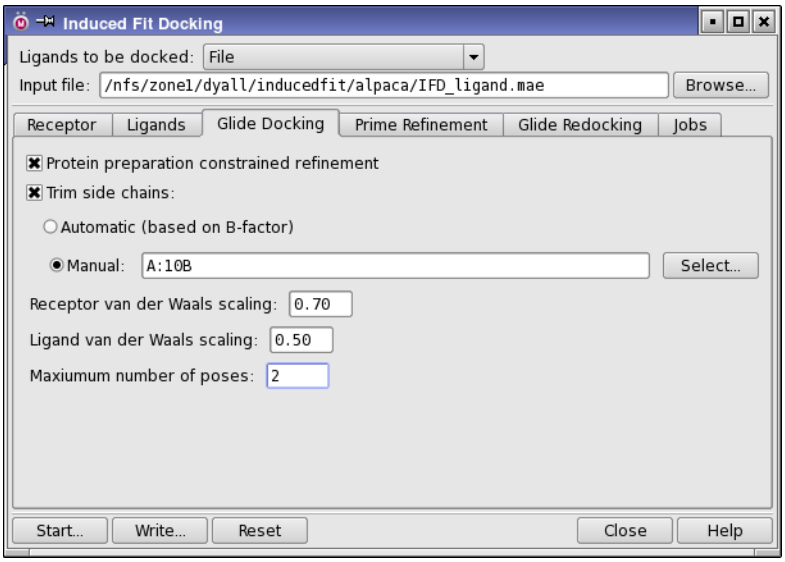

Figure 2.3. The Glide Docking tab of the Induced Fit Docking panel.

1. Select Protein Preparation constrained refinement.

   This option is not selected by default, and is not needed if you include it when you use the Protein Preparation Wizard to prepare your protein.

2. Select Trim side chains.

   The Receptor van der Waals scaling factor is automatically changed to 0.70. Removing side chains from active-site residues provides more room for ligand docking, so the receptor does not need to be quite as soft. The side chains are restored after docking.

3. Select Manual, and click the Select button to the right of the text box.

   The Atom Selection dialog box opens.

4. Click Clear under the ASL text box to clear the previous selection if necessary.

5. In the Residue folder, select Sequence.

6. In the Entry (Chain) list, select 1(A).

7. In the Sequence list, select ILE 10 B.

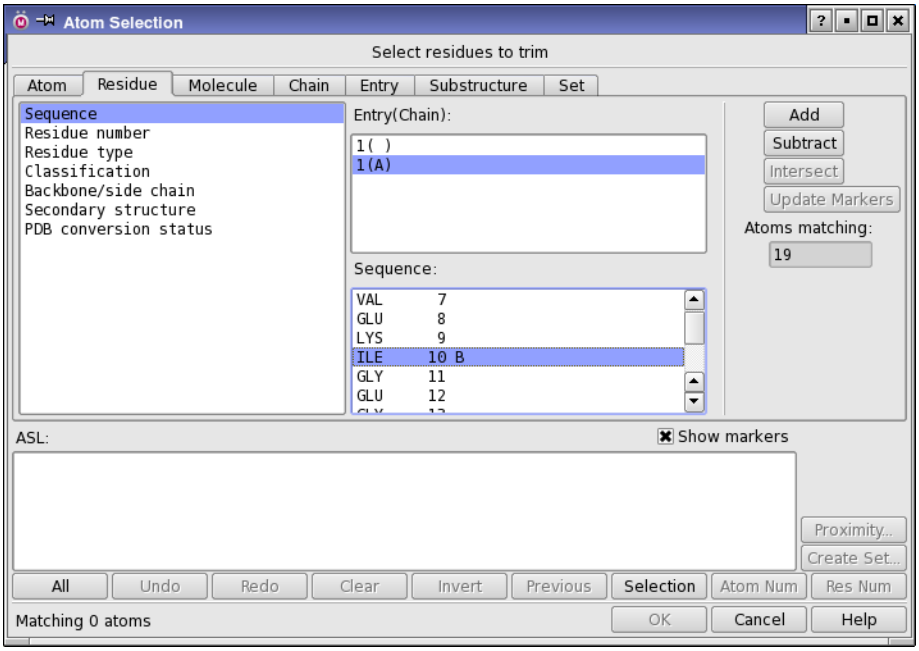

Figure 2.4. The Atom Selection dialog box showing the residue selection.

8. Click Add, then click OK.

   Residue 10, isoleucine, is selected for side-chain removal. This residue is temporarily mutated to alanine during the initial Glide docking.

9. Enter 2 in the Maximum number of poses text box.

   Normally you would leave this value at the default. This choice is solely to obtain results in a fairly short time.

2.3.4 Specifying Prime Induced Fit Options

In the Prime Refinement tab, you will reduce the distance from the ligand that defines residues for refinement. This is done solely to speed up the calculations. In real applications, you should not in general reduce this distance below the default of 5.0 Å.

1. Enter `3.4` in the Refine residues within *m* Å of ligand poses text box.
2. Ensure that Optimize side chains is selected.

The remaining options in this tab and the Glide Redocking tab are left at their default values. Most of the changed values are for the purpose of shorter execution time. In real applications, you would not make these settings. You are now ready to run the job.

2.4 Running the Induced Fit Docking Job

The Induced Fit Docking protocol is basically a series of Glide and Prime jobs. By default, the job is run serially on your local machine. In the Jobs tab you can specify the number of processors used for each kind of job. The maximum number of processors you should use is the number of poses, which in this tutorial is 2.

1. If you want to choose a multiprocessor host, enter 2 in the Number of Glide CPUs text box and the Number of Prime CPUs text box in the Jobs tab.
2. Click Start.

   A dialog box opens, warning that the protein has not been prepared with the Protein Preparation Wizard. In this tutorial, a constrained minimization is being run on a structure that is otherwise well defined, so you can go ahead with the IFD job.

3. Click Continue.

   The Induced Fit Docking - Start dialog box opens.

4. Change the job name to `InducedFit1`.
5. Select a host from the Host menu.

6. Click Start.

   The following files and directories should be present in your working directory:

```
InducedFit1.inp        InducedFit1_lig.mae    InducedFit1.log
InducedFit1_rec.mae    InducedFit1.restart    InducedFit1_workdir/
```

The job log is written to *jobname*`.log`. It lists the input parameters and the ligands to be docked, then reports progress on the job stages of the protocol. The job stages are described in Table 5.2 on page 34. The log also lists the subjobs that the job launches and the output files they produce. You can monitor the progress of the job in the Monitor panel. An example of the log file for this tutorial is shown in Section 5.3 on page 46. The total time for this job on a 1.8 GHz Pentium 4 processor with 256 kB cache was about 30 minutes.

2.5 Viewing Results

In ordinary flexible-ligand Glide docking, a "pose" or "ligand pose" is a particular conformation of the ligand with respect to the receptor. Because ordinary Glide uses a rigid receptor, each pose combines a unique ligand conformation with an identical receptor structure. In the context of rigid-receptor docking, the term "pose" is equivalent to "ligand pose." In Induced Fit Docking, both the ligand and the receptor conformation are different for each pose. In the context of induced fit docking, the term "pose" is used for each unique complex structure.

The output structures from the Induced Fit Docking job are stored in a compressed Maestro file, *jobname*`-out.maegz`.

1. Click the Import button on the Project toolbar.

   

   The Import panel opens.

2. Choose Maestro (or Common) from the Files of type option menu.
3. Navigate to your working directory and select `InducedFit1-out.maegz`.
4. Click Open.

   The first pose from the output file is displayed in the Workspace. By default, only the ligand and the residues that were refined are displayed. If you cannot see the structure, click the Fit button.

   

If the Project Table panel is not displayed, click the Table button on the Project toolbar.

The input structure, with the native hymenialdisine ligand, and the best staurosporine pose are displayed in Figure 2.5. To display images like these, follow the instructions below.

5. If the Display Atoms toolbar is not visible, click Display Atoms on the Manager toolbar, or choose Window → Toolbars → Display Atoms.
6. Choose Molecules from the Display Sel button menu on the Display Atoms toolbar.

   

7. Pick an atom in the ligand.

   The protein residues are undisplayed, and only the ligand is visible.
8. Choose +6 Å from the Within button menu on the Display Atoms toolbar.

   

   Protein residues within 6 Å of the ligand are displayed. There are now sufficient residues displayed to see the context of the ligand, but not too many to obscure the view.
9. Click the Apply button on the Style toolbar.

   

   If the Style toolbar is not displayed, click Style on the Manager toolbar or choose Window → Toolbars → Style.

   The representation of the pose in the Workspace is updated to highlight the ligand and the binding site residues.

   The Style default representation settings can be adjusted by choosing Settings from the Apply button menu. For more information about the Workspace Style toolbar, see Section 7.6 of the *Maestro User Manual*.
10. Include the CDK2-frag entry in the Workspace and repeat the instructions starting from Step 6.
11. Add the first staurosporine pose to the Workspace (control-click the In column).

    Both entries should now be in the Workspace.

