Glide 5.5

User Manual

Schrödinger Press

Pharmacyclics LLC
Exhibit 2037
BeiGene v. Pharmacyclics
PGR2024-00009
EX2037.0001

Glide User Manual Copyright © 2009 Schrödinger, LLC. All rights reserved.

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

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

Schrödinger software includes software and libraries provided by third parties. For details of the copyrights, and terms and conditions associated with such included third party software, see the Legal Notices for Third-Party Software in your product installation at $SCHRODINGER/docs/html/third_party_legal.html (Linux OS) or %SCHRODINGER%\docs\html\third_party_legal.html (Windows OS).

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

June 2009

Contents

Document Conventions vii

Chapter 1: Introduction 1

1.1 Running Schrödinger Software 1

1.2 Glide Panels 2

1.3 Using Glide 2

1.4 Documentation 3

1.5 Citing Glide in Publications 3

Chapter 2: Glide Overview 5

2.1 Introduction to Glide 5

2.2 Glide Constraints 10

2.3 Glide Extra-Precision Mode 10

2.4 Glide/Prime Induced Fit Docking 12

Chapter 3: Protein and Ligand Preparation 13

3.1 Choosing the Most Appropriate Protein Site 13

3.2 Protein Preparation 14

3.3 Checking the Protein Structures 17

3.3.1 Checking the Orientation of Water Molecules 17

3.3.2 Checking for Steric Clashes 17

3.3.3 Resolving H-Bonding Conflicts 18

3.3.4 Docking the Native Ligand 19

3.4 Ligand Preparation 19

3.4.1 Using LigPrep for Ligand Preparation 20

3.4.2 Using Other Programs for Ligand Preparation 22

Chapter 4: Receptor Grid Generation 25

4.1 The Receptor Grid Generation Panel 25

EX2037.0003

4.2 The Receptor Tab 27
4.2.1 Defining the Receptor 27
4.2.2 Van der Waals Radius and Charge Scaling 27
4.2.3 Selection of Partial Charges 29
4.3 The Site Tab 30
4.3.1 Selecting a Box Center 32
4.3.2 Setting the Box Sizes 33
4.4 The Constraints Tab 34
4.4.1 Setting Positional Constraints 35
4.4.2 Setting H-Bond and Metal Constraints 37
4.4.3 Setting Hydrophobic Constraints 39
4.4.3.1 Generating the Hydrophobic Map 39
4.4.3.2 Defining Hydrophobic Constraint Regions 39
4.5 The Rotatable Groups Tab 42

Chapter 5: Ligand Docking **45**

5.1 The Ligand Docking Panel 45
5.2 The Settings Tab 47
5.2.1 Specifying the Receptor Grid 47
5.2.2 Selecting the Docking Precision 48
5.2.3 Setting Docking Options 48
5.2.4 Writing XP Descriptor Information 50
5.2.5 Adding Epik State Penalties to the Docking Score 50
5.2.6 Advanced Options 50
5.2.6.1 Selection of Initial Poses 51
5.2.6.2 Energy Minimization Settings 52
5.3 The Ligands Tab 52
5.3.1 Specifying the Source of the Ligands 53
5.3.2 Selecting the Source of Ligand Partial Charges 54
5.3.3 Setting Restrictions on the Type of Ligands 54
5.3.4 Van der Waals Radii Scaling 55
5.4 The Core Tab 55

EX2037.0004

5.5 The Constraints Tab 58
5.5.1 Setting Constraints 59
5.5.2 Defining Ligand Features 61
5.5.2.1 Loading and Saving Feature Sets 62
5.5.2.2 Adding, Editing, and Deleting Patterns 62
5.5.2.3 Excluding Functional Groups from a Feature 64
5.5.2.4 Visualizing Patterns 64
5.5.3 Choosing When To Apply Constraints 64
5.5.4 Using Multiple Constraints 64
5.6 The Similarity Tab 65
5.6.1 Setting up Similarity Scoring 66
5.6.2 Creating a Trained Similarity Model 67
5.6.3 Similarity Scoring Function 68
5.7 The Output Tab 70
5.7.1 Structure Output Options 70
5.7.2 Post-Docking Minimization 72
5.7.3 Per-Residue Interaction Scores 73
5.7.4 Advanced Settings 73
5.8 Docking Output Properties 75

Chapter 6: Visualizing Docking Results **77**

6.1 Viewing Poses 77
6.2 The Glide XP Visualizer 79
6.2.1 The Ligands Table 80
6.2.2 Controlling the Display 82
6.2.3 Exporting XP Terms 83
6.2.4 XP Terms and Their Visualizations 83

Chapter 7: Running Glide from the Command Line **87**

7.1 Running Jobs 88
7.2 Job Files and Directories 89

7.3 Receptor Grid Generation and Ligand Docking 90
7.3.1 The glide Command 90
7.3.2 The impact Command 97
7.3.3 The para_glide Command 97
7.3.4 Restarting Glide Docking Jobs 100
7.4 Glide Utilities 101
7.4.1 glide_sort 101
7.4.2 glide_merge 103
7.4.3 glide_rescore 104
7.5 Examining Results From the Command Line 104
7.6 Customizing Glide Calculations 105
7.6.1 Changing the Glide Atom Typing 105
7.7 Docking Log Messages 105
References **107**
Getting Help **109**
Index **113**

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, and screen 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`.

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].

EX2037.0007

EX2037.0008

Introduction

The *Glide User Manual* is intended to help you perform ligand database screening and high-accuracy docking with Glide. Glide is run primarily from the Maestro graphical user interface, but can also be run from the command line. Online help for Glide is available in Maestro, although the information in this manual is generally more comprehensive.

Chapter 2 introduces the scientific methods and computational procedures used in Glide.

Chapter 3 describes the preparation of the protein and the ligands for use in Glide.

Chapter 4 describes the use of the Receptor Grid Generation panel to calculate the grids that represent the receptor.

Chapter 5 describes the use of the Ligand Docking panel to set up and run docking jobs, and the use of atom-pair similarity, Glide constraints, and distributed processing of multiple-ligand docking calculations.

Chapter 6 contains information on visualizing the results of Glide docking runs, using the Glide Pose Viewer, the Project Table, and the Glide XP Visualizer.

Chapter 7 contains information about running Glide, and its associated applications and utilities, from the command line.

The *Glide Quick Start Guide* contains tutorials intended to familiarize you with protein preparation, receptor grid generation, ligand docking, and the Pose Viewer.

1.1 Running Schrödinger Software

To run any Schrödinger program on a UNIX platform, or start a Schrödinger job on a remote host from a UNIX 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 start Maestro with the following command:

EX2037.0009

```
$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. For more information on starting Maestro, including starting Maestro on a Windows platform, see Section 2.1 of the *Maestro User Manual*.

1.2 Glide Panels

This section describes common features of the two main Glide panels. Controls that are specific to a panel appear in the middle section of the panel and are contained in a series of tabs.

In the lower part of the two main panels, the following buttons appear:

- Start—Opens the Start dialog box, in which you can set up job parameters and submit jobs. The general description is given in Section 2.2 of the *Job Control Guide*. Specific features that appear in this dialog box are described in the chapters on the Glide panels.
- Write—Writes out all the files required for the job, but does not start the job. The Glide job can then be run from the command line using the `impact` command. See Chapter 7 for details.
- Reset—Resets the options in all the tabs to the default values.

1.3 Using Glide

The Glide task most frequently performed is ligand docking. The grid files produced by a single receptor grid generation task can be used for any number of jobs that dock ligands to that receptor. Before generating receptor grids, it is strongly recommended that you prepare the protein. Therefore, the first steps of a typical project beginning with an unprepared protein-ligand complex structure (e.g., from PDB) might proceed using the Glide panels as follows:

1. After correcting formal charges and bond orders in the ligand, cofactors, and nonstandard residues, set up and start the automated preparation and refinement portions of the protein preparation procedure using the Protein Preparation panel. Details of protein preparation are given in Chapter 3.
2. Ensure that the ligands to be docked are in the right form. Details of this process are given in Chapter 3.
3. With the prepared receptor-ligand complex in the Workspace, use the Receptor Grid Generation panel to specify settings, optionally define constraints, and start the receptor grid generation job. Details of setting up receptor grid generation jobs are given in Chapter 4.

EX2037.0010

4. Specify the base name for the receptor grid files you want to use in the Ligand Docking panel, and use the other settings and options in the panel to set up and start a ligand docking job. As many docking jobs as you want can be set up in this panel, using the current receptor grids or specifying a different set of grids to use. Details of setting up ligand docking jobs are given in Chapter 5.

Glide also has a web interface that you can use to submit docking jobs with a simplified set of input options, for one of a set of previously prepared grids. The results are presented in your browser. The web interface can be obtained from the Download page on the Schrödinger web site. Instructions are included in the download in an HTML document.

1.4 Documentation

For information related to the installation and use of Glide, see the following documentation:

- The *Installation Guide*, which includes installation instructions for all Schrödinger products and documentation.
- The *Job Control Guide*, which includes instructions for running and managing Schrödinger jobs.
- The *Glide Quick Start Guide*, which contains tutorials intended to familiarize you with protein preparation, receptor grid generation, ligand docking, and visualization of results.
- The *Impact Command Reference Manual*, which contains syntax and keywords for Impact command input files.
- The *Maestro User Manual*, which describes how to use the features of Maestro, including the Atom Selection dialog box. An appendix describes command-line utilities, many of which are used in Glide.
- The *Maestro Command Reference Manual*, which contains commands, options, and arguments for running Maestro from the command line, including the Atom Specification Language (ASL) and the Entry Specification Language (ESL).

1.5 Citing Glide in Publications

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

Glide, version 5.5, Schrödinger, LLC, New York, NY, 2009.

EX2037.0011

EX2037.0012

Glide Overview

This chapter contains an overview of the Glide (Grid-based Ligand Docking with Energetics) program, its scientific methods and computational procedures.

Glide searches for favorable interactions between one or more ligand molecules and a receptor molecule, usually a protein. Each ligand must be a single molecule, while the receptor may include more than one molecule, e.g., a protein and a cofactor. Glide can be run in rigid or flexible docking modes; the latter automatically generates conformations for each input ligand. The combination of position and orientation of a ligand relative to the receptor, along with its conformation in flexible docking, is referred to as a *ligand pose*.

The ligand poses that Glide generates pass through a series of hierarchical filters that evaluate the ligand's interaction with the receptor. The initial filters test the spatial fit of the ligand to the defined active site, and examine the complementarity of ligand-receptor interactions using a grid-based method patterned after the empirical ChemScore function [1].

Poses that pass these initial screens enter the final stage of the algorithm, which involves evaluation and minimization of a grid approximation to the OPLS-AA nonbonded ligand-receptor interaction energy.

Final scoring is then carried out on the energy-minimized poses. By default, Schrödinger's proprietary GlideScore multi-ligand scoring function is used to score the poses. If GlideScore was selected as the scoring function, a composite *Emodel* score is then used to rank the poses of each ligand and to select the poses to be reported to the user. Emodel combines GlideScore, the nonbonded interaction energy, and, for flexible docking, the excess internal energy of the generated ligand conformation.

2.1 Introduction to Glide

Glide uses a hierarchical series of filters to search for possible locations of the ligand in the active-site region of the receptor. The shape and properties of the receptor are represented on a grid by several different sets of fields that provide progressively more accurate scoring of the ligand poses.

Conformational flexibility is handled in Glide by an extensive conformational search, augmented by a heuristic screen that rapidly eliminates unsuitable conformations, such as conformations that have long-range internal hydrogen bonds.

EX2037.0013

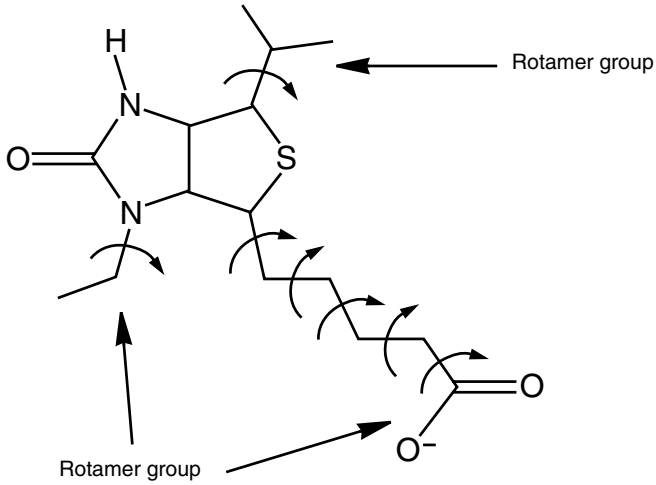


Figure 2.1. Definition of core and rotamer groups.

As illustrated in Figure 2.1, each ligand is divided into a *core* region and some number of *rotamer groups*. Each rotamer group is attached to the core by a rotatable bond, but does not contain additional rotatable bonds. The core is what remains when each terminus of the ligand is severed at the "last" rotatable bond. Carbon and nitrogen end groups terminated with hydrogen (—CH_3, —NH_2, —NH_3^+) are not considered rotatable because their conformational variation is of little significance. In Figure 2.1, the four central torsions are part of the core, and the methyl groups are not considered rotatable.

During conformation generation, each core region is represented by a set of core conformations, the number of which depends on the number of rotatable bonds, conformationally labile 5– and 6–membered rings, and asymmetric pyramidal trigonal nitrogen centers in the core. This set typically contains fewer than 500 core conformations, even for quite large and flexible ligands, and far fewer for more rigid ligands. Every rotamer state for each rotamer group attached to the core is enumerated. The core plus all possible rotamer-group conformations is docked as a single object. Glide can also dock sets of pre-computed conformations. However, Glide offers its greatest value when conformations are generated internally.

For each core conformation (or for rigid docking, each ligand), an exhaustive search of possible locations and orientations is performed over the active site of the protein. The search begins with the selection of "site points" on an equally spaced 2 Å grid that covers the active-site region (Stage 1 in Figure 2.2). This selection is made as follows. Distances from the site point to the receptor surface are evaluated at a series of pre-specified directions and sorted into

EX2037.0014

distance ranges ("bins") of width 1 Å. Likewise, distances from the ligand center (the midpoint of the two most widely separated atoms) to the ligand surface are sorted into bins of width 1 Å. For a given site point, the distance ranges from the site point to the receptor are compared with those from the ligand center to the ligand surface. Glide positions the ligand center at the site point if there is a good enough match, but skips over the site point if there is not.

The second stage of the hierarchy begins by examining the placement of atoms that lie within a specified distance of the line drawn between the most widely separated atoms (the *ligand diameter*). This is done for a pre-specified selection of possible orientations of the ligand diameter (Step 2a). If there are too many steric clashes with the receptor, the orientation is skipped.

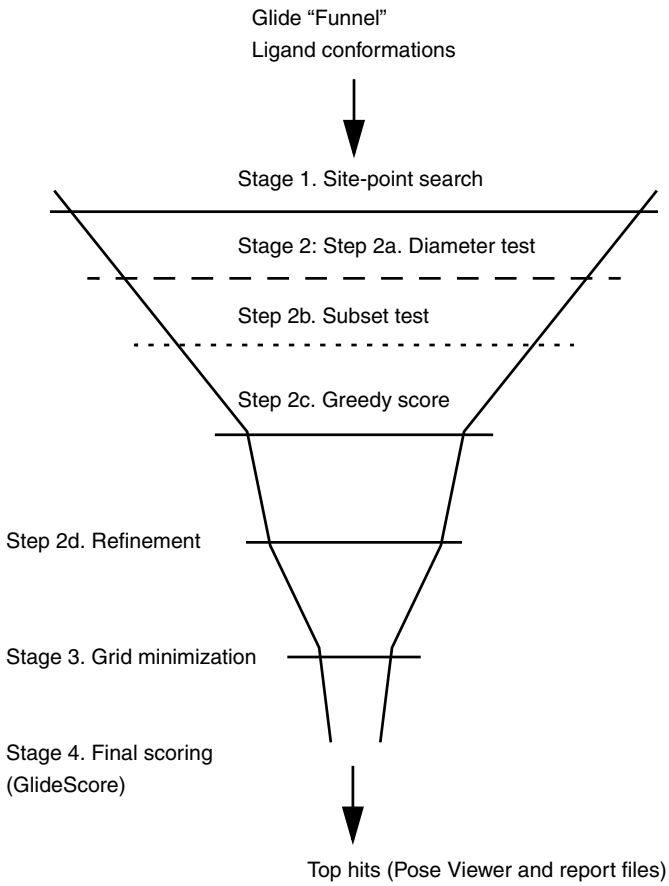


Figure 2.2. The Glide docking hierarchy.

EX2037.0015

Next (Step 2b), rotation about the ligand diameter is considered, and the interactions of a subset consisting of all atoms capable of making hydrogen bonds or ligand-metal interactions with the receptor are scored (*subset test*). If this score is good enough, all interactions with the receptor are scored (Step 2c).

The scoring in these three tests is carried out using Schrödinger's discretized version of the ChemScore empirical scoring function [1]. Much as for ChemScore itself, this algorithm recognizes favorable hydrophobic, hydrogen-bonding, and metal-ligation interactions, and penalizes steric clashes. This stage is called "greedy scoring," because the actual score for each atom depends not only on its position relative to the receptor but also on the best possible score it could get by moving ±1 Å in *x*, *y*, or *z*. This is done to mute the sting of the large 2 Å jumps in the site-point/ligand-center positions. The final step in Stage 2 is to re-score the top greedy-scoring poses via a "refinement" procedure (Step 2d), in which the ligand as a whole is allowed to move rigidly by ±1 Å in the Cartesian directions.

Only a small number of the best refined poses (typically 100-400) is passed on to the third stage in the hierarchy—energy minimization on the pre-computed OPLS-AA van der Waals and electrostatic grids for the receptor. The energy minimization typically begins on a set of van der Waals and electrostatic grids that have been "smoothed" to reduce the large energy and gradient terms that result from too-close interatomic contacts. It finishes on the full-scale OPLS-AA nonbonded energy surface ("annealing"). This energy minimization consists only of rigid-body translations and rotations when external conformations are docked. When conformations are generated internally, however, the optimization also includes torsional motion about the core and end-group rotatable bonds. Unless otherwise specified, a small number of the top-ranked poses are then subjected to a sampling procedure in which alternative local-minima core and rotamer-group torsion angles are examined to try to improve the energy score.

Finally, the minimized poses are re-scored using Schrödinger's proprietary *GlideScore* scoring function. GlideScore is based on ChemScore, but includes a steric-clash term, adds buried polar terms devised by Schrödinger to penalize electrostatic mismatches, and has modifications to other terms:

GScore = 0.065*vdW + 0.130*Coul + Lipo + Hbond + Metal + BuryP + RotB + Site

The components of the GlideScore (GScore) are described in Table 2.1.

The choice of best-docked structure for each ligand is made using a model energy score (*Emodel*) that combines the energy grid score, the binding affinity predicted by GlideScore, and (for flexible docking) the internal strain energy for the model potential used to direct the conformational-search algorithm. Glide also computes a specially constructed Coulomb-van der Waals interaction-energy score (CvdW) that is formulated to avoid overly rewarding charge-charge interactions at the expense of charge-dipole and dipole-dipole interactions. This

EX2037.0016

Table 2.1. GlideScore components.

Component	Description
vdW	Van der Waals energy. This term is calculated with reduced net ionic charges on groups with formal charges, such as metals, carboxylates, and guanidiniums.
Coul	Coulomb energy. This term is calculated with reduced net ionic charges on groups with formal charges, such as metals, carboxylates, and guanidiniums.
Lipo	Lipophilic term derived from hydrophobic grid potential. Rewards favorable hydrophobic interactions.
HBond	Hydrogen-bonding term. This term is separated into differently weighted components that depend on whether the donor and acceptor are neutral, one is neutral and the other is charged, or both are charged.
Metal	Metal-binding term. Only the interactions with anionic acceptor atoms are included. If the net metal charge in the apo protein is positive, the preference for anionic ligands is included; if the net charge is zero, the preference is suppressed.
BuryP	Penalty for buried polar groups.
RotB	Penalty for freezing rotatable bonds.
Site	Polar interactions in the active site. Polar but non-hydrogen-bonding atoms in a hydrophobic region are rewarded.

score is intended to be more suitable for comparing the binding affinities of different ligands than is the “raw” Coulomb-van der Waals interaction energy. In the final data work-up, you can combine the computed GlideScore and “modified” Coulomb-van der Waals score values to give a composite score that usually helps improve enrichment factors in database screening applications.

This hierarchical search gives Glide exceptionally high accuracy in predicting the binding mode of the ligand. At the same time, the computational cost is dramatically reduced compared to what would be required for a complete systematic search. The key to this reduction is that the algorithm allows the rotamer groups to be optimized one at a time for a given core conformation and location of the ligand. For example, if there are five rotamer groups and each has three rotamer states, the total number of conformers in the ensemble based on this core conformation/location is $3^5 = 243$. However, if the rotamer groups are optimized one at a time, the number of conformational combinations is only $3 \times 5 = 15$, for a savings of about a factor of 15 in computational effort. While many other time-saving algorithms in Glide contribute to its performance advantages, this fundamental qualitative feature allows large libraries to be screened at an affordable computational cost.

EX2037.0017

2.2 Glide Constraints

A *Glide constraint* is a ligand-receptor interaction requirement. The constraint usually means that a ligand atom must lie within a certain region defined in relation to features of the receptor that are responsible for ligand binding. To use Glide constraints, you must specify receptor sites for possible ligand interactions when you set up a receptor grid generation job. When you run a ligand docking job, you can select Glide constraints to apply from the list of receptor constraint sites that you defined for the receptor.

In Glide constraint docking jobs, Glide incorporates satisfaction of these constraints into several of its hierarchical filters, allowing prompt rejection of docked poses that fail to meet the requirements.

The first constraint filter is a simple one involving the atoms of the ligand. If a ligand does not contain atoms of the right types to make the required interactions with the receptor constraint atoms, Glide simply skips that ligand—for instance if a selected receptor constraint atom is a polar hydrogen and the ligand has no hydrogen-bond acceptors. If there are qualifying atoms in the ligand, Glide keeps a list of the possible "partner atoms" for each constraint, for use in subsequent filters. For hydrophobic constraints, this filter checks that the ligand contains a sufficient number of hydrophobic heavy atoms; the sum of the minimum numbers specified for all the selected constraints of this type. In addition, Glide locates centroids of ligand hydrophobic groups, which take the role of partner atoms in subsequent filters.

Other filters operate further down the funnel. These include matching distances between partner atoms (or hydrophobic group centroids) for different constraints against the corresponding distances between receptor atoms; matching distances from the ligand center to partner atoms against those from receptor site points to constraint atoms or hydrophobic regions, and restricting orientations of, and rotations about, the ligand diameter to those that bring partner atoms into appropriate proximity with receptor atoms or regions. In addition, a restraining potential is used in grid optimizations (for hydrogen-bond and metal constraints only) to bring or keep constraint-satisfying ligand-receptor atom pairs at appropriate distances, and torsional sampling moves are not tried if they would move any ligand atom currently in a constraint-satisfying position. Finally, grid-optimized poses are rejected if they don't strictly satisfy all selected constraints.

For information on using Glide constraints, see Section 4.4 and Section 5.5.

2.3 Glide Extra-Precision Mode

The extra-precision (XP) mode of Glide combines a powerful sampling protocol with the use of a custom scoring function designed to identify ligand poses that would be expected to have unfavorable energies, based on well-known principles of physical chemistry. The presumption

EX2037.0018

is that only active compounds will have available poses that avoid these penalties and also receive favorable scores for appropriate hydrophobic contact between the protein and the ligand, hydrogen-bonding interactions, and so on. The chief purposes of the XP method are to weed out false positives and to provide a better correlation between good poses and good scores.

Extra-precision mode is a refinement tool designed for use only on good ligand poses. The more extensive XP docking method and specialized XP scoring method are strongly coupled: the more precise poses produced by XP docking are necessary for the more demanding XP scoring method. Because XP docking mode requires considerably more CPU time, you should screen large sets of ligands first in standard-precision (SP) mode or in high-throughput virtual screening (HTVS) mode. Only the top-scoring ligands should be docked using XP mode.

In any grid-based docking method, the receptor is essentially frozen. Some degree of flexibility can be introduced by scaling parts of the potential, as is done in SP mode, but such techniques cannot represent larger adjustments to the receptor, like conformational changes. XP mode is less forgiving than SP mode so that it can screen out false positives, and is designed to locate active compounds that bind to a particular conformation of the receptor. Active compounds can be prevented from docking if these compounds are not compatible with the particular conformation of the receptor that is being used. To ensure that actives are not eliminated you should if possible dock into multiple receptor conformations and combine the results of the individual docking runs.

The XP sampling method is based on an anchor and refined growth strategy. Anchor fragments of the docked ligand, typically rings, are chosen from the set of SP poses and the molecule is re-grown bond by bond from these anchor positions. Complete minimizations and XP scoring are carried out on the large ensemble of poses generated by this growing method. At various cycles the growing is focused to attempt to relieve any XP scoring penalties as well as to optimize the best scoring poses. This focused sampling is essential for allowing the use of the hard XP scoring function as well as for finding the best scoring basins of attraction. It is important to note that the coupling between the extra sampling and the XP scoring means that it is not recommended to just score the SP poses with XP scoring.

The scoring function (GlideScore XP) includes additional terms over the SP scoring function, and a more complete treatment of some of the SP terms, as described below.

To model solvation, explicit water molecules are docked into a list of protein-ligand complexes that otherwise receive good GlideScores, and descriptors based on the interaction of these water molecules with various charged and polar groups of the ligand and protein are used as a measure of whether the complex is physically realistic. Penalties are assigned to structures where statistical results indicate that one or more groups is inadequately solvated. A large database of co-crystallized structures has been used to optimize the parameters associated with the

EX2037.0019

penalty terms. The explicit-water technology and descriptors are also used in Glide SP scoring. However, the improved sampling allows XP docking to assign substantially higher penalties to serious violations of physical principles.

GlideScore XP specifically rewards occupancy of well-defined hydrophobic pockets by hydrophobic ligand groups. Hydrophobic reward terms are employed in empirical scoring functions such as ChemScore and the SP version of GlideScore in the form of lipophilic-lipophilic pair terms, while other empirical scoring functions use lipophilic surface-area contact terms for this purpose. Investigations have shown that simple pair terms underestimate hydrophobic effects in certain well-defined cases. The hydrophobic term in GlideScore XP was developed to offset this underestimation. The term can confer up to several kcal/mol of additional binding energy in favorable cases, and substantially improves enrichment factors GlideScore XP also includes improvements to the scoring of hydrogen bonds as well as detection of buried polar groups, and detection of pi-cation and pi-pi stacking interactions.

XP mode can be used when the active site of the complex contains a metal and often works well, but has not yet been optimized for such applications.

For information on using XP mode, see Section 5.2. For information on visualizing the various XP scoring terms, see Section 6.2.

2.4 Glide/Prime Induced Fit Docking

Glide docking uses the basic assumption of a rigid receptor. Scaling of van der Waals radii of nonpolar atoms, which decreases penalties for close contacts, can be used to model a slight “give” in the receptor or the ligand or both, and specified hydroxyl groups can be allowed to reorient to optimize hydrogen bonding. This may not be sufficient to treat systems where ligand binding induces substantial conformation changes in the receptor (“induced fit.”) Schrödinger has developed a procedure for such cases which uses Prime and Glide to perform induced fit docking. For more information about Schrödinger’s Induced Fit Docking protocol, see the document *Induced Fit Docking*.

EX2037.0020

Protein and Ligand Preparation