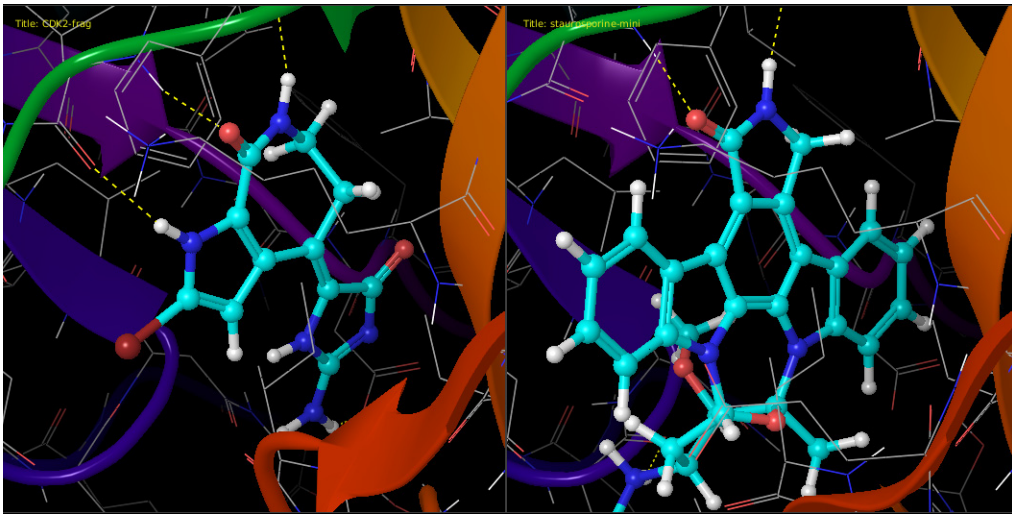

Figure 2.5. The native hymenialdisine structure (left) and the best staurosporine structure (right).

12. Click the Tile button on the Workspace toolbar.

13. The structures are displayed side-by-side. You will be able to see the details better if you expand Maestro to fill the screen.
14. Click the Contacts button on the Style toolbar.

15. The Contacts and H-Bonds dialog box opens.
16. Select Display receptor-ligand hydrogen bonds, and click OK.

    The hydrogen bonds between the receptor and ligand are shown as dashed yellow lines.
17. Zoom in on the ligand, translate and rotate it to a suitable view.

    The view changes are applied to both structures simultaneously.

For more information on changing how structures are represented in the Workspace, see Chapter 7 of the *Maestro User Manual*.

Preparing Structures

Before you run an Induced Fit Docking job, you must prepare the receptor and the ligands. Instructions for these preparation tasks are given below.

3.1 Preparation of the Receptor

Proper preparation of the protein or protein-ligand complex to be used as the receptor is critical to the success of Induced Fit Docking. Both Glide and Prime have certain requirements, and in addition there are requirements for Induced Fit Docking.

In general, the Induced Fit Docking procedure requires a complete, all-atom structure (explicit hydrogens present) with correct bond orders and formal charges. If you are starting with a typical PDB structure (heavy atoms only), the hydrogen atoms are usually implicit, and there may be missing atoms or residues in the structural information, and atom or residue labels that are incorrect. All of these issues must be addressed before proceeding.

When you import a PDB structure into Maestro, it is color coded according to various problems detected in the input data. You can use these color codes to identify and fix the structure. For more information, see Section 3.1.6 of the *Maestro User Manual*. Apart from problems indicated by the color coding, other structural problems may exist that must be fixed. These problems are often only detected by careful inspection of the structure, particularly in the active site.

Maestro provides the means to automatically fix most of these problems in the Protein Preparation Wizard panel, which you open from the Workflows menu. A description of how to use this panel is given in Chapter 2 of the *Protein Preparation Guide*. Procedures for manually fixing structures, including problems that are not fixed by the Protein Preparation Wizard, are given in Chapter 3 of the *Protein Preparation Guide*.

You do not need to perform the refinement step in the Protein Preparation Wizard, as this part can be done in the Induced Fit Docking procedure.

The general steps in the protein preparation procedure for Induced Fit Docking are:

1. Import the PDB protein structure.
2. Examine the structure for problems, including noting the color code.
3. Fix bond orders, formal charges, and atom names.

The ligand residues are usually the ones that are colored orange. Fixing the ligand is a prerequisite for any Prime refinement calculation that you may do to fix other problems. This is not the same as preparation of the ligands for docking, which is treated in the next section. You must fix the bond orders, formal charges, and PDB atom names before you can run Prime.

Check also for groups that you expect to have formal charges, to ensure that these charges are correct. Note that the symmetry of nitro and carboxylate groups is automatically accounted for. You should assign the formal charges and bond orders according to the Lewis structure.

4. Fix residues that are missing atoms with a Prime side-chain prediction.

   These residues are colored red on PDB import. See Section 3.6 of the *Protein Preparation Guide* for instructions on fixing these residues. When you come to setting up the Induced Fit Docking job, you should also consider selecting any of these residues that are close to the active site for mutation—see Section 4.5 on page 25.

5. Check for missing residues.

   The residues at the breaks are not usually color-coded. If there are breaks, you will need to do a Prime calculation to predict the structure of the missing residues. See Section 3.7 of the *Protein Preparation Guide*.

6. Check the active site for incorrect side-chain geometry, protonation state or tautomerization, and fix as appropriate.

   This task is performed by the Wizard; to perform the changes manually, see Section 3.8 and Section 3.9 of the *Protein Preparation Guide*.

7. Run a Prime Energy calculation to check that all the problems were fixed.

   If the Prime energy does not look reasonable, it is likely that some problems have not been fixed. You should then inspect the protein for possible remaining problems.

To run a Prime energy calculation:

1. From the Prime submenu of the Applications menu, choose Refinement.

   The Prime Refinement panel opens.

2. From the Task option menu, choose Energy Analysis.
3. Click Start.

   The Start dialog box opens.

4. Make any job settings, then click Start.

3.2 Preparation of the Ligands To Be Docked

Each ligand that will be docked to the receptor must also meet certain requirements. Like the receptor, it must have correct bond orders, formal charges, and a complete set of hydrogens for a valid ionization state. You can run the ligands through the Schrödinger application LigPrep to produce one or more desired ligand conformations and ionization states. See the *LigPrep User Manual* for information on this application.

In addition to the structure preparation, the names of the atoms in each ligand to be docked must satisfy two conditions:

- All atoms in the ligand must have the same PDB Residue Name, Residue Number, and Chain Name. This condition is satisfied automatically during job execution, so you do not need to do anything. The program sets the PDB residue name to "UNK", the chain name to Z, the residue number to 999, and the insertion code to blank. These values ensure that there is no conflict with the receptor.
- All atoms must have PDB Atom Names that are unique within the ligand residue, as required for parameter generation. Prime attempts to ensure that atoms have unique PDB atom names, but if it does not succeed in this task, the job will fail. It is therefore highly recommended to correct the ligand if it does not satisfy this condition. If it does not, you must first ensure that the ligand is a single residue.

To check whether a ligands satisfies these conditions and correct it if it does not, use the procedures below. Before doing so, ensure that only the ligand is displayed in the Workspace.

The ligands to be docked must be in a single Maestro file. If you have prepared ligands in Maestro, you should export them to a Maestro file.

To check that all atoms in the ligand residue have the same residue information:

1. Choose Workspace > Atom Labels.

   The Atom Labels panel opens.

2. In the Composition folder, clear all selections, then select Residue name, Residue number, and Chain name.
3. Ensure that Add labels is selected.
4. Click All in the Label Atoms section.

The labels are displayed, which you can examine to ensure that they are the same.

To correct the residue information:

1. Choose Edit > Build > Residue Properties.

   The Build panel opens at the Residue Properties folder.

2. Choose the property from the Property option menu.
3. Enter a value in the appropriate text box.
4. Click All.

To check that all atoms have unique PDB atom names within the residue:

1. Include only the ligand in the Workspace.
2. Choose Workspace > Atom Labels.

   The Atom Labels panel opens.

3. In the Composition folder of the Atom Labels panel, clear all selections, then select PDB atom name.
4. Ensure that Add labels is selected.
5. Click All in the Label Atoms section.

To correct ligands that do not have unique PDB atom names:

1. Open the Build panel.
2. In the Atom Properties folder, choose PDB Atom Name from the Property option menu.
3. In the Set unique PDB atom names within residues section, click All.

Running Induced Fit Docking from Maestro

The Induced Fit Docking protocol is run from Maestro using the Induced Fit Docking panel. To open the panel:

- **Maestro:** Choose Workflows → Induced Fit Docking or Tasks → Docking → Induced Fit Docking.
- **BioLuminate:** Choose Tasks → Ligand Tasks → Docking → Induced Fit Docking.

Before you run the Induced Fit Docking protocol, you must prepare the protein and the ligands. Instructions for these preparation tasks are given in Chapter 3.

To run the protocol on a receptor, the receptor must be displayed in the Workspace. You should ensure that only the desired protein structure or protein-ligand complex is included in the Workspace. If you are using a protein-ligand complex, you must ensure that the complex is a single entry. If it is not, choose Entry > Merge in the Project Table panel to merge the ligand and the receptor into a single entry, and display the merged entry. Pose viewer files from Glide, for example, have the receptor and the ligands in separate entries, so the receptor entry must be merged with the chosen ligand entry.

The features available in the Induced Fit Docking panel are described in the following subsections. Along with each panel feature, some related details about using the protocol are discussed.

If you want to run the Induced Fit Docking protocol from the command line, see Chapter 5. The command line also allows you to run the protocol on more than one receptor.

4.1 General Panel Layout

The Induced Fit Docking panel is divided into six tabs, Receptor, Ligands, Glide Docking, Prime Refinement, Glide Redocking, and Jobs. In the upper part of the panel you can specify the ligands to be docked.

Below the tabs are three buttons:

- Start—Starts the induced fit docking job.
- Write—Generates input files that can be used with the `ifd` command to launch induced fit docking from the command line. Does not start the job.
- Reset—Resets all the settings in the Induced Fit Docking panel to their defaults.