The quality of Glide results depends on reasonable starting structures for both the protein and the ligand. Schrödinger offers a comprehensive protein preparation facility in the Protein Preparation Wizard, which is designed to ensure chemical correctness and to optimize protein structures for use with Glide and other products. Likewise, Schrödinger offers a comprehensive ligand preparation facility in LigPrep. It is strongly recommended that you process protein and ligand structures with these facilities in order to achieve the best results.

3.1 Choosing the Most Appropriate Protein Site

Glide relies on the rigid-receptor approximation in order to treat protein-ligand binding. However, protein conformations are in general flexible and can occupy a continuum of states. Proteins can exhibit induced fit effects on binding of a ligand, in which the protein conformation changes significantly. This effect can be seen in cases where more that one co-crystallized complex is available. Since Glide docking experiments employ only a single protein geometry, two approaches are commonly taken. One is to select a single well-suited representative protein structure to dock into. The other is to use an ensemble of representative structures, into which each of the candidate ligands is docked.

If more than one co-crystallized complex is available, you must decide whether to select a single protein site or to choose two or more sites for use in independent docking experiments. This choice may depend on whether the protein site is rigid or mobile, as well as on the resources available and the objectives of the docking study. A single representative site should suffice for a rigid protein. Cases in which the site changes substantially as different ligands bind may require the use of two or more sites, if finding the maximum number of promising ligands is the main objective.

Some proteins are known to be rigid. To determine whether a single site is likely to be sufficient in other cases, you can transpose the known protein-ligand complexes into the coordinate frame of a “reference” complex. The objective is to judge whether the reference site is compatible with all the co-crystallized ligands or whether another site appears more suitable. A procedure for making this determination is as follows:

1. Choose a reference complex and superimpose all the other complexes to it.

   You can perform this task in Maestro with the Superposition panel. The atoms you choose to superimpose could be the alpha carbons, or all the backbone atoms for residues

EX2037.0021

in common, or the Cα or backbone atoms for selected active-site residues. Alternatively, you can use the Protein Structure Alignment facility to perform the superposition in an automated fashion.

2. Display the protein for the reference complex and the ligand for each of the other complexes, in turn. Examine the active-site region to determine whether the superimposed ligand can fit into the reference site without steric clashes that could not reasonably be relieved by minor repositionings. You can use the tools in the Contacts folder of the Measurements panel to examine contacts.
3. Display the protein for the reference complex and the protein for each of the other complexes in turn. Note whether any residues in the superimposed protein differ appreciably in position or conformation from those in the reference site.
4. From the above steps, judge whether the reference site appears compatible with all the co-crystallized ligands or, if not, whether another site appears more compatible.
5. Choose a most representative (i.e., fairly generous) site for docking—or choose two or more sites if there are large differences between the sites and the objective is to find as many prospective strong binders as possible.
6. Write out a separate file for the protein or proteins that will be prepared. Also write out separate files for the superimposed ligands.

For example, an initial screening of the CDK-2 kinase receptor used 1hck as the docking site, the co-crystallized ligand for which is ATP. Other known co-crystallized complexes include 1aq1, 1di8, 1dm2, 1fvt, and 1fvv. Superimposing these five complexes onto 1hck using all Cα atoms in common revealed that at least four of the five ligands (all active binders) cannot fit into the 1hck site because its active-site channel is too short. The reason is that ATP and a Mg^{2+} ion bound to its terminal phosphate group pull glutamate and lysine sidechains more deeply into the 1hck cavity, where they form an ion pair that closes off the cleft. Based on this visual examination, the 1dm2 site was chosen instead. This site is considerably more open than the 1hck site, though somewhat less so than the 1aq1 or 1fvv sites. Glide was far more successful in docking the known binders into the 1dm2 site than into 1hck, and was even more successful when the still more open 1aq1 site was used.

3.2 Protein Preparation

A typical PDB structure file consists only of heavy atoms, can contain waters, cofactors, and metal ions, and can be multimeric. The structure generally has no information on bond orders, topologies, or formal atomic charges. Terminal amide groups can also be misaligned, because the X-ray structure analysis cannot usually distinguish between O and NH_2. Ionization and tautomeric states are also generally unassigned. Glide calculations use an all-atom force field

EX2037.0022

for accurate energy evaluation. Thus, Glide requires bond orders and ionization states to be properly assigned and performs better when side chains are reoriented when necessary and steric clashes are relieved.

This section provides an overview of the protein preparation process. The entire procedure can be performed in the Protein Preparation Wizard panel, which you open from the Workflows menu on the main toolbar. This tool and its use is described in detail in Chapter 2 of the *Protein Preparation Guide*.

After processing, you will have files containing refined, hydrogenated structures of the ligand and the ligand-receptor complex. The prepared structures are suitable for use with Glide. In most cases, not all of the steps (outlined below) need to be performed. See the descriptions of each step to determine whether it is required, and make the appropriate selections in the Protein Preparation Wizard panel.

You may on occasion want to perform some of these steps manually. Detailed procedures for manual preparation are described in Chapter 3 of the *Protein Preparation Guide*.

The steps in the procedure, which can be performed either in the Protein Preparation Wizard panel or manually, are listed below.

1. Import a ligand/protein cocrystallized structure, typically from PDB, into Maestro.

   The preparation component of the protein preparation facility requires an identified ligand.

2. Simplify multimeric complexes.

   For computational efficiency it is desirable to keep the number of atoms in the complex structure to a minimum. If the binding interaction of interest takes place within a single subunit, you should retain only one ligand-receptor subunit to prepare for Glide. If two identical chains are both required to form the active site, neither should be deleted.

   - Determine whether the protein-ligand complex is a dimer or other multimer containing duplicate binding sites and duplicate chains that are redundant.
   - If the structure is a multimer with duplicate binding sites, remove redundant binding sites and the associated chains by picking and deleting molecules or chains.

3. Locate any waters you want to keep, then delete all others.

   These waters are identified by the oxygen atom, and usually do not have hydrogens attached. Generally, all waters (except those coordinated to metals) are deleted, but waters that bridge between the ligand and the protein are sometimes retained. If waters are kept, hydrogens will be added to them by the preparation component of the protein preparation job. Afterwards, check that these water molecules are correctly oriented.

EX2037.0023

4. Adjust the protein, metal ions, and cofactors.

   Problems in the PDB protein structure may need to be repaired before it can be used. Incomplete residues are the most common errors, but may be relatively harmless if they are distant from the active site. Structures that are missing residues near the active site should be repaired.

   Metal ions in the protein complex cannot have covalent bonds to protein atoms. The MacroModel atom types for metal ions are sometimes incorrectly translated into dummy atom types (Du, Z0, or 00) when metal-protein bonds are specified in the input structure. Furthermore, isolated metal ions may erroneously be assigned general atom types (GA, GB, GC, etc.).

   It may be necessary to adjust the protonation of the protein, which is crucial when the receptor site is a metalloprotein such as thermolysin or an MMP. In such a case, Glide assigns a special stability to ligands in which anions coordinate to the metal center. To benefit from this assignment, groups such as carboxylates, hydroxamates, and thiolates must be anionic. The protein residues that line the approach to the metal center (such as Glu 143 and His 231 in thermolysin) need to be protonated in a manner compatible with the coordination of an anionic ligand such as a carboxylate or hydroxamate. The co-crystallized complex therefore needs to be examined to determine how the protein and the ligands should be protonated. In some cases, two or more protonation states of the protein may need to be used in independent docking experiments to cover the range of physically reasonable ligand dockings.

   Cofactors are included as part of the protein, but because they are not standard residues it is sometimes necessary to use Maestro's structure-editing capabilities to ensure that multiple bonds and formal charges are assigned correctly.

   - Fix any serious errors in the protein.
   - Check the protein structure for metal ions and cofactors.
   - If there are bonds to metal ions, delete the bonds, then adjust the formal charges of the atoms that were attached to the metal as well as the metal itself.
   - Set charges and correct atom types for any metal atoms, as needed.
   - Set bond orders and formal charges for any cofactors, as needed.

5. Adjust the ligand bond orders and formal charges.

   If the complex structure contains bonds from the ligand or a cofactor to a protein metal, they must be deleted. Glide models such interactions as van der Waals plus electrostatic interactions. Glide cannot handle normal covalent bonds to the ligand, such as might be found in an acyl enzyme.

EX2037.0024

If you are working with a dimeric or large protein and two ligands exist in two active sites, the bond orders have to be corrected in both ligand structures.

6. Run a restrained minimization of the protein structure.

   This is done with `impref`, and should reorient side-chain hydroxyl groups and alleviate potential steric clashes. The minimization is restrained to the input protein coordinates by a user-selected RMSD tolerance.

7. Review the prepared structures.

   - If problems arise during the restrained minimization, review the log file, correct the problems, and rerun.
   - Examine the refined ligand/protein/water structure for correct formal charges, bond orders, and protonation states and make final adjustments as needed.

3.3 Checking the Protein Structures

After you have completed the protein preparation, you should check the completed ligand and protein structures.

3.3.1 Checking the Orientation of Water Molecules

You only need to perform this step if you kept some structural waters. Reorienting the hydrogens is not strictly necessary, as their orientation should have been changed during refinement, but it is useful to check that the orientation is correct.

If the orientation is incorrect, reorient the molecules by using the procedure outlined in Section 3.9 of the *Protein Preparation Guide*.