4.2 Ligand Input

At the top of the panel you can specify the source of the ligands to be docked, by choosing one of the items from the Ligands to be docked option menu:

- File—Enter the file name in the Input file text box, or use the Browse button to open a file selector and navigate to the file. The file must contain one or more ligand structures to be docked. The file must be a Maestro file or an SD file, either compressed or uncompressed. The file name is displayed in the File text box.
- Selected entries—Use the selected entries in the Project Table as the source of ligands. These entries are written to a file named *jobname*`_lig.maegz`.

4.3 Receptor Options

The Receptor tab has options for defining the size and position of the receptor region for which grids are generated, and for defining H-bond and metal constraints.

Box center

- Centroid of the ligand—This option centers receptor grids at the centroid of the molecule you select as the ligand. To define the ligand, select Pick and pick a ligand atom in the Workspace.

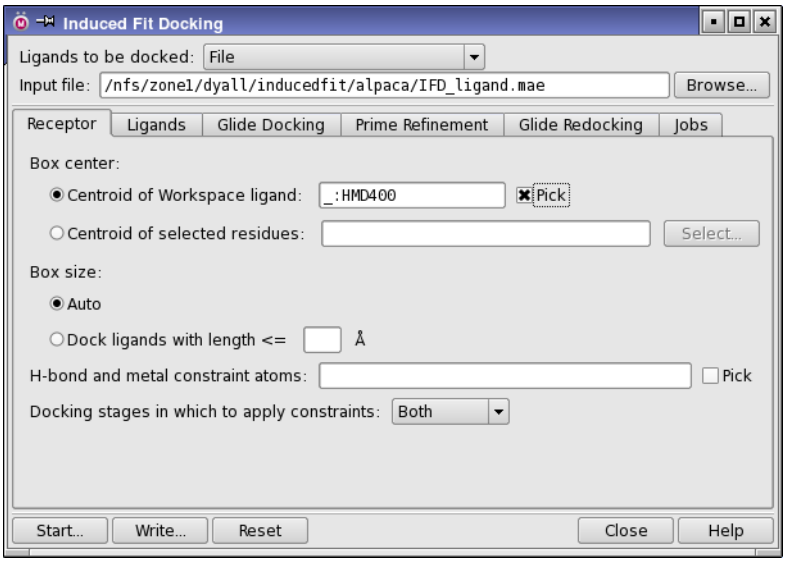

Figure 4.1. The Receptor tab of the Induced Fit Docking panel.

- Centroid of the residues—This option, which must be used if no ligand is present in the receptor structure in the Workspace, centers receptor grids at the centroid of a set of residues that you define. To choose the residues, click Select, and use the tools in the Atom Selection dialog box.

The inner box is drawn in green in the Workspace and the outer box is drawn in purple.

Box size

- Auto—If the Box center option is Centroid of the ligand, the enclosing box size is calculated automatically from the size of the ligand. If the Box center option is Centroid of the residues, the enclosing box size is set to 26 Å on a side.
- Specify—Select this option to specify the length of each edge of the enclosing box.

H-bond and metal constraint atoms

Select Pick to pick atoms in the receptor structure to be used for Glide H-bond or metal constraints. The atoms should be hydrogen-bond acceptors (e.g. O, N, S, with lone pairs available), hydrogen-bond donors (e.g. H in OH, NH, SH groups), or metal atoms. The atoms that are picked are listed in the text box. Symmetry-related atoms (such as the other O atom in a carboxylate group) are automatically included as constraints, so you only need to pick one. The constraints are applied using the default feature sets for hydrogen-bond donors and acceptors, and metal atoms. For more information on H-bond and metal constraints in Glide, see Section 4.4.2 and Section 5.5.1 of the *Glide User Manual*.

Docking stages in which to apply constraints

This option menu allows you to choose the Glide docking stages in which to apply the constraints, from Both, Initial, or Redocking.

4.4 Ligand Options

In the Ligands tab, you can select options and set values for the conformational sampling of the ligands:

Dock rigidly

Select this option to dock the ligands as they are in the input file, without doing any conformational (torsional) sampling. This option is useful, for example, if you want to pregenerate the ligand conformers, or if you want to dock a native ligand in its native conformation.

Sample ring conformations

Select this option to sample the conformations of rings. These conformations are not sampled in the conformation generation, which focuses on sampling of rotatable bonds, leaving the core fixed. This option is selected by default. Deselect this option if you want rings to remain in their input conformations throughout docking.

Energy window

Discard ring conformations whose energy is higher than that of the lowest conformation by the amount specified in this text box.

Amide bonds

This option menu provides a choice of how to treat amide bonds. The choices are:

- Penalize non-planar conformation—penalize amide bonds that are not cis or trans (default)
- Vary conformation—allow nonplanar amide bonds
- Retain original conformation—freeze amide bonds in their input conformation throughout docking
- Allow trans conformation only—enforce trans conformation within a small angle range (20°)

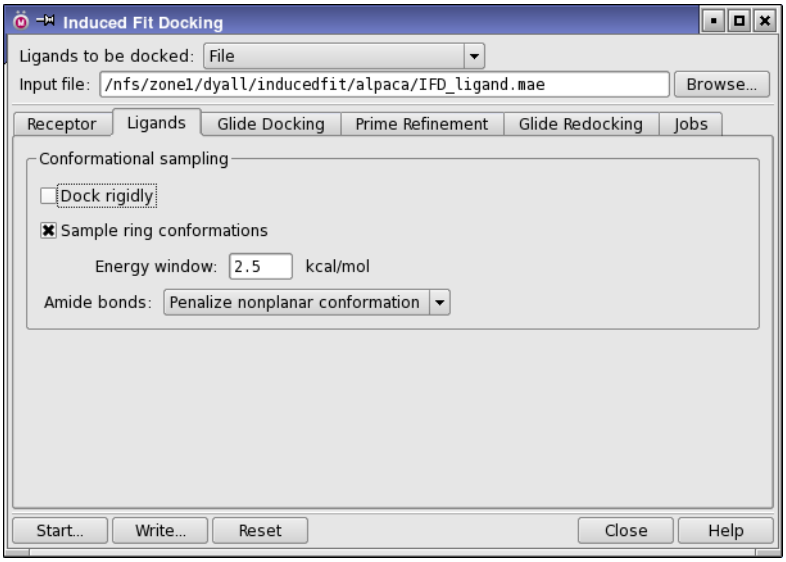

Figure 4.2. The Ligands tab of the Induced Fit Docking panel.

4.5 Initial Glide Docking Options

In the Glide Docking tab you set up the initial docking stage. This stage is intended to generate poses that can dock to the receptor when it adjusts to the presence of the ligand. To do this, side chains can be removed and interactions with hydrophobic groups can be adjusted to allow more room for the ligand.

4.5.1 Protein preparation

Select this option to run the constrained refinement part (impref) of the protein preparation procedure on the receptor. The constrained minimization ends when the RMSD is 0.18 Å or less. If you have already run the Protein Preparation Wizard for your receptor, you do not need to select this option. It is recommended that you run the Protein Preparation Wizard instead of relying on this option, and a warning is posted if the receptor does not appear to have been processed by the Protein Preparation Wizard.

4.5.2 Removing Side Chains

In the initial docking, you can remove the side chains of some residues to ensure that they do not prevent the ligands from docking in the preferred orientation. These residues are mutated to alanine for the initial docking stage. The original residue types are retained, and used to restore the original side chains for the Prime refinement stage of the Induced Fit Docking process.

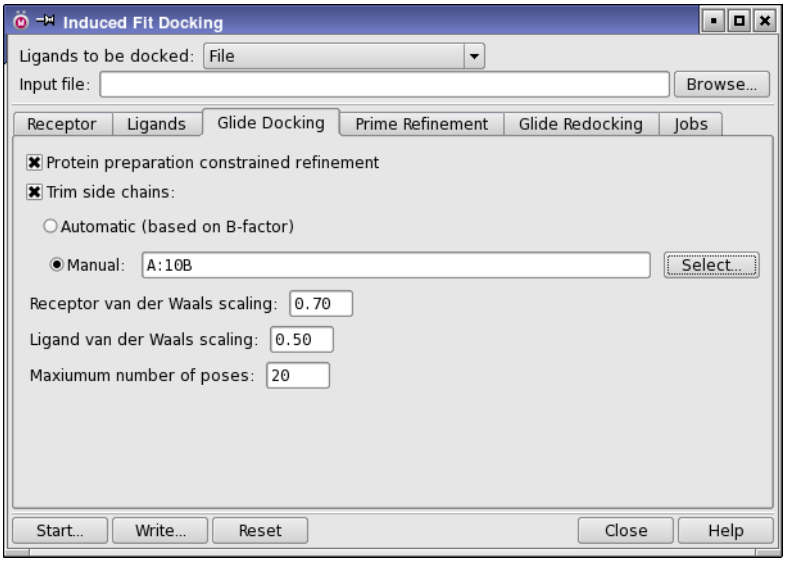

Figure 4.3. The Glide Docking tab of the Induced Fit Docking panel.

The removal of side chains is important if the side chains move significantly upon docking. It is also important if there is more than one binding mode, and if there were problems in the PDB structure.

To remove ("trim") side chains, select Trim side chains and select one of the two options for trimming the side chains:

- Automatic (based on B-factor)—Selects up to 3 residues that are within 5 Å of the ligand and have the highest B-factors; the B-factors must be above 40.
- Manual—Select the residues manually. Click Select to open the Atom Selection dialog box for selection of the desired residues.

As a general rule-of-thumb, you should select no more than three side chains. The residues containing these side chains are temporarily mutated to alanine for the initial docking step. To remove the side chains, select Trim side chains, and then select an option for trimming. For most cases, you should select Manual.