When you have corrected the orientation of the retained water molecules, you should run a refinement on the adjusted protein-ligand complex.

3.3.2 Checking for Steric Clashes

You should make sure that the prepared site accommodates the co-crystallized ligand in the restraint-optimized geometry obtained from the structure preparation.

Steric clashes can be detected by displaying the ligand and protein in Maestro and using the Contacts folder in the Measurements panel to visualize bad or ugly contacts. Maestro defines bad contacts purely on the basis of the ratio of the interatomic distance to the sum of the van der Waals radii it assigns. As a result, normal hydrogen bonds are classified as bad or ugly contacts. By default, Maestro filters out contacts that are identified as hydrogen bonds, and displays only the genuine bad or ugly contacts.

EX2037.0025

If steric clashes are found, repeat the restrained optimization portion of the protein preparation procedure, but allow a greater rms deviation from the starting heavy-atom coordinates than the default of 0.3 Å. Alternatively, you can apply an additional series of restrained optimizations to the prepared ligand-protein complex to allow the site to relax from its current geometry.

3.3.3 Resolving H-Bonding Conflicts

You should look for inconsistencies in hydrogen bonding to see whether a misprotonation of the ligand or the protein might have left two acceptor atoms close to one another without an intervening hydrogen bond. One or more residues may need to be modified to resolve such an acceptor-acceptor or donor-donor clash.

Some of these clashes are recognized by the preparation process but cannot be resolved by it. The preparation process may have no control over other clashes. An example of the latter typically occurs in an aspartyl protease such as HIV, where both active-site aspartates are close to one or more atoms of a properly docked ligand. Because these contact distances fall within any reasonable cavity radius, the carboxylates are not subject to being neutralized and will both be represented as negatively charged by the preparation process. However, when the ligand interacts with the aspartates via a hydroxyl group or similar neutral functionality, one of the aspartates is typically modeled as neutral.

If residues need to be modified, follow these steps:

1. Place the refined protein-ligand complex in the Workspace.
2. Examine the interaction between the ligand and the protein (and/or the cofactor).
3. Use your judgment and chemical intuition to determine which protonation state and tautomeric form the residues in question should have.
4. Use the structure-editing capabilities in Maestro to resolve the conflict (see Section 3.8 of the *Protein Preparation Guide* for procedures).
5. Re-minimize the structure.

It is usually sufficient to add the proton and perform about 50 steps of steepest-descent minimization to correct the nearby bond lengths and angles. Because this optimizer does not make large-scale changes, the partial minimization can be done even on the isolated ligand or protein without danger of altering the conformation significantly. However, if comparison to the original complex shows that the electrostatic mismatch due to the misprotonation has appreciably changed the positions of the ligand or protein atoms during the protein-preparation procedure, it is best to reprotonate the original structure and redo the restrained minimization.

EX2037.0026

3.3.4 Docking the Native Ligand

Once you have prepared the protein and generated grids (see Chapter 4), you should dock the native ligand both rigidly and flexibly. If either run fails to produce a low-rms structure, the structure or structures obtained from the docking should be scored in place, and the outcomes in the `*.rept` files for the dockings and in the `*.scor` files for the score-in-place runs should be checked.

Before running a docking job, you can run a score-in-place Glide calculation on the complex and check that the metal-ligation energy is reasonable. If it is highly positive, you may have to re-adjust the charge and protonation states in the active site manually.

- If rigid docking fails to give a low-rms pose, there may be a problem with the structure preparation. Double-check the work done to prepare the protein. Such a failure may also reflect incomplete sampling by Glide. If the problem is sampling, there are some settings in the GUI, and others that can be made in the input file for Glide, that can be used to increase the amount of sampling.
- If rigid docking succeeds but flexible docking gives a high-rms pose that has a poor Coulomb-vdW energy, there may also be a sampling problem.
- If rigid docking succeeds but flexible docking finds a high-rms solution that has a more negative Coulomb-vdW energy than the ligand pose generated in the structure preparation, there may be a scoring problem, that is, Glide may have chosen the wrong solution because its Coulomb-vdW and Emodel energies are more favorable than those for the correct solution. There usually is no way you can fix such a problem. However, if the GlideScore value computed for the misdocked ligand is similar to that obtained for the co-crystallized ligand, the database screen may still be effective in identifying ligands that can bind tightly. The ligand may just have found a plausible alternative binding mode.

3.4 Ligand Preparation

To give the best results, the structures that are docked must be good representations of the actual ligand structures as they would appear in a protein-ligand complex. This means that the structures supplied to Glide must meet the following conditions:

1. They must be three-dimensional (3D).
2. They must have realistic bond lengths and bond angles.

   Glide only modifies the torsional internal coordinates of the ligand during docking, so the rest of the geometric parameters must be optimized beforehand.
3. They must each consist of a single molecule that has no covalent bonds to the receptor,

EX2037.0027

with no accompanying fragments, such as counter ions and solvent molecules.

4. They must have all their hydrogens (filled valences).

5. They must have an appropriate protonation state for physiological pH values (around 7).

   For example, carboxylic acids should be deprotonated and aliphatic amines should be protonated. Otherwise a neutral aliphatic amine could improperly act as a hydrogen-bond acceptor in the docking calculations, or could occupy a hydrophobic region without incurring the large desolvation penalty that XP Glide docking would have assessed if the amine had been properly protonated.

   Protonation states are particularly crucial when the receptor site is a metalloprotein such as thermolysin or a MMP. If the metal center and its directly coordinated protein residues have a net charge, Glide assigns a special stability to ligands in which anions coordinate to the metal center. To benefit from this assignment, groups such as carboxylates, hydroxamates, and thiolates must be anionic. If there is no net charge, Glide gives no preference to anions over neutral functional groups.

6. They must be supplied in Maestro, SD, Mol2, or PDB format.

   Maestro transparently converts SD, MacroModel, Mol2, PDB, and other formats to Maestro format during structure import. However, Glide has no direct support for other formats, so you should ensure that your structures are in Maestro, SD, Mol2, or PDB format before starting Glide jobs.

All of the above conditions can be met by using LigPrep to prepare the structures. Use of LigPrep is described in the next section.

3.4.1 Using LigPrep for Ligand Preparation

The Schrödinger ligand preparation product LigPrep is designed to prepare high quality, all-atom 3D structures for large numbers of drug-like molecules, starting with 2D or 3D structures in SD, Maestro, or SMILES format. LigPrep can be run from Maestro or from the command line. For detailed information on LigPrep, see the *LigPrep User Manual*.

To run LigPrep, you must have a LigPrep license. The MacroModel commands `premin` and `bmin` require LigPrep licenses when run in a LigPrep context, and are limited to a restricted set of commands when run using a LigPrep license. LigPrep can be run from Maestro or from the command line.

The LigPrep process consists of a series of steps that perform conversions, apply corrections to the structures, generate variations on the structures, eliminate unwanted structures, and optimize the structures. Many of the steps are optional, and are controlled by selecting options in the LigPrep panel or specifying command-line options. The steps are listed below.

EX2037.0028

1. Convert structure format.
2. Select structures.
3. Add hydrogen atoms.
4. Remove unwanted molecules.
5. Neutralize charged groups.
6. Generate ionization states.
7. Generate tautomers.
8. Filter structures.
9. Generate alternative chiralities.
10. Generate low-energy ring conformations.
11. Remove problematic structures.
12. Optimize the geometries.
13. Convert output file.

The LigPrep panel allows you to set up LigPrep jobs in Maestro. Choose LigPrep from the Applications menu to open the panel. For details of panel options and operation, see Chapter 2 of the *LigPrep User Manual*.

The simplest use of LigPrep produces a single low-energy 3D structure with correct chiralities for each successfully processed input structure. LigPrep can also produce a number of structures from each input structure with various ionization states, tautomers, stereochemistries, and ring conformations, and eliminate molecules using various criteria including molecular weight or specified numbers and types of functional groups present.

The default options in the LigPrep panel remove unwanted molecules, add hydrogens, and minimize the ligand structure (performing a 2D-3D conversion, if necessary). Below are notes on panel options that produce more than one output structure per input structure.

The Ionization options allow you to generate all the ligand protonation states that would be found in the specified pH range. The Ionization options are:

- Retain original state—Retain the original ionization state rather than attempting to ionize the ligand.
- Neutralize—Neutralize the ligand by adding or removing protons from ionizable groups.
- Generate possible states at target pH *target* +/- *range*—This is the default, and can generate several different output structures for each input structure. The default pH *target* is 7.0

EX2037.0029

with a +/- *range* of 2.0, so the default pH range is 5.0 – 9.0. Both the target and range settings can be changed. You can use either the `ionizer` or Epik to generate ionization states. Epik generally yields more accurate states, because it uses a more sophisticated algorithm and performs ionization and tautomerization together. In addition, Epik generates a set of states that are more appropriate for metal binding when you select Add metal binding states, and calculates penalties for these states to use in the docking score. Epik is a separate product, so you must purchase this product to use it.

Generate low energy ring conformations: *number* per ligand. The default is to generate only the lowest energy conformation.

Desalt is selected by default.

Generate tautomers is selected by default. The `tautomerizer` generates up to 8 tautomers per ligand, selecting the most likely tautomers if more than 8 are possible. If you are sure that the input structures are already in the correct tautomeric form for docking to a particular target, then the `tautomerizer` should be turned off by deselecting Generate tautomers.

The `stereoizer` can generate two stereoisomers per chiral center in the ligand, up to a specified maximum. There are three Stereoisomers options:

The first two options, Retain specified chiralities (the default) and Determine chiralities from 3D structure, generate both isomers only at chiral centers where chirality is unspecified or indeterminate; centers with known chirality retain that chirality. The difference is that Retain specified chiralities takes its chirality data from the input file (SD or Maestro), while Determine chiralities from 3D structure ignores input file chiralities and takes chirality information from the 3D geometry.

Generate all combinations varies the stereochemistry up to a maximum number of structures specified by Generate at most *max* per ligand. The default maximum is 32.

3.4.2 Using Other Programs for Ligand Preparation

If you prefer to prepare the ligands with other programs, you can do so. Schrödinger software installations include a number of utilities that can be used to perform some of the above tasks. These utilities are also used by LigPrep. One of these, the Ionizer, can be used to prepare ligands in the required protonation states. Some of the other tasks can be performed as follows:

- Hydrogen atoms can be added in Maestro with either the Add hydrogens toolbar button:

  or the Hydrogen Treatment panel (choose Hydrogen Treatment from the Edit menu).

EX2037.0030

Hydrogen atoms can also be added (or removed) using the utility `applyhtreat`, which is described in Appendix D of the *Maestro User Manual*.

- Structure file format conversion can be done from the command line with utilities such as `structconvert`, `pdbconvert`, `sdconvert`, and `maemmod`—see Appendix D of the *Maestro User Manual*.

EX2037.0031

EX2037.0032

Receptor Grid Generation

Glide searches for favorable interactions between one or more ligand molecules and a receptor molecule, usually a protein. The shape and properties of the receptor are represented on a grid by several different sets of fields that provide progressively more accurate scoring of the ligand poses. For receptors that adopt more than one conformation on binding, you might want to prepare grids for each conformation, to ensure that possible actives are not missed. Glide can, however, handle different hydroxyl conformations with a single grid generation.

The options in each tab of the Receptor Grid Generation panel allow you to define the receptor structure by excluding any cocrystallized ligand that may be present, determine the position and size of the active site as it will be represented by receptor grids, and set up Glide constraints. Ligand docking jobs cannot be performed until the receptor grids have been generated.

Receptor grid generation requires a "prepared" structure: an all-atom structure with appropriate bond orders and formal charges. Information on structure preparation is given in Chapter 3.

This chapter contains the following information:

- Instructions for using the Receptor Grid Generation panel in preparation for Glide ligand docking, including the setup of hydrophobic constraints.
- A detailed description of the Receptor Grid Generation panel and each of its tabs.

Much of the information in this chapter is also available in the Maestro online help.

4.1 The Receptor Grid Generation Panel

To open the Receptor Grid Generation panel, choose Receptor Grid Generation from the Glide submenu of the Applications menu. The Receptor Grid Generation panel has three tabs, which you use to specify settings for the receptor grid generation job:

- Receptor
- Site
- Constraints

These tabs are described in later sections of this chapter.

EX2037.0033

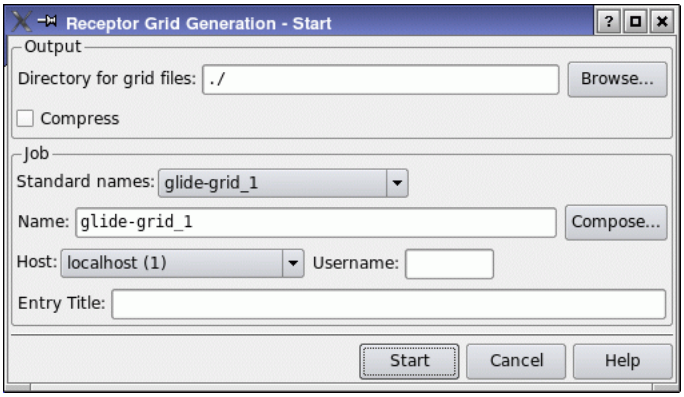


Figure 4.1. The Start dialog box for grid generation jobs.

Below the tabs are three buttons:

Start

When you click the Start button, the Receptor Grid Generation - Start dialog box opens. In this dialog box you can specify where you want the output to be saved, by typing the path into the Directory for grid files text box, or browsing for the directory. You can also choose to compress the grid files into a `.zip` file. You can name your job by typing the name into the Name text box, by choosing from the Standard Names option menu, or by clicking Compose and choosing and editing a job name or a file name. You can specify the Username and the Host, and provide an entry title for the receptor. To start the job, click Start.

Write

Click Write to write the input files without starting the job. A dialog box opens, in which you can specify the job name by typing the name into the Name text box or by choosing from the Standard Names option menu; select the Compress option to compress the grid files into a `.zip` archive at the end of the grid generation job; and choose to write a Glide simplified input file (`.in`) or an Impact (DICE) input file (`.inp`).

Reset

Click Reset to restore the default settings in all tabs.

4.2 The Receptor Tab

In this tab you define the part of the Workspace system for which receptor grids should be calculated. You can also scale receptor atom van der Waals radii in this tab, and choose whether to use partial charges from the force field or from the input structure. The tab has three sections, Define receptor, Van der Waals radii scaling, and Per-atom Van der Waals radius and charge scaling.

4.2.1 Defining the Receptor

The Define receptor section contains options for defining the part of the system in the Workspace to be treated as the receptor. If only the receptor is included in the Workspace, and no ligand is present, you can ignore the Define receptor options.

- If the Workspace includes both a receptor and a ligand, use these options to pick the ligand molecule. The ligand will be excluded from receptor grid generation. Everything not defined as the ligand will be treated as part of the receptor. The ligand can be identified either as a molecule or as an entry in the Workspace.

If you want to use a binding site from a SiteMap calculation, you can do so by treating it as a ligand and including it in the Workspace. Each site from SiteMap is a separate entry. However, the site points in the site are not connected, so the site must be selected as an entry.

To select the ligand, ensure that Pick to identify ligand molecule is selected, choose an option from the option menu, then pick an atom in the ligand molecule or SiteMap site. The ligand (or site) is now distinguished from the receptor. If Show markers is selected, when the ligand molecule is picked, it is marked with dark green markers. Deselect the option to remove the markers.

4.2.2 Van der Waals Radius and Charge Scaling

Glide does not allow for receptor flexibility in docking (apart from hydroxyl rotations), but scaling of van der Waals radii of nonpolar atoms, which decreases penalties for close contacts, can be used to model a slight “give” in the receptor and the ligand. Scaling of other interactions can also help to model flexibility of parts of the receptor.

If you have a receptor in which there is substantial movement upon docking, such as a change in side-chain conformation, backbone location or loop conformation, you should consider docking to multiple protein conformations. You can also use the Induced Fit Docking protocol to account for receptor flexibility. This protocol uses Glide and Prime, and is much more computationally demanding than Glide docking alone. It is therefore mainly useful for docking a small number of ligands. See the document *Induced Fit Docking* for more information.

EX2037.0035

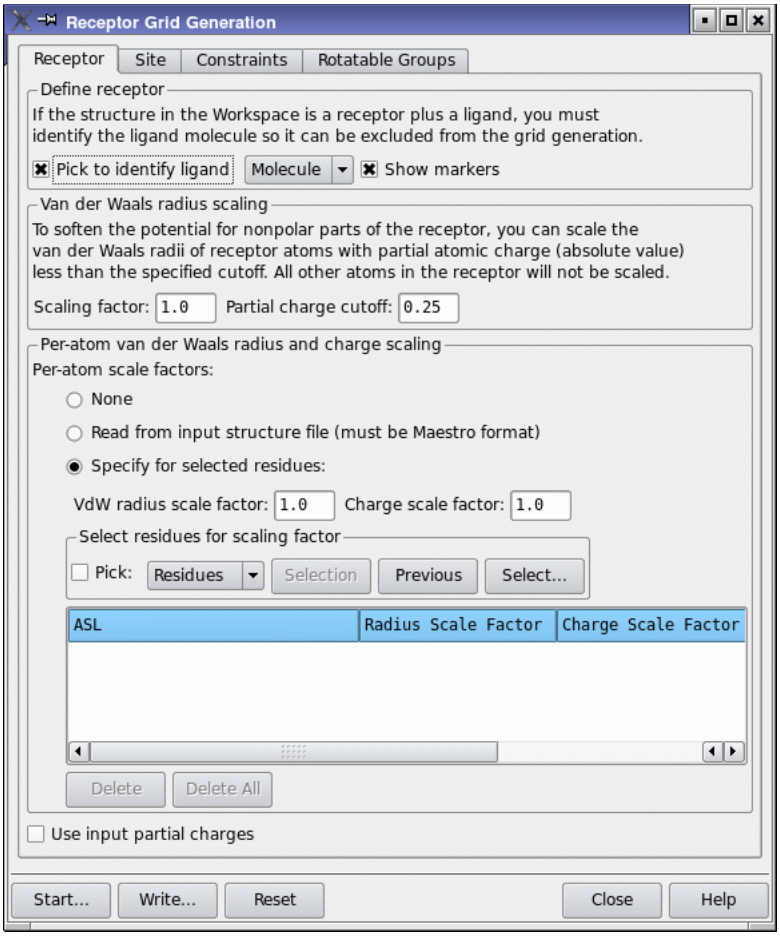


Figure 4.2. The Receptor tab of the Receptor Grid Generation panel.

If you have a receptor that has Ser, Thr, or Tyr residues in the active site and the rotation of the hydroxyls on these groups is important, you can specify the hydroxyls as rotatable groups in the Rotatable Groups tab—see Section 4.5 on page 42 for details.

Glide provides two means of accounting for protein flexibility by scaling of van der Waals radii.

For nonpolar receptor atoms, you can use the controls in the Van der Waals radius scaling section to scale the van der Waals radius of those receptor atoms. The definition of nonpolar atoms is determined by a partial charge threshold that you can set. For ordinary Glide docking, it is recommended that receptor radii be left unchanged, and any scaling be carried out on

EX2037.0036

ligand atoms. Receptor scaling is probably most useful when the active site is tight and encapsulated.

The Scale by text box specifies the scaling factor. Van der Waals radii of nonpolar receptor atoms are multiplied by this value. The default value is 1.00, for which no scaling is done. Scaling of van der Waals radii is performed only on nonpolar atoms, defined as those for which the absolute value of the partial atomic charge is less than or equal to the number in the text box. Since this is an absolute value, the number entered must be positive. The default is 0.25.

For a more flexible method of softening the receptor potential, you can specify van der Waals radii and charges on a per-atom basis. This is done in the Per-atom van der Waals radius and charge scaling section. If you specify per-atom scaling, the global scaling of nonpolar atoms is only performed on those atoms for which per-atom scaling factors are not specified.

There are three choices for the source of per-atom scale factors:

- None—do not use per-atom scale factors.
- Read from input structure file—Read the radius and charge scaling factors from the input file for the receptor. This file must be in Maestro format, and must have the properties VdW radius scale factor (`r_glide_atom_vdwscale`) and Charge scale factor (`r_glide_atom_coulscale`) defined. These properties must have a value for each atom in the receptor. They are defined if you have previously set per-atom scaling factors.
- Specify for selected residues—Specify the scaling factors for selected residues. You can make multiple selections and apply scaling factors to each, using the Select residues for scaling factor picking tools. The selections should be mutually exclusive. A row is added to the table for each pick, with the scale factors specified in the van der Waals radius scale factor text box and Charge scale factor text box. You can subsequently edit the factors in the table.

4.2.3 Selection of Partial Charges

You can use partial charges from the input structure, rather than from the force field, by selecting Use input partial charges. This option is useful if, for example, you have obtained improved partial charges around the active site, such as those from a QSite calculation or a QM-Polarized Ligand Docking calculation.

EX2037.0037

4.3 The Site Tab

The settings in the Site tab determine where the scoring grids are positioned and how they are prepared from the structure in the Workspace. To make these settings, you need to understand how Glide sets up grids.

Glide uses two “boxes” to organize the calculation:

- The grids themselves are calculated within the space defined by the *enclosing box*, or *outer box*. This is also the box within which all the ligand atoms must be contained.
- Acceptable positions for the ligand center during the site point search lie within the *ligand diameter midpoint box*, or *inner box*. This box gives a truer measure of the effective size of the search space. However, ligands can move outside this box during grid minimization.

The only requirement on the enclosing box is that it be big enough to contain all ligand atoms when the ligand center is placed at an edge or vertex of the inner box. Enclosing boxes that are larger than this are not useful: they take up more space on disk and in memory for the scoring grids, which take longer to compute. The maximum size of the enclosing box is 50 Å.