Some situations in which you would want to choose residues for mutation are described below.

a. If the protein is apo and there are existing holo proteins, superimpose the apo structure on one of the holo proteins and select residues in the active site that adopt significantly different positions.

b. If there are side chains with alternate positions (colored green on PDB import) or have missing density (colored red on PDB import), and either are within 5 Å of the ligand, they should be included in the side chain mutation.

c. If there are multiple structures in the unit cell (that have been independently solved in the X-ray structure determination, for example), superimpose these structures with the Protein Structure Alignment panel (Tools menu), and look at the active site residues. Any residues for which the side chains are in different locations should be considered for mutation.

d. Any side chain with a temperature factor (B) greater than about 40 should be considered for mutation, but not if the whole structure has high B values. If the whole structure has high temperature factors, then rank the residues in order of decreasing temperature factors and chose from the top of this list until a maximum of 3 residues is chosen. You can do this automatically by selecting Automatic based on B-factor for Trim side chains.

4.5.3 Softening the Potential

In addition to trimming side chains, you can soften the potentials by scaling the van der Waals radii of the receptor and ligand non-polar atoms. This simulates a small amount of flexibility in the ligand and the receptor.

Receptor van der Waals scaling

Specify the scaling factor for the receptor van der Waals terms in this text box. The default value of 0.50 was chosen to permit enough flexibility for the ligand to dock in the best poses. If side chains are trimmed, this value is changed to 0.70. The binding-site residue mutation is expected to reduce the need to soften the receptor potential by van der Waals radii scaling.

Ligand van der Waals scaling

Specify the scaling factor for the ligand van der Waals terms. The default value of 0.50 was chosen to permit enough flexibility for the ligand to dock in the best poses.

4.5.4 Limiting the Number of Poses

You can limit the number of poses per ligand to retain from the initial docking, by entering the desired limit in the Maximum number of poses text box. These poses are passed to Prime for the Prime refinement step.

4.6 Prime Refinement Options

In the Prime Refinement tab, you set options for the refinement of the protein to adjust to the poses of the ligand. The main task is to select the residues to refine. In general, you should choose residues for refinement that are within 5 Å of the active site, which is the default. To these you should add residues beyond this limit that have large motion—for example, if they are part of a helix or loop that goes close to the active site.

It is usually not necessary to exclude residues from refinement. If you are confident that the side chains are fixed— for example, if they are bound to a metal ion—you could leave these residues out of the refinement. In the case that there is a metal ion in the active site, the protein side chains that are ligating the metal should be excluded from the refinement.

Refine residues within *N* Å of ligand poses

Set a distance from the ligand for selecting residues to refine in this text box. Entire residues that have any atoms within the specified distance of any ligand atom are included in the refinement. The default, 5.0 Å, is recommended. With smaller values, jobs will run faster but the results may not be good if significant side-chain movement is necessary to accommodate the new ligand. With larger values, jobs will run slower but not necessarily yield better results. While 5.0 Å is the recommended value, values ranging from about 4 – 8 Å are reasonable to try.

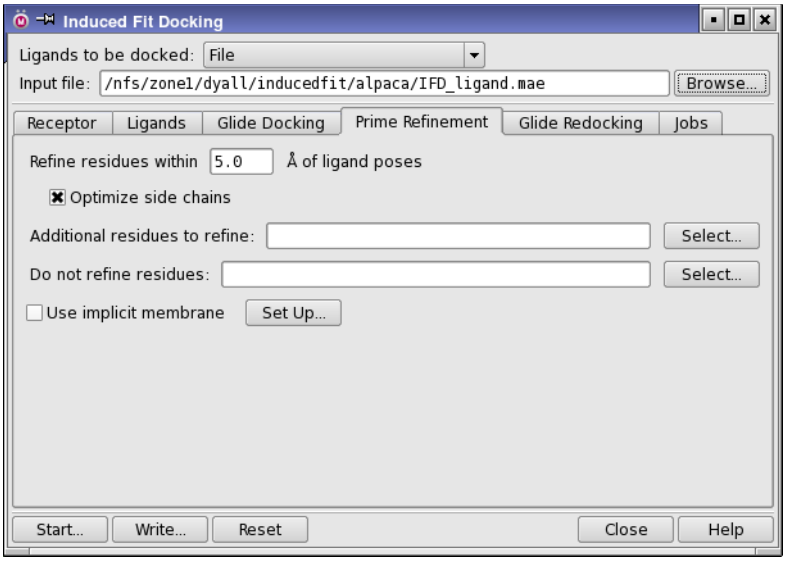

Figure 4.4. The Prime Refinement tab of the Induced Fit Docking panel.

Optimize side chains

By default, side chains are optimized. If this option is deselected, Prime skips the optimization of side chains and proceeds with the minimization of the selected residues and ligand. Skipping the side-chain optimization results in a faster calculation. Apart from speed, you might want to deselect this option if you are confident that the side chain conformations are essentially correct, or want to relax the structure without risking putting the side chains in new, and possibly incorrect, conformations.

Additional residues to refine

Click the Select button to choose residues that should undergo Prime refinement even if they are more than the specified distance from any ligand atom—for example, a loop that must adjust to the presence of the ligand. This button opens the Atom Selection dialog box, in which you can select the residues. The ASL expression for the residues is displayed in the text box.

Do not refine residues

Select residues to leave out of the refinement. Click Select to specify residues that need not undergo Prime refinement even if they are within *N* Å of a ligand pose. This button opens the Atom Selection dialog box. The ASL expression for the residues is displayed in the text box.

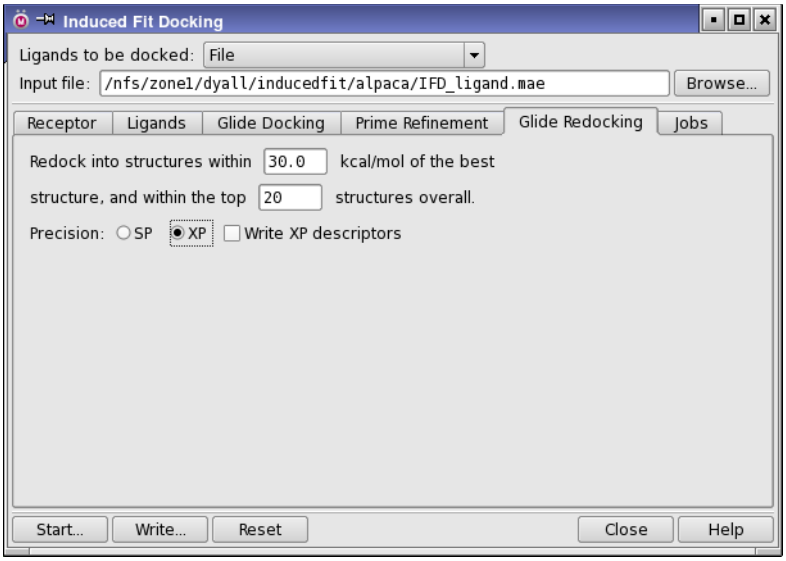

Figure 4.5. The Glide Redocking tab of the Induced Fit Docking panel.

Use implicit membrane

Select this option to use the Prime implicit membrane model for all Prime calculations. If the Workspace structure does not already have a membrane model, click Set Up to open the Prime Membrane Setup panel and set up the membrane. See Section 6.3 of the *Prime User Manual* for information on the implicit membrane model.

4.7 Glide Redocking Options

In the Glide Redocking tab, you select the receptor-ligand complex structures that are used for redocking, and set options for the redocking of the ligands into the refined receptors.

Redock into structures within *N* kcal/mol of the best structure, and within the top *M* structures overall

Set a threshold for eliminating high-energy structures from the Prime refinement step and a limit on the number of structures. Structures whose Prime energy is more than the specified amount above the lowest-energy structure and are not among the lowest *M* structures are eliminated. The default for the window *N* is 30.0 kcal/mol, and the maximum number of structures *M* defaults to 20, though of course the actual number of structures cannot exceed the number of ligand poses generated in the initial Glide docking step.

Precision

Select the docking precision for the redocking.

- SP—Standard-precision Glide docking. This is the default.
- XP—Extra-precision Glide docking. This option is recommended only when you are redocking a small number of low-energy structures. To ensure that this is the case, you can make the redocking energy window *N* smaller than the default, reduce the maximum number of structures *M* to be redocked, or both.

  If you select this docking precision, you can also choose to write out XP descriptors for use in the XP Visualizer (see Section 6.2 of the *Glide User Manual*), by selecting Write XP descriptors. You must have the required license to use this feature.

In the redocking, a small number of extra conformations are generated, to improve the final pose. This is done by default; if you want to turn it off, remove the `MULTI_LIG_CONF` keyword from the input file, and then run the job from the command line.

4.8 Running the Job

Induced fit docking jobs can be distributed across multiple processors. If you want to do so, you can specify the number of processors to use for Glide and the number to use for Prime in the Jobs tab. These numbers can be different.

- Number of Prime CPUs—The number of CPUs over which to distribute Prime subjobs. Each pose is run as a separate subjob. There is no benefit in specifying more CPUs than the number of poses.
- Number of Glide CPUs—The maximum number of CPUs on which to run Glide subjobs simultaneously. If multiple ligands are being docked, each can be run as a separate subjob. There is no benefit in specifying more CPUs than the number of poses.

If a subjob from a distributed job fails, the job continues with the remaining subjobs and stages. You can request that the job terminates immediately, which stops the job at the current stage, retaining the results of any completed subjob, by selecting Terminate job immediately if any subjob fails (this starts the job with the `-STRICT` option). When a job fails, you can restart it from the command line—see Section 5.1 on page 33.

To start the job, click Start. The Induced Fit Docking—Start dialog box opens, in which you can choose a host and specify the job name. If you requested multiple CPUs, you should choose a host that has the required number of CPUs available. Click Start in this dialog box to submit the job to the host.

4.9 Induced Fit Docking Results

When an Induced Fit Docking job finishes, it creates a compressed Maestro file named *jobname*`-out.maegz` in the launch directory. This file contains the output poses with their IFDScore. This score is the sum of the GlideScore from the redocking step and 5% of the Prime energy from the refinement calculation.

To view the results, you can import the structures from the Maestro file. By default, only the ligand and the protein residues that were refined are displayed in the Workspace.

Running Induced Fit Docking from the Command Line

If you simply want to run the standard Induced Fit Docking protocol, the Induced Fit Docking panel in Maestro provides the easiest means of running the calculations, with a range of options. If you want to set up options that are not available from Maestro, customize the steps in the protocol, add steps or rearrange steps, you can do so by editing the input file and submitting the job from the command line. The input file is designed so that you can tailor the protocol to your specific task, within the scope of the tools available. On Windows, you can use a Schrodinger Command Prompt window, which you open from the Start menu.

5.1 The ifd Command

You can use the `ifd` command to launch an Induced Fit Docking job from the command line:

```
ifd [options] jobname.inp
```

The job settings are specified on the command line. You can include them in the input file, but the options specified on the command line take precedence. You must run this script in the directory in which *jobname*`.inp` resides. The command supports the standard Job Control options—see Section 2.3 of the *Job Control Guide*. The remaining options are given in Table 5.1. The format of the input file is described in the next section.

Table 5.1. Options for the ifd command

Option	Description
`-HOST` *host*	Run the master job (driver) on the specified host. If the job launch directory is not available on that host, you must also use `-NOLOCAL` or the job will fail.
`-SUBHOST` *host*	Run the subjobs on the specified host.
`-NGLIDECPU` *N*	Maximum number of CPUs to use for Glide subjobs.
`-NPRIMECPU` *M*	Maximum number of CPUs to use for Prime subjobs.
`-STRICT`	Terminate the job if any subjob fails.
`-RETRIES` *N*	Number of times to try rerunning a subjob if it fails. Default: 0.
`-OVERWRITE`	Overwrite the input file
`-RESTART`	Restart the job with the specified input file. The file *jobname*`.restart` must exist: this file contains the current state of a job.

5.2 The ifd Input File

The input file for the `ifd` command defines the stages of the protocol and the order in which they are executed in addition to defining the input settings for each stage. This means that you can define your own protocol with the available stages. You can specify any stage multiple times, and the stages are run in the order in which you specify them.

The input file is structured as follows:

```
<Global Settings section>
STAGE stage-name1
  <Stage settings>
STAGE stage-name2
  <Stage settings>
...
```

Each group of settings consists of a line containing a keyword and its value. The settings for a stage apply only to that stage. If you repeat a stage, the settings revert to their defaults unless you explicitly set them. The list of stages is given in Table 5.2.

Table 5.2. Description of stages in the induced fit docking protocol.

Stage	Description
COMPILE_RESIDUE_LIST	Compile a list of residues for refinement.
GLIDE_DOCKING	Dock ligands using Glide. This stage includes both grid generation and ligand docking.
PPREP	Prepare the protein using the refinement part of the Glide protein preparation facility.
PRIME_ECALC	Perform a Prime energy calculation.
PRIME_HELIX	Perform a Prime rigid-body helix refinement for a specified helix.
PRIME_LOOP	Perform a Prime loop prediction for the specified loop.
PRIME_MINIMIZATION	Perform a Prime minimization on the compiled list of residues.
PRIME_REFINEMENT	Perform a Prime refinement on the compiled list of residues.
PRIME_SIDECHAIN	Perform a Prime side-chain prediction on the compiled list of residues.
SCORING	Calculate scores for the poses.
SORT_AND_FILTER	Sort and filter poses.
TRIM_SIDECHAINS	Temporarily mutate the specified residues to alanine, to remove side chains for a following docking stage.

Some keywords take a residue, a list of residues, or a list of atoms as their value. Residues are specified in the format *chain*:*residue*, where *chain* is the single-letter chain name and *residue* is the residue number and insertion code, for example `A:151C`. Atoms are specified in the format *chain*:*residue*:*atom*, where *atom* is the PDB atom name, for example `A:151:_N__`, and underscores are used instead of blanks. In the descriptions below, *residue-spec* and *atom-spec* are used to denote these specifications.

5.2.1 Global Settings

This section contains settings that affect the whole job, and consists of host settings and receptor input file definitions. The keywords for this section are given in Table 5.3. Note that there is no benefit in specifying more CPUs than the number of poses.

Table 5.3. Keywords for the global settings section

Keyword	Description
`INPUT_FILE`	Specify the file name for the receptor. The file must be a Maestro file (`.mae`, `.maegz`, or `.mae.gz`). Multiple receptors can be specified, either in a single file or by including multiple instances of this keyword. The IFD protocol is applied to each receptor independently and the results collated.
`NUM_GLIDE_CPUS`	The maximum number of CPUs over which to distribute Glide subjobs. When docking multiple ligands, each can be run as a separate subjob. Default: 1.
`NUM_PRIME_CPUS`	Specify the number of CPUs over which to distribute Prime subjobs. Each pose is run as a separate subjob. Default: 1.
`SUBJOB_HOST`	Specify the host on which to run Glide and Prime subjobs. Default is `localhost`.

5.2.2 The PPREP Stage

The `PPREP` stage has one setting, `RMSD` *value*, which specifies the convergence threshold for the constrained minimization of the Glide protein preparation. The minimization ends when the RMSD is less than or equal to *value*.

5.2.3 The TRIM_SIDECHAINS Stage

This stage specifies the residues whose side chains should be temporarily removed by mutating the residues to alanine. The next Glide docking step uses the mutated residues. The original residue types are retained, and used to restore the original side chains later in the Induced Fit Docking process. For more information on selecting side chains, see Section 4.5 on page 25.

Two methods are available for automatic selection of side chains for mutation: by B-factor and by flexibility. The first selects the side chains with the highest B-factors. This method appears

to be insensitive to surface-exposed residues, and is the preferred method. The second selects side chains based on whether it has rotamer states that do not clash with the rest of the receptor, and can select surface-exposed residues for mutation.

The keywords for this stage are given in Table 5.4.

Table 5.4. Keywords for the TRIM_SIDECHAINS stage.

Keyword	Description
RESIDUES {*spec*|AUTO}	Residues for temporary mutation to alanine. Either a list of residues must be given, or AUTO.
METHOD *method*	Method for automatic specification of residues. Allowed values: BFACTOR—Mutate residues near the ligand with the highest B factors, subject to a cutoff on the B-factor and the number of residues. FLEXIBILITY—Mutate residues for which other rotamers exist that don't have steric clashes with the rest of the protein. Default: BFACTOR.
BFACTOR_CUTOFF *cutoff*	B-factor cutoff above which residues are selected for mutation. Used only with METHOD BFACTOR. Default: 40.0.
MAX_RESIDUES *N*	The maximum number of residues to mutate. If more than *N* residues have B-factors greater than the value of BFACTOR_CUTOFF, the *N* residues with the highest B-factors are selected. Used only with METHOD BFACTOR. Default: 3.
MAX_FLEXIBILITY *dist*	If any heavy atom in the side chain is capable of moving more than this distance, the side chain is considered flexible and is mutated. Used only with METHOD FLEXIBILITY. Default: 5.0 Å.
RESOLUTION *angle*	Granularity of the rotamer search for steric clashes, in degrees. Used only with METHOD FLEXIBILITY. Default: 30.0.

5.2.4 The GLIDE_DOCKING Stage

This stage performs the Glide grid generation and ligand docking. Settings for Glide are made in this section. If multiple receptors are specified in the input files, the settings are applied to each receptor. This means that the binding site is defined by the same residue specifications for each receptor. For example, if a ligand is used, it must have the same residue name, residue number, and chain ID in each complex.

Keywords that are not listed in Table 5.5 are passed directly to Glide, which looks them up in its internal list (MMIM keywords). This allows the use of any available Glide options via these keywords. Some of these, which are set in the GUI, are listed below:

- MMIM_DOCKING_METHOD rigid—dock ligands rigidly.

- MMIM_GLIDE_CONFGEN_RINGCONF {TRUE|FALSE}—sampling of ring conformations.
- MMIM_LIG_RINGCONFCUT *energy*—energy window for eliminating ring conformations.

Table 5.5. Keywords for the GLIDE_DOCKING stage.