The ligand center is defined in a rigid-docking run as the midpoint of the line drawn between the two most widely separated atoms. The definition changes slightly for flexible docking, where the ligand center becomes the midpoint between the two most widely separated atoms of the *core region*—the part of the ligand remaining after each of the end-groups has been stripped off at the terminal end of the connecting rotatable bond.

The two boxes share a common center. Thus, the operations in the tab that center one box also center the other.

Each rigidly docked ligand or flexibly docked conformation has an associated length, L, which can be defined as twice the distance from the ligand center to its farthest atom. The required relationship between L and the lengths E and B of the enclosing and inner boxes for successful placement of the ligand center anywhere within the inner box is:

$$E \geq B + L$$

The enclosing box must be large enough in each dimension to hold the length of the inner box plus the maximum length of any ligand. If a larger ligand is encountered, not all positions for the center of the ligand in the inner box are accessible. The effective inner box for that ligand will be smaller than the dimension nominally specified. In any docking job using these receptor grids, ligands are confined to the enclosing box.

EX2037.0038

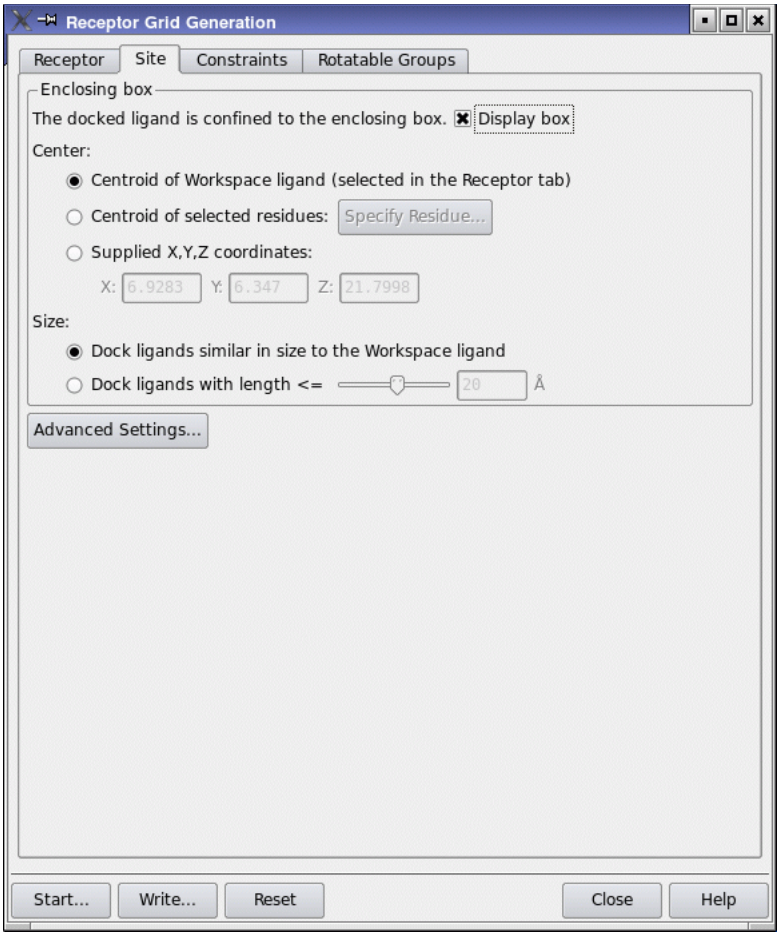


Figure 4.3. The Site tab of the Receptor Grid Generation panel.

If the structure in the Workspace consists of a receptor and the ligand molecule you identified in the Receptor tab, Glide uses the position and size of the ligand to calculate a default center and a default size for the enclosing box. When you open the Site tab, the Workspace displays the center of the enclosing box as a set of coordinate axes colored bright green, and the boundaries of the region as a purple wire-frame cube.

If the Workspace includes a SiteMap binding site, you might want to reduce the size of the enclosing box, because it is likely that the site is larger than defined by a ligand. You might also want to specify the center of the box by selection of a few residues from the receptor, as the centroid of the site might not be in the optimal location. This is particularly so if the site is not well defined or the site points extend over a broad region.

EX2037.0039

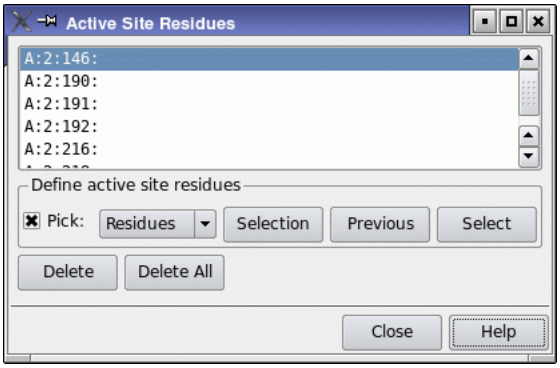


Figure 4.4. The Active Site Residues panel.

If the Workspace structure consists of a receptor only, there is no default center for the enclosing box. The box will not be displayed until you have specified a grid center by selecting residues.

By default, the purple enclosing box outline and the green axes at the center are displayed when you enter the tab. Deselect Display Box to undisplay the box and its center.

4.3.1 Selecting a Box Center

Select one of the options under Center to determine how the center of the grid is defined:

- Centroid of Workspace ligand

  This option centers grids at the centroid of the ligand molecule that was defined in the Receptor tab, also called the Workspace ligand. If a Workspace ligand has been defined, this option is the default. The Advanced Settings button is available with this option.

- Centroid of selected residues

  This option centers grids at the centroid of a set of residues that you select. With this option you can define the active site (where grids should be centered) with only the receptor in the Workspace. The Specify Residue button is only available when you choose this option; the Advanced Settings button is not available with this option.

  To select the residues, click Specify Residue. The Active Site Residues panel opens. Using the picking controls, you can pick the residues that best define the active site. Picked residues are marked in pink when the Active Site Residues panel is open. The center and the default boundaries of the enclosing box are calculated and displayed after each residue pick. The list of selected residues is displayed at the top of the panel. You can delete residues by selecting them in the list, and clicking Delete. To delete all residues, click Delete All.

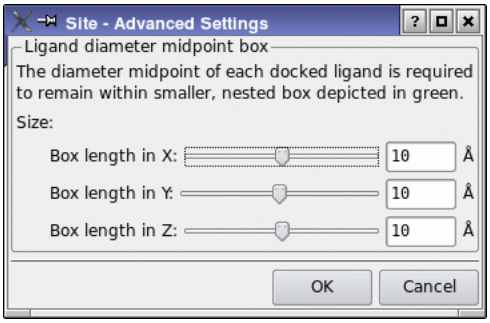


Figure 4.5. The Site - Advanced Settings dialog box.

- Supplied X, Y, Z coordinates

  This option centers the grid at the Cartesian coordinates that you specify in the X, Y, and Z text boxes. These text boxes are only available when you choose this option. The Advanced Settings button is available with this option.

4.3.2 Setting the Box Sizes

The Size section provides options for the size of the enclosing box. The default option is Dock ligands similar in size to the Workspace ligand when a ligand is used for the grid center. This choice is suitable when the ligands to be docked are of the same size as, or smaller than, the Workspace ligand. If you expect to dock larger ligands, or if there is no Workspace ligand, select Dock ligands with length <= and use the slider to choose an appropriate maximum ligand length. The slider is set to 20 Å by default. If you use the centroid of specified residues, or supplied coordinates, Dock ligands with length <= is the default option.

To change the size of the inner box, click Advanced Settings. The Site - Advanced Settings dialog box opens, and the inner box, or *ligand diameter midpoint box,* is displayed as a cube outlined in bright green. The diameter midpoint of each docked ligand remains within this box in the site-point search stage, but can move outside this box in the grid minimization stage (see Figure 2.2 on page 7). You can use the Size sliders to increase or decrease the dimensions of each side of the box. The default is 10 Å on each side; the allowed range is 6 Å to 14 Å.

A larger ligand diameter midpoint box can be useful to allow ligands to find unusual or asymmetric binding modes in the active site. Conversely, if the default ligand diameter midpoint box allows ligands to stray into regions you know to be unfruitful, you can confine their midpoints to a smaller box, eliminating some of the less useful poses and saving calculation time. Changing the shape of the box can be useful when the active site is spatially extended in one or more directions.

4.4 The Constraints Tab

The Constraints tab of the Receptor Grid Generation panel is used to define Glide constraints for the receptor grids to be generated. Glide constraints are receptor-ligand interactions that you believe to be important to the binding mode, based on structural or biochemical data. Setting constraints enables Glide to screen out ligands, conformations, or poses that do not meet these criteria early on in their evaluation for docking suitability.

There are four types of Glide constraints available in Glide 5.5: positional constraints, H-bond constraints, metal constraints, and hydrophobic constraints.

- A positional constraint is a requirement that one or more ligand atoms occupy a spherical volume that is centered at a particular position.
- An H-bond constraint is a requirement that a particular receptor-ligand hydrogen bond be formed.
- A metal constraint is a requirement that a particular metal-ligand interaction is present when the ligand is docked.
- A hydrophobic constraint is a requirement that a user-defined number of hydrophobic heavy atoms in the ligand occupy a hydrophobic region in the active site.

The constraints tab has three subtabs, Positional, H-bond/Metal and Hydrophobic, for setting up the constraints. The number of constraints of each type is reported on the subtab tab, and the total number of constraints is reported at the top of the tab. The maximum number of constraints that you can define for a given grid is ten constraints, distributed among positional, H-bond, metal, and hydrophobic constraints.

When constraints setup is complete and the grid generation job is run, Glide writes a file containing the information about the constraints. Subsequent docking jobs use this file to determine whether a given ligand pose satisfies the constraints. If the base name for writing grid files is *gridbase*, the constraints file is named *gridbase*`.cons`. Not all of these constraints are used in a given docking job: when you set up the docking job, you can choose which constraints to apply in ligand docking.

To use Glide constraints, you must define the constraints during grid generation setup and then apply the constraints during docking setup. Any Glide constraints that you want to apply in docking must be defined when the receptor grids are generated.

Note: In this release, you cannot use flexible receptor groups for constraints.

When you are setting up constraints, it may be helpful to undisplay most of the receptor, leaving only residues within a short distance of the ligand visible.

EX2037.0042

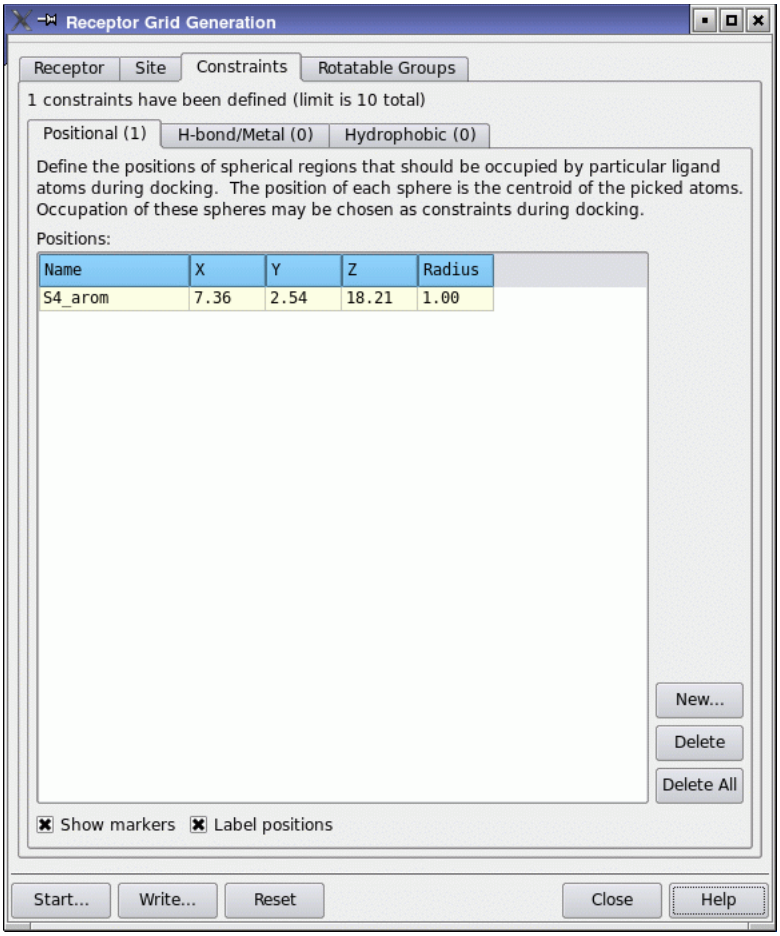


Figure 4.6. The Constraints tab of the Receptor Grid Generation panel showing the Positional subtab.

4.4.1 Setting Positional Constraints

Positional constraints define a region that must contain a particular kind of ligand atom. The specific kind of atom is defined during docking setup, using SMARTS patterns. Positional constraints allow you to require interactions between any kind of receptor and ligand atoms, while at the same time placing tighter restrictions on the ligand atom position than is typical with other constraint types.

EX2037.0043

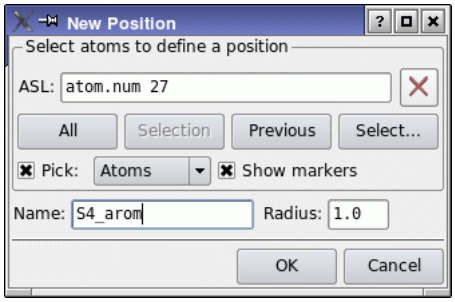


Figure 4.7. The New Position dialog box.

For example, a hydrogen-bond acceptor in the receptor might be capable of forming hydrogen bonds in two directions, but only one of these results in good binding. While setting an H-bond constraint in this case allows a ligand hydrogen atom to lie in either of these directions, a positional constraint can require it to be in the "good" direction. The constraint could be set by selecting a hydrogen atom that bonds in the "good" direction to define the position of the constraint. For reasons such as this, we suggest that you display a model ligand in the Workspace to aid in selecting appropriate positions for constraints.

To add a positional constraint, click New. This button opens the New Position dialog box, in which you can pick atoms with the standard picking controls to define a position, name the positional constraint, and specify its radius. The position is the centroid of the selected atoms, and must be inside the enclosing box. While picking is in progress, the constraint is marked with a gray sphere. When you click OK, the constraint is added to the Positions table if it is inside the enclosing box; otherwise a warning is displayed.

To delete a constraint, select it in the Positions table and click Delete; to delete all constraints, click Delete all.

The Positions table displays the positional constraints you have chosen, giving the name, coordinates, and radius of the constraint sphere for each constraint. The coordinates and the radius are given in angstroms. You can select a single constraint in the table, and edit the name, coordinates, and radius of the sphere by clicking in the table cell and changing the value.

To view the positional constraints in the Workspace, select Show markers. The selected constraint is marked by a yellow sphere. The other positional constraints are marked by red spheres. If Show markers is selected, selecting the Label positions option displays the name of the constraint in the Workspace. The labels are colored the same as the constraints.

4.4.2 Setting H-Bond and Metal Constraints

Up to ten symmetry-distinct receptor atoms can be chosen as possible H-bond or metal constraint sites.