Keyword	Description
AMIDE_MODE	Amide bond sampling mode. penal—penalize nonplanar conformation free—vary conformation fixed—retain original conformation trans—allow trans conformation only
BINDING_SITE	Grid center. The center can be specified in one of the following ways: coords *x*,*y*,*z* ligand *residue-spec* residues *residue-spec*,*residue-spec*,... coords specifies the grid center directly; ligand specifies the centroid of the ligand, and residues specifies the centroid of the listed residues.
CONSTRAINT_ATOMS	Comma-separated list of Glide H-bond constraint atoms: *atom-spec*,*atom-spec*,... . Any constraints that are set are applied in docking: there are no optional constraints.
CV_CUTOFF	Threshold for rejecting poses based on Coulomb-van der Waals energy. Poses are rejected if the energy is greater than the threshold. Default: 0.00
HBOND_CUTOFF	Threshold for rejecting poses based on hydrogen bonding energy. Poses are rejected if the energy is greater than the threshold. Default: 0.00
INCLUDE_NATIVE	Run a standard Glide initial docking job and add the top pose to the set of poses kept. Can take values Yes or No. Default: No.
INNER_BOX	Dimension of ligand bounding box. Default: 10.0
LIGAND_CCUT	Partial charge threshold for scaling ligand van der Waals radii. Default: 0.15
LIGAND_FILE	Name of ligand file. Must be in Maestro format (.mae, .mae.gz, .maegz). Required.
LIGAND_SCALE	Scaling factor for ligand van der Waals radii. Default: 0.80
LIGANDS_TO_DOCK	List of ligands to dock from the ligand file. Can take the values all, self (ligand that was last docked with this receptor), or a comma-separated list of integers with no white space. Default: all
MAX_LIG_ATOMS	Maximum number of ligand atoms. Ligands that do not meet this criterion are discarded. Default: 200
MAX_POSESPERLIG	Maximum number of poses per ligand. Default: 1
MAX_ROT_BONDS	Maximum number of rotatable bonds. Ligands that do not meet this criterion are discarded. Default: 35

Table 5.5. Keywords for the GLIDE_DOCKING stage. (Continued)

Keyword	Description
MAX_TOTALPOSES	Maximum total number of poses. Default: 100000
MINIMUM_POSES	Redock without H-bond filtering if less than this number of poses is found. Default: 0
MULTI_LIG_CONF	Generate 4 additional conformations for input to Glide docking; these extra conformations improve the ligand pose. Can take values Yes or No. Default: No.
OUTER_BOX	Dimension of grid enclosing box. Can take the value auto or a number. The value auto computes the box size from the size of the ligand, if the grid is centered on the ligand, or sets it to 26 Å if the grid is centered on the centroid of a set of residues. Default: auto
PRECISION	Glide docking precision. Can take the values SP or XP. Default: SP
RECEPTOR_CCUT	Partial charge threshold for scaling receptor van der Waals radii. Default: 0.25
RECEPTOR_SCALE	Scaling factor for receptor van der Waals radii. Default: 1.00

5.2.5 The COMPILE_RESIDUE_LIST Stage

The list of residues for Prime refinement is compiled in this section. The initial list includes all residues within a prescribed distance of the ligand (whose identity can be specified in terms of a set of residues). To this list, specified residues that lie outside this cutoff can be added, and specified residues inside the cutoff can be omitted.

Table 5.6. Keywords for the COMPILE_RESIDUE_LIST stage.

Keyword	Description
CENTER	List of residues from which to measure the cutoff distance. Default: Z:999, which is the default for the ligand.
DISTANCE_CUTOFF	Cutoff distance (in angstroms) from the ligand pose, within which residues that have any atoms are included in the refinement list. Default: 5.0
RESIDUES_TO_ADD	Comma-separated list of residues to add to the refinement list. These should be residues that lie outside the distance cutoff
RESIDUES_TO_OMIT	Comma-separated list of residues to omit from the refinement list.

5.2.6 The PRIME_REFINEMENT and PRIME_MINIMIZATION Stages

These stages perform a Prime refinement or a Prime minimization. In Prime refinement, the side chains of the residue list compiled previously are optimized, then the residues are minimized along with the ligand. There are two stage settings: `NUMBER_OF_PASSES` *n*, and `USE_MEMBRANE` {`true`|`false`} (optional). However, you can also add keywords for the `prime` program, as described in Section 10.2 of the *Prime User Manual*.

By default, only one pass through Prime refinement is performed, consisting of three steps:

1. Optimize side chains.
2. Minimize residues.
3. Minimize residues and ligand.

If multiple passes are requested, the first two steps are executed the number of times specified specified by *n*. Multiple passes could be useful if, for example, the active site is extremely packed. The minimization would open up the structure somewhat and allow a better side-chain prediction.

In Prime minimization, only the last two steps are performed, and they are performed only once. There is only one optional setting for `PRIME_MINIMIZATION`: `USE_MEMBRANE` {`true`|`false`}. However, you can also add keywords for the `prime` program, as described in Section 10.2 of the *Prime User Manual*. A Prime minimization is faster than a Prime refinement. You might want to use a Prime minimization if you are confident that the side chain conformations are essentially correct, or want to relax the structure without risking putting the side chains in new, and possibly bad, conformations.

5.2.7 The PRIME_ENERGY and PRIME_SIDECHAIN Stages

These stages perform a Prime energy calculation or a Prime side-chain prediction. The energy calculation is done on the entire complex. The side-chain prediction is run on the previously compiled residue list. These stages do not have any specific settings, but you can add keywords for the `prime` program, as described in Section 10.2 of the *Prime User Manual*. The `PRIME_SIDECHAIN` stage allows you to run a side-chain prediction independently of the `PRIME_REFINEMENT` or `PRIME_MINIMIZATION` stages.

5.2.8 The PRIME_LOOP Stage

This stage performs a Prime loop prediction. If the receptor has a particularly flexible loop that might preclude ligand binding even with a softened potential, you could consider doing a loop prediction before the initial Glide docking. If there are more subtle loop movements associated with ligand binding that cannot be reached by minimization alone, you could consider adding a loop prediction after the Prime refinement.

Loop prediction is performed with the full input structure, including the ligand, if present. The input structure is the structure from the previous stage of the protocol. If the loop prediction is done as the first stage, the input structure is defined by `INPUT_FILE`. If you want to do an initial loop prediction on an apo protein, the structure defined by `INPUT_FILE` should be an apo structure, not a complex. If you perform the loop prediction after docking, the ligand from the docking calculation is present and cannot be removed.

In addition to predicting the loop itself, you can refine the side chains of other residues along with the loop. These residues can be selected beforehand with a `COMPILE_RESIDUE_LIST` stage, and added by setting `INCLUDE_RESIDUE_LIST` to `TRUE`, or added as a shell of residues within a distance specified by `DISTANCE_CUTOFF`. If you specify extra residues by both mechanisms, all members of both sets are included.

Table 5.7. Keywords for the PRIME_LOOP stage

Keyword	Description
`START_RESIDUE`	First residue in the loop.
`END_RESIDUE`	Last residue in the loop.
`DISTANCE_CUTOFF`	Threshold for inclusion of residues for side-chain refinement. Any residues with atoms within this distance are included. Default: 5.0.
`MAX_ENERGY_GAP`	Energy threshold for predicted loop structures (in kcal/mol). Structures are discarded if their energy is more than this amount above the lowest-energy structure. Default: 10000.0.
`MAX_STRUCTURES`	Maximum number of structures to retain. Default: 1000
`INCLUDE_RESIDUE_LIST`	Include the residues from `COMPILE_RESIDUE_LIST` for side-chain refinement. Can take values `TRUE` or `FALSE`. Default: `FALSE`.
`USE_MEMBRANE`	Use the implicit membrane model. Can take values `TRUE` or `FALSE`. Default: `FALSE`.

5.2.9 The PRIME_HELIX Stage

This stage performs a rigid-body refinement of a region containing a helix. The keywords specific to this stage are defined in Table 5.8. Other keywords are passed on to `refinestruct`, the Prime executable.

Table 5.8. Keywords for the PRIME_HELIX stage.

Keyword	Description
START_RESIDUE	First residue of mobile region that contains the helix
START_HELIX_RESIDUE	First residue of helix
END_HELIX_RESIDUE	Last residue of helix
END_RESIDUE	Last residue of mobile region that contains the helix
DISTANCE_CUTOFF	Distance cutoff for prediction of side chains close to helix.
MAX_ENERGY_GAP	Maximum energy gap for saved structures.
MAX_STRUCTURES	Maximum number of structures to store.
USE_MEMBRANE	Use implicit membrane model (must be set up).

5.2.10 The SORT_AND_FILTER Stage

The sorting and filtering stage first groups all structures by the ligand contained within each structure. The poses for a particular ligand are then sorted by the property specified by POSE_FILTER. POSE_KEEP can then be used to keep the best poses, defined as those that have the smallest (most negative) value of the property, and discard the rest.After this filtering step, the groups of poses for each ligand are sorted by the property specified by LIGAND_FILTER for the top pose in each group. LIGAND_KEEP can then be used to discard entire ligand groups, in the same way as with POSE_KEEP.

5.2.11 SCORING Settings

In this stage you can define the scoring function in terms of Maestro properties that are available in the output file from each stage. The default scoring function when you run Induced Fit Docking from Maestro is a two-term function that adds 0.05 of the Prime energy to the GlideScore. You can provide a name for the property, which is written to the output Maestro file and can be displayed in the Project Table.

To define the scoring function, include TERM settings for each property that you want to include in the scoring function. The property must come from the Maestro output file of one of the previous stages. For the purpose of generating a scoring function, the stages are indexed by

Table 5.9. Keywords for the SORT_AND_FILTER stage.

Keyword	Description
POSE_FILTER	Name of Maestro property for filtering poses, for example, r_psp_Prime_Energy
POSE_KEEP	Threshold on property for filtering poses. The syntax is as follows: *n*% Keep the *n*% of poses with the lowest property values *n*# Keep the *n* poses with the lowest property values *n* Keep poses with property values within *n* of the lowest value.
LIGAND_FILTER	Name of Maestro property for filtering ligands, for example, r_psp_Prime_Energy
LIGAND_KEEP	Threshold on property for filtering ligands. The syntax is the same as for POSE_KEEP.

counting stages that produce output files backwards from the current stage, starting from zero. As an example, the indexes of the stages are shown to the left for the following sequence of stages.

Table 5.10. Keywords for the scoring settings section

Keyword	Description
SCORE_NAME	Name of property to add to Maestro files. Must be in the format r_psp_*name*, where *name* is the property name displayed in Maestro.
TERM	Add a term to the scoring function. You can include multiple TERM keywords to define the scoring function. Format: *coeff*, *property*, *stage,* where *coeff* is the coefficient, *property* is the property from the Maestro output file, and *stage* is the index of the property-generating stage, counting backwards in the input file with 0 for the previous stage.
REPORT_FILE	CSV file containing ligand number, score, score terms, and file name of the structure. Default: scores.csv.

```
4    STAGE PPREP
3    STAGE PRIME_LOOP
2    STAGE GLIDE_DOCKING
     STAGE COMPILE_RESIDUE_LIST
1    STAGE PRIME_REFINEMENT
     STAGE SORT_AND_FILTER
     STAGE SORT_AND_FILTER
0    STAGE GLIDE_DOCKING
     STAGE SCORING
```

The stages that do not generate a Maestro output file are COMPILE_RESIDUE_LIST, SORT_AND_FILTER, and SCORING. The SCORING stage adds the score to the Maestro file from the last stage.

The output from the scoring stage is a comma-separated value (CSV) file containing (in order) the ligand number, (which is always 1), the score, the list of terms in the score, and the file-name for the structure. The rows are sorted by the score. The CSV file and the files for the structures are in the *jobname*_workdir subdirectory. The default name is scores.csv.

5.2.12 Sample Input File

A sample input file is given below. This file was generated for the tutorial example by clicking Write in the Induced Fit Docking panel. The comments are generated when the file is written.

```
# Global Variables
#  These variables affect the entire job, and must all appear
#  before the first STAGE declaration.  Multiple INPUT_FILE
#  entries are supported, as are files containing multiple
#  receptor structures.
#
#  If beginning with an existing Pose Viewer file, simply specify
#  it as the INPUT_FILE (making sure the name ends in "_pv.mae")
#  and ensure that the first GLIDE_DOCKING stage is commented out.
#  The ligand used in producing the Pose Viewer file must also be
#  provided to the second GLIDE_DOCKING stage, using the LIGAND_FILE
#  keyword.
INPUT_FILE InducedFit1_rec.mae

# Protein Preparation
#  Run a simple constrained minimization of the receptor
#  structure(s).
STAGE PPREP
  RMSD 0.18

# Prime Loop Prediction
#  Perform a loop prediction on the specified loop, including
#  side chains within the given distance.  Only return
#  structures within the specified energy range from the
#  lowest energy prediction, up to the maximum number of
#  conformations given.
#
#  Note: This stage is disabled by default.  Uncomment the
#   lines below and edit the fields appropriately to enable it.
#STAGE PRIME_LOOP
#  START_RESIDUE A:11
#  END_RESIDUE A:16
#  RES_SPHERE 7.5
#  MAX_ENERGY_GAP 30.0
```

```
#  MAX_STRUCTURES 5
#  USE_MEMBRANE no
#  FIX_ATOM_NAMES yes

# In order to temporarily remove the side chains of residues
# (i.e., mutate to Ala) that are blocking the binding site,
# uncomment the following STAGE line, and then specify the
# sidechains to be removed using either one of the two Methods
# described below.
#
STAGE TRIM_SIDECHAINS
  RESIDUES A:10B

# Glide Docking
#  Perform the initial Glide docking, producing a
#  ligand-receptor complex for each pose requested/found.
#  If multiple receptor structures are used, the requested
#  number of poses will be generated for each structure.
STAGE GLIDE_DOCKING
  RECEPTOR_CCUT 0.25
  LIGAND_FILE InducedFit1.mae
  LIGANDS_TO_DOCK all
  MULTI_LIG_CONF no
  LIGAND_CCUT 0.15
  CV_CUTOFF 100.0
  HBOND_CUTOFF -0.05
  INNER_BOX 10.0
  MINIMUM_POSES 1
  MMIM_LIG_RINGCONFCUT 2.5
  AMIDE_MODE penal
  BINDING_SITE ligand _:HMD400
  OUTER_BOX auto
  RECEPTOR_SCALE 0.70
  LIGAND_SCALE 0.50
  MAX_POSESPERLIG 2
  PRECISION SP

# Determine Residue to Refine
#  Compile a list of all residues within the specified
#  distance of any pose of the ligand.
STAGE COMPILE_RESIDUE_LIST
  DISTANCE_CUTOFF 3.4

# Prime Refinement
#  Optimize the side chains of the residue list compiled
#  previously, then minimize them along with the ligand.
STAGE PRIME_REFINEMENT
  NUMBER_OF_PASSES 1
  USE_MEMBRANE no
  FIX_ATOM_NAMES yes
```

```
# Sort and Filter
#  Only retain poses with Prime Energies within the
#  specified range from the lowest energy pose.
STAGE SORT_AND_FILTER
  POSE_FILTER r_psp_Prime_Energy
  POSE_KEEP 30.0

# Sort and Filter
#  Only retain the top number of poses specified.
STAGE SORT_AND_FILTER
  POSE_FILTER r_psp_Prime_Energy
  POSE_KEEP 20#

# Glide Docking
#  Redock the ligand back into the newly optimized receptor,
#  using default Glide settings.
STAGE GLIDE_DOCKING
  BINDING_SITE ligand Z:999
  RECEPTOR_SCALE 1.00
  RECEPTOR_CCUT 0.25
  LIGAND_FILE InducedFit1.mae
  LIGANDS_TO_DOCK self
  MULTI_LIG_CONF yes
  LIGAND_SCALE 0.80
  LIGAND_CCUT 0.15
  CV_CUTOFF 0.0
  HBOND_CUTOFF 0.0
  INNER_BOX 10.0
  MAX_POSESPERLIG 1
  MMIM_LIG_RINGCONFCUT 2.5
  AMIDE_MODE penal
  OUTER_BOX auto
  PRECISION SP

# Scoring
#  Compile the IFD Score, consisting of the GlideScore for
#  the Glide Redocking plus 5% of the Prime Energy from the
#  Prime Refinement.
STAGE SCORING
  SCORE_NAME  r_psp_IFDScore
  TERM 1.0,r_i_glide_gscore,0
  TERM 0.05,r_psp_Prime_Energy,1
  REPORT_FILE report.csv
```

5.3 Files

An Induced Fit docking run requires an input file, as described in the previous section, a file of ligands in Maestro format, and one or more receptor files, also in Maestro format. The structure files can be compressed (`.maegz`, `.mae.gz`) or uncompressed (`.mae`).

As the induced fit docking job proceeds, input files and results files are written to the *jobname*`_workdir` subdirectory of the output directory by default. If you ran the job with the `-NOLOCAL` option, this subdirectory is created in a scratch directory instead, and its files are not available after the job finishes. If you are only interested in the final results, you need not be concerned with the intermediate files. If you want to examine the results for a given stage, however, you will need to know how these files are named.

5.3.1 Intermediate Files

The names of the intermediate files start with *jobname*, and each stage in the protocol appends a descriptive suffix to the names of the files that pass through it.

The input ligand file is named *jobname*`_lig.maegz` and each input receptor is written to a file named *jobname*`_rec-`*N*`.maegz`, where *N* is an index starting from 1.

The suffixes that are added and passed on are listed in Table 5.11. The usual suffix for an output or log file is added to the stem inherited from the previous stage. For example, a docking stage adds `_grid.log`, `_dock.log`, `_pv.maegz`, `.rept`, `.log` for the various log and output files. The final file stem for a run that included all optional stages—protein preparation, loop refinement, and side-chain mutation—would be *jobname*`_rec-`*N*`_ref-out-`*M*`-trim_pv-`*L*`-out-`*K*.

Table 5.11. Suffixes appended to the stem of the file name by each stage.

Suffix	Stage or process
`_ref`	Impref minimization
`-trim`	Side-chain mutation
`_pv`	Docking, used for pose-viewer output
`-`*M*	Index added for each receptor, pose, or loop.
`-out`	Prime loop prediction or refinement

5.3.2 Final Output Files

The final Maestro output file is copied to the launch directory with the name *jobname*-out.maegz. The score is present as a Maestro property in these files. The scoring stage generates a comma-separated-values file in the *jobname*_workdir directory, named report.csv by default.

The files produced by an Induced Fit Docking run in the output directory are listed in Table 5.12.

Table 5.12. FIles produced by Induced Fit Docking run.

File	Description
jobname.log	Log file, containing details of settings for each stage and execution of stages.
jobname-out.maegz	Maestro file containing the results.
*jobname*_restart	Restart file. Stores the current state of the job. Rerunning the job detects this file and prompts for a restart or a new run.

A sample log file, for the tutorial Induced Fit Docking run, is shown below.

```
==================================
  Induced Fit Docking Calculation
==================================

Started at: Tue Feb 14 14:39:20 2012
Job ID: myhost-0-4f3ae28d

Output will be written to: /nfs/dyall/inducedfit/alpaca/InducedFit1-out.maegz

Here are the parameters that will be used:

    General:
        JobHost: localhost
        #Prime CPUs: 2
        #Glide CPUs: 2
        Working Dir: /nfs/dyall/inducedfit/alpaca/InducedFit1_workdir

    Stage: Protein Preparation
        Max RMSD:   0.18

    Stage: Trimming Sidechains
        Residues:            A:10B

    Stage: Glide Docking
        Receptor Scaling:        0.70
```

```
    Receptor Scaling Cutoff: 0.25
    Ligand Source:           InducedFit1.maegz
    Ligand To Dock:          all
    Precision:               SP
    Ligand Scaling:          0.50
    Ligand Scaling Cutoff:   0.15
    CV Cutoff:               100.0
    H-Bond Cutoff:           -0.05
    Poses Per Ligand:        2
    Minimum Poses:           1
    Include Standard Docking no
    Amide Rotations:         penal

Stage: Determine Residues for Refinement
    Distance Cutoff: 3.4 A
    Additional Residues:
    Omit Residues:

Stage: Prime Active Site Optimization
    Number of Passes:  1
    Implicit Membrane: no

Stage: Sorting and Filtering
    Pose Filter:     r_psp_Prime_Energy
        Keep:        30.0
    Ligand Filter:   <none>
        Keep:        <none>

Stage: Sorting and Filtering
    Pose Filter:     r_psp_Prime_Energy
        Keep:        20#
    Ligand Filter:   <none>
        Keep:        <none>

Stage: Glide Docking
    Receptor Scaling:        1.00
    Receptor Scaling Cutoff: 0.25
    Ligand Source:           InducedFit1.maegz
    Ligand To Dock:          self
    Precision:               SP
    Ligand Scaling:          0.80
    Ligand Scaling Cutoff:   0.15
    CV Cutoff:               0.0
    H-Bond Cutoff:           0.0
    Poses Per Ligand:        1
    Minimum Poses:           0
    Include Standard Docking no
    Amide Rotations:         penal
```

```
    Stage: Scoring
        Score =   + 1.0 r_i_glide_gscore(0)  + 0.05 r_psp_Prime_Energy(1)
        Report File: report.csv

Number of initial structures: 1

================================================
Stage: Protein Preparation
================================================

Running subjobs on hosts:
  localhost (Max: 1)
Number of jobs:       1
Max retries per job:  0
Max allowed failures: 1000000
Run with -LOCAL:      False
Verbosity:            normal

Starting JobDJ...

Keep one job on localhost: False
JobDJ columns:
  C: Number of completed subjobs
  A: Number of active subjobs (e.g., submitted, running)
  W: Number of waiting/pending subjobs

C A W | Activity  JobId  JobName  JobHost
- - - | --------- -----------------------
0 1 0 | launched  myhost-0-4f3ae2a5 InducedFit1_rec-1 localhost [myhost]
1 0 0 | finished  myhost-0-4f3ae2a5 InducedFit1_rec-1 localhost [myhost]

All jobs are done.

    Structures to be carried forward: 1
Stage completed.  Elapsed time: 66.8 seconds
================================================

================================================
Stage: Trimming Sidechains
================================================
    Structures to be carried forward: 1
Stage completed.  Elapsed time: 0.2 seconds
================================================

================================================
Stage: Glide Docking
================================================

Running subjobs on hosts:
```

```
  localhost (Max: 2)
Number of jobs:       2
Max retries per job:  0
Max allowed failures: 1000000
Run with -LOCAL:      False
Verbosity:            normal

Starting JobDJ...

Keep one job on localhost: False
JobDJ columns:
  C: Number of completed subjobs
  A: Number of active subjobs (e.g., submitted, running)
  W: Number of waiting/pending subjobs

C A W | Activity  JobId  JobName  JobHost
- - - | --------- -----------------------
0 1 1 | launched  myhost-0-4f3ae2df InducedFit1_rec-1_ref-trim localhost [myhost]
0 2 0 | launched  myhost-0-4f3ae2e4 InducedFit1_rec-1_ref-trim_noHbond localhost
[myhost]
1 1 0 | finished  myhost-0-4f3ae2df InducedFit1_rec-1_ref-trim localhost [myhost]
2 0 0 | finished  myhost-0-4f3ae2e4 InducedFit1_rec-1_ref-trim_noHbond localhost
[myhost]

All jobs are done.

    Structures to be carried forward: 2
Stage completed.  Elapsed time: 151.1 seconds
================================================

================================================
Stage: Determine Residues for Refinement
================================================
    Calculating residue distances...
    Structures to be carried forward: 2
Stage completed.  Elapsed time: 0.3 seconds
================================================

================================================
Stage: Prime Active Site Optimization
================================================

Running subjobs on hosts:
  localhost (Max: 2)
Number of jobs:       2
Max retries per job:  0
Max allowed failures: 1000000
Run with -LOCAL:      False
Verbosity:            normal
```

```
Starting JobDJ...

Keep one job on localhost: False
JobDJ columns:
  C: Number of completed subjobs
  A: Number of active subjobs (e.g., submitted, running)
  W: Number of waiting/pending subjobs

C A W | Activity  JobId  JobName  JobHost
- - - | --------- -----------------------
0 1 1 | launched  myhost-0-4f3ae377 InducedFit1_rec-1_ref-trim_pv-1 localhost
[myhost]
0 2 0 | launched  myhost-0-4f3ae37c InducedFit1_rec-1_ref-trim_pv-2 localhost
[myhost]
1 1 0 | finished  myhost-0-4f3ae377 InducedFit1_rec-1_ref-trim_pv-1 localhost
[myhost]
2 0 0 | finished  myhost-0-4f3ae37c InducedFit1_rec-1_ref-trim_pv-2 localhost
[myhost]

All jobs are done.

    Structures to be carried forward: 2
Stage completed.  Elapsed time: 166.6 seconds
================================================

================================================
Stage: Sorting and Filtering
================================================
    Structures to be carried forward: 2
Stage completed.  Elapsed time: 0.1 seconds
================================================

================================================
Stage: Sorting and Filtering
================================================
    Structures to be carried forward: 2
Stage completed.  Elapsed time: 0.1 seconds
================================================

================================================
Stage: Glide Docking
================================================

Running subjobs on hosts:
  localhost (Max: 2)
Number of jobs:       2
Max retries per job:  0
Max allowed failures: 1000000
```

```
Run with -LOCAL:      False
Verbosity:            normal

Starting JobDJ...

Keep one job on localhost: False
JobDJ columns:
  C: Number of completed subjobs
  A: Number of active subjobs (e.g., submitted, running)
  W: Number of waiting/pending subjobs

C A W | Activity  JobId  JobName  JobHost
- - - | --------- -----------------------
0 1 1 | launched  myhost-0-4f3ae41e InducedFit1_rec-1_ref-trim_pv-2-out localhost
[myhost]
0 2 0 | launched  myhost-0-4f3ae423 InducedFit1_rec-1_ref-trim_pv-1-out localhost
[myhost]
1 1 0 | finished  myhost-0-4f3ae41e InducedFit1_rec-1_ref-trim_pv-2-out localhost
[myhost]
2 0 0 | finished  myhost-0-4f3ae423 InducedFit1_rec-1_ref-trim_pv-1-out localhost
[myhost]

All jobs are done.

    Structures to be carried forward: 2
Stage completed.  Elapsed time: 156.8 seconds
================================================

================================================
Stage: Scoring
================================================
    Structures to be carried forward: 2
Stage completed.  Elapsed time: 0.4 seconds
================================================

IFD Job Completed at: Tue Feb 14 14:48:23 2012
Output available in: /nfs/zone1/dyall/inducedfit/alpaca/InducedFit1-out.maegz
Total elapsed time: 542.9 seconds
```

5.4 Running Induced Fit Docking from Pregenerated Glide Results

It can be useful to run a separate Glide job instead of running the initial docking through IFD—for example, if you want to use options not available through IFD, such as addition of metal, hydrophobic, or positional constraints. To use the results of a separate Glide run, follow the procedure below:

1. Ensure that the protein is properly prepared for both Glide and Prime.

   To prepare the protein, use the Protein Preparation Wizard panel, which you can open from the Workflows menu.

2. (Optional) Mutate receptor side chains to ALA in your Glide job, by using one of the following approaches:

   - Mutate the residues to ALA in the Build panel, and then change the residue name back to the original residue name.
   - Delete all the side chain atoms beyond the CB, and then apply hydrogen treatment to the CB so that it has three hydrogens.

   Both approaches produce ALA sidechains while retaining the original residue name, which is required in order for Prime to rebuild the full residue during the optimization stage of induced fit docking.

3. Run grid generation and serial Glide docking jobs.

   You can select the desired constraints or other options for the Glide jobs. You should save multiple poses per ligand (the IFD default is 20) and applying IFD scalings (0.5 for the ligand, and 0.7/0.5 for the receptor with/without mutations).

   You must run this initial Glide docking job serially because IFD uses the lignum Glide property to determine which poses were generated from a particular ligand. Running Glide in parallel results in duplicated lignum values, and this would interfere with ligand and pose tracking by `ifd`.

4. Set up an Induced Fit Docking job in Maestro as if you were running all stages from the beginning.

   You can choose any settings for the Initial Glide docking stage, because this stage will not be used. For the ligands to be docked, browse for the ligand file used as input for the Glide job.

5. Write the job files with the Write button, but do not start the job.

6. Edit the *jobname*.inp file:
   a. Comment out the STAGE PPREP section, the STAGE TRIM_SIDECHAINS section and the first STAGE GLIDE_DOCKING section.
   b. For the INPUT_FILE (at the top), change the reference from *jobname*_rec.mae to the pose viewer file from your Glide job, *glidejob*_pv.mae file.
7. Run the Induced Fit Docking job from the command line:

   ```
   $SCHRODINGER/ifd jobname.inp
   ```

You can monitor the job's progress in the *jobname*.log file, or in the Monitor panel. The results can be seen in the structure output file *jobname*-out.maegz.

5.5 Restarting IFD Jobs

If an IFD job fails for some reason, you can restart it from the last completed stage with the following command:

```
ifd -RESTART [options] jobname.inp
```

The file *jobname*.restart must be in the same directory as the input file. This file is generated during the failed run, and contains information on the last completed stage. In addition, the temporary files must be in the *jobname*_workdir subdirectory of the job launch directory. This subdirectory is only present if the driver of the failed job had access to the job launch directory. If it did not (the job was run with -NOLOCAL), the intermediate files are written to a scratch directory, and the job cannot be restarted.

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
- Induced Fit Docking purchaser (company, research institution, or individual)
- Primary Induced Fit Docking 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`.

Index

S

T

V

W

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®