For hydrogen-bonding interactions, the receptor atom must be a polar hydrogen (including thiol H in cysteine), nitrogen, or oxygen. If you choose an atom with one or more symmetry-equivalent atoms in its functional group, the symmetry-equivalent atoms are all selected as well, and collectively count as one constraint. For example, if you create a constraint by picking one oxygen atom of a carboxylate group, Glide includes the other oxygen atom in the same constraint. A ligand interaction with either oxygen atom satisfies that single constraint.

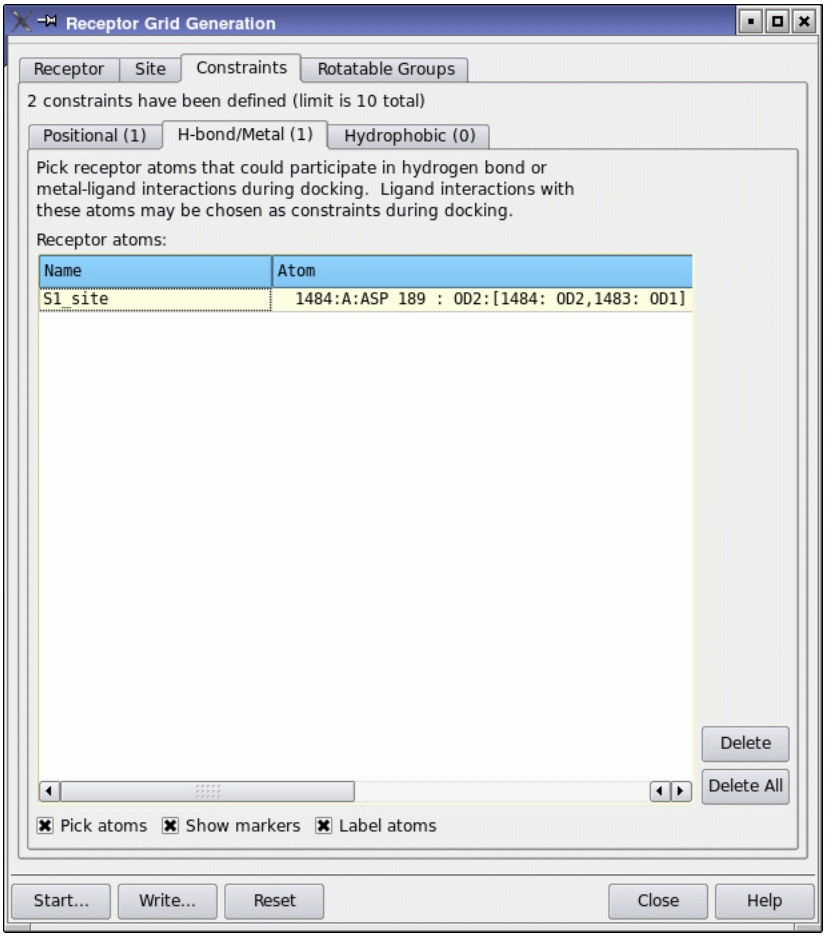


Figure 4.8. The Constraints tab of the Receptor Grid Generation panel showing the H-bond/Metal subtab.

EX2037.0045

For metal-ligand interaction constraints, the receptor atom must be a metal ion.

The receptor atoms selected must also be close enough to the ligand that satisfying the constraints is possible. You do not need to specify limits on distances or angles between receptor and ligand atoms for the constraint: Glide sets these values internally, to H-acceptor distances of 1.2 to 2.5 Å; donor angles greater than 90°, and acceptor angles greater than 60°.

These values are looser than those employed by Maestro. General distance requirements are incorporated using the enclosing box for the ligand. The receptor atoms selected for constraints must be inside the enclosing box (which is displayed in purple) or within bonding range of it.

To display hydrogen bonds in the Workspace, choose Inter from the Display H-bonds button menu and click on a ligand atom. The hydrogen bonds between the ligand and the receptor are displayed. This should make it easier to locate the relevant receptor atoms. If you want to use the looser criteria for hydrogen bonds employed by Glide, select Measurements from the Tools menu and change the values in the H-bonds tab.

To set H-bond or metal constraints, ensure that Pick atoms is selected in the H-bond/Metal constraints tab, and pick the desired atoms in the Workspace. Glide automatically includes symmetry-equivalent atoms such as the other oxygen in a carboxylate group. If Show markers is selected, a red cross and red padlock appear next to each atom picked. If the picked atom is one of a set of symmetry-equivalent atoms, all the atoms in the set are marked. If Show markers is selected, selecting Label atoms displays the constraint name in the Workspace.

As you select atoms in the receptor, they appear in the Receptor atoms table. The Atom column gives detailed information on the atom selected in the format

atom-number : *chain* : *residue-name* *residue-number* : *atom-name* : *symmetry-set*

where

- *atom-number* is the Maestro atom number
- *chain* is the chain name
- *residue-name* is the name of the residue
- *residue-number* is the residue number and the insertion code, if any
- *atom-name* is the PDB atom name
- *symmetry-set* is the atom name or symmetry-equivalent atom set

for example,

341:C:ASN 239 : OD1 : ODn

If the picked atom is part of a symmetry equivalent set, its identification is followed by square brackets enclosing the number and name of each atom in the set, separated by commas.

EX2037.0046

The Name column gives a label that is used when displaying the possible constraints from the Ligand Docking panel. The format of the name is:

chain : *residue-name* : *residue-number* : *atom-name* (*type*)

where the quantities except type are defined above, and *type* is either `hbond` or `metal`, for example,

C:ASN:239:OD1(hbond)

This name can be changed by editing the table cell.

To delete a single H-bond or metal constraint, select it in the list and click Delete; to delete all the listed constraints, click Delete All.

4.4.3 Setting Hydrophobic Constraints

A hydrophobic constraint requires that a hydrophobic region of the receptor be occupied by one or more hydrophobic heavy atoms in the ligand. The possible hydrophobic regions are identified from a hydrophobic map of the receptor site. You can select one or more of the hydrophobic regions as a constraint, and determine the size of the region that must be occupied by adding or deleting cubic volumes ("cells") to the region. When you set up a docking job, you can select one or more of these regions and specify how many atoms must occupy each region. The tools for generating the hydrophobic map and defining the hydrophobic constraint regions are in the Hydrophobic subtab of the Constraints tab.

4.4.3.1 Generating the Hydrophobic Map

To set up a hydrophobic constraint, you must first run a job to generate the hydrophobic map of the receptor site. To generate the hydrophobic map, click Locate Hydrophobic Cells in the Setup section. A map generation job is started, which takes a few minutes. The icon on the job status button turns green and spins; when the job finishes it turns pink again.

Translucent gray cubes that represent the hydrophobic regions around the active site should be displayed when the job finishes. If they are not, adjust the Threshold phobic potential slider until they appear. A less negative threshold enlarges the hydrophobic region; a more negative threshold diminishes the hydrophobic region. The threshold corresponds to the isovalue contour at which the hydrophobic map is displayed, and has a default value of –0.5.

4.4.3.2 Defining Hydrophobic Constraint Regions

After the hydrophobic map has been generated, the table in the Define regions section contains a single default constraint region, `region1`, with 0 in the Num. Cells column (equivalent to no constraint). You can change the name of a region in the Name column, or turn off the visualization markers for a region by deselecting the box in the V column.

EX2037.0047

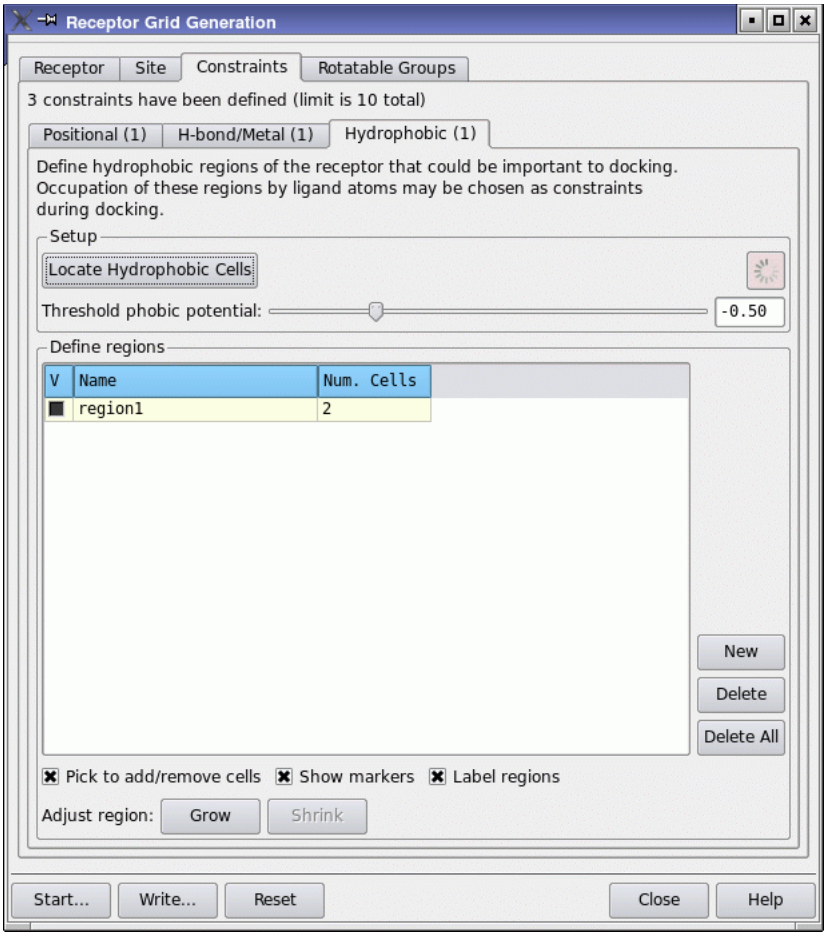


Figure 4.9. The Constraints tab of the Receptor Grid Generation panel showing the Hydrophobic subtab.

A region is defined by a set of hydrophobic cells. The cells in the region do not have to be contiguous. To add individual cells to a region, you can select Pick to add/remove cells and click on cells in the Workspace. When you select this option, the cursor has the label C, to indicate that cell picking is active. The cell color changes to red when it is added. The last picked cell is outlined in yellow. To delete a cell that has already been added, click on the cell in the Workspace. Its color changes to gray. (See Figure 4.10.)

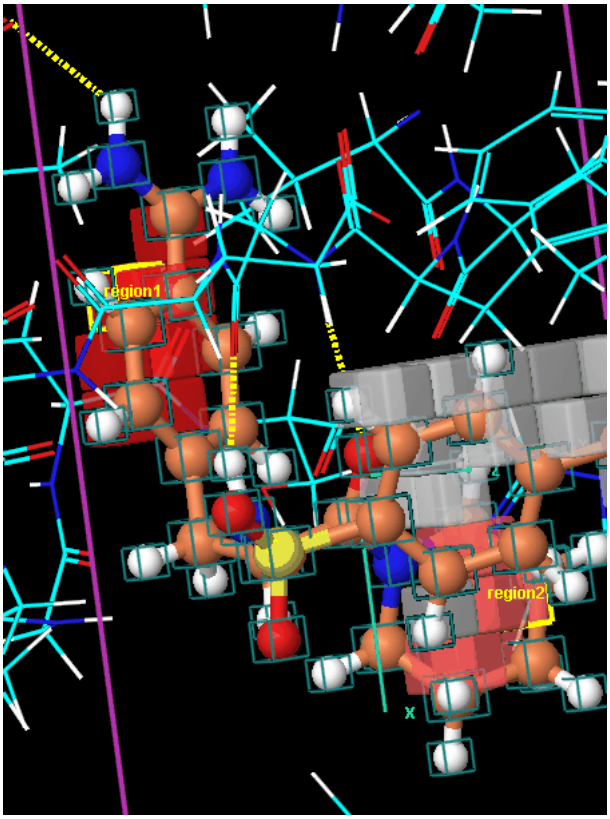


Figure 4.10. Hydrophobic constraint regions displayed in the Workspace.

You can also add cells and remove cells using the Grow and Shrink buttons. When you click Grow, the cells that are nearest neighbors to the cell outlined in yellow are added to the defined region (colored red). Clicking Grow a second time selects a layer of cells adjacent to those most recently selected. This process can be repeated until the entire block of cells is selected. To remove layers of cells added using Grow, click Shrink. Each click removes one layer.

To add a new hydrophobic constraint region, click New. A row is added to the table with a default name of `region`*n*. To delete a constraint region, click Delete.

When the Label regions option is selected (the default), regions are labeled in the Workspace with their Name from the Define regions table.

EX2037.0049

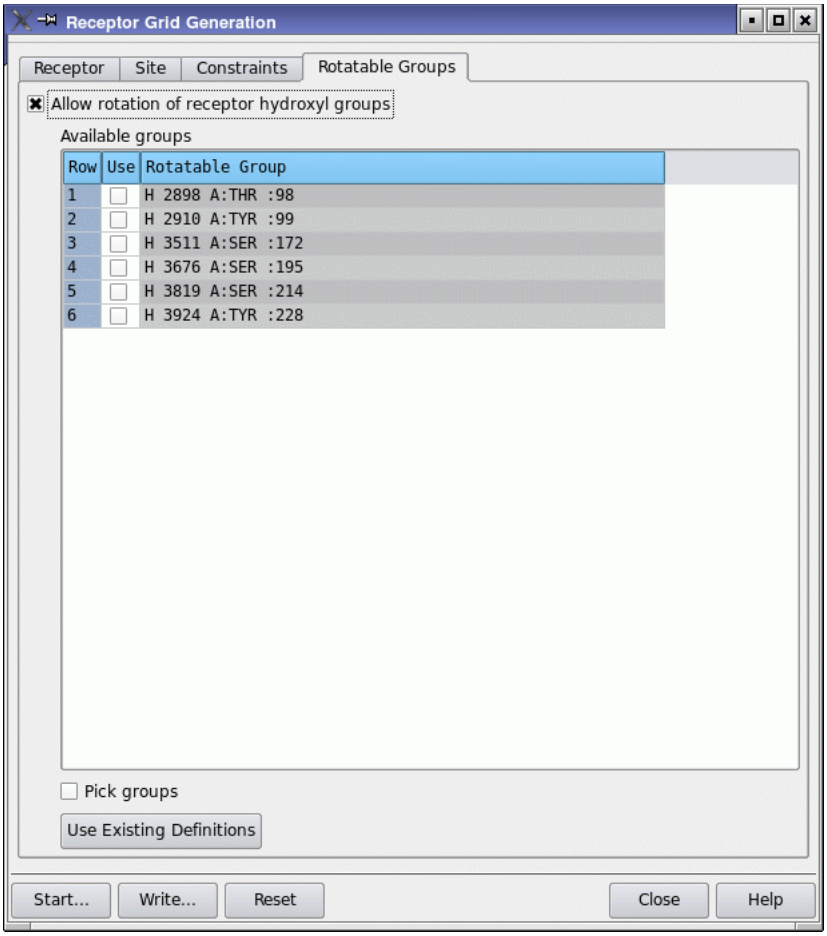


Figure 4.11. The Rotatable Groups tab of the Receptor Grid Generation panel.

4.5 The Rotatable Groups Tab

The hydroxyl groups in residues such as Ser, Thr, and Tyr can adopt different orientations with different ligands. Glide can allow such groups to adopt different orientations when ligands are docked, to produce the most favorable interaction. For Ser and Thr, the hydroxyls can be oriented in any of the three local minima; for Tyr, they can be in either of the two local minima. The Rotatable Groups tab allows you to choose which of these hydroxyl groups should be treated flexibly.

To enable the generation of grids that include flexible hydroxyls, select Allow rotation of receptor hydroxyl groups. When you do so, the Available groups table is loaded with a list of all

rotatable groups within the enclosing box. You can then select which of these groups you want to include as rotatable groups either by selecting them in the Use column of the table, or selecting Pick groups and picking the hydrogens of the groups in the Workspace. The selected groups are marked in the Workspace. Clicking a second time, either in the column or in the Workspace, deselects the group.

If you have previously prepared grids with flexible hydroxyls, you can click Use Existing Definitions to load the existing selection.

In this release, you cannot create constraints to flexible hydroxyls.

Note: Once you have set up a grid with flexible hydroxyls, the flexibility is used in docking, and cannot be turned off.

EX2037.0051

EX2037.0052

Ligand Docking

Glide ligand docking jobs require a set of previously calculated receptor grids and one or more ligand structures. The ligand structures must satisfy the conditions listed on page 19. Information on setting up grid generation jobs is given in Chapter 4. For a docking tutorial example, see Chapter 3 of the *Glide Quick Start Guide*.

Preparation of the ligands before docking is strongly recommended. LigPrep or MacroModel can be used to prepare ligands—see Chapter 3 for more information.

If a correct Lewis structure cannot be generated for a ligand, it is skipped by the docking job. Glide also automatically skips ligands containing unparametrized elements, such as arsenic, or atom types not supported by the OPLS force fields, such as explicit lone pair "atoms."

This chapter contains a detailed description of the Ligand Docking panel in Maestro and each of its tabs, including instructions for applying Glide constraints, using similarity, using extra-precision Glide docking (Glide XP), and distributed processing.

5.1 The Ligand Docking Panel

To open the Ligand Docking panel, choose Ligand Docking from the Glide submenu of the Applications menu. The Ligand Docking panel has six tabs:

- Settings
- Ligands
- Core
- Constraints
- Similarity
- Output

These tabs are described in the following sections of this chapter. You use the options in these tabs to specify settings for the ligand docking job. When you have completed your setup, the following buttons allow you to process your job:

- Start—Open the Ligand Docking - Start dialog box to start the job
- Write—Write the input files to disk without starting the job
- Reset—Discard settings and restore the default settings in all tabs

EX2037.0053

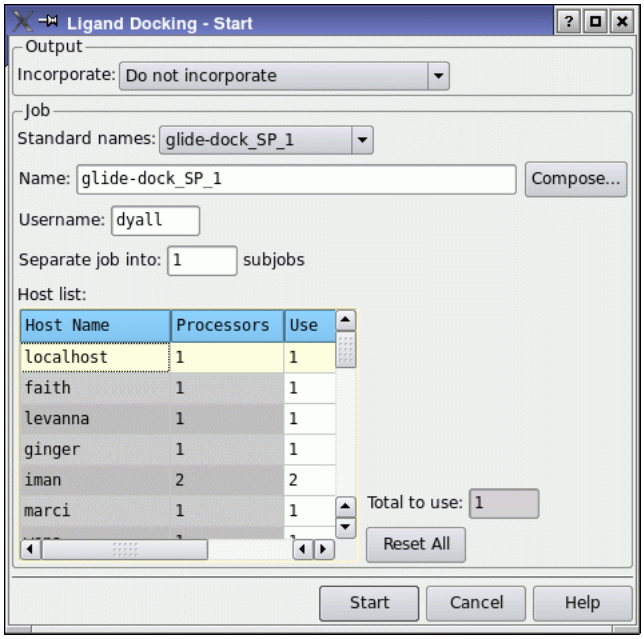


Figure 5.1. The Start dialog box for ligand docking jobs.

The Ligand Docking - Start dialog box provides controls for the running of the job. You can name the job by entering text in the Name text box, by choosing a name from the Standard Names option menu, or by clicking Compose and choosing and editing a job name or a file name. You can specify the user name on the specified hosts. If you want to incorporate the results into the current Maestro project, you can select an option from the Incorporate option menu. Note that incorporation can take a long time if you have a large number of ligands.

You can also specify options for distributed processing. By default, the job is run serially. To separate the docking job into subjobs that are run in parallel, enter the number of subjobs in the Separate docking into *n* subjobs text box, then select the hosts and enter number of processors to use on each host from the Host list table. The number of subjobs cannot be set to less than the number of processors. You can edit the cells in the Use column to set the number of processors to use for each host. For more details on the use of this table, see Section 2.2 of the *Job Control Guide*.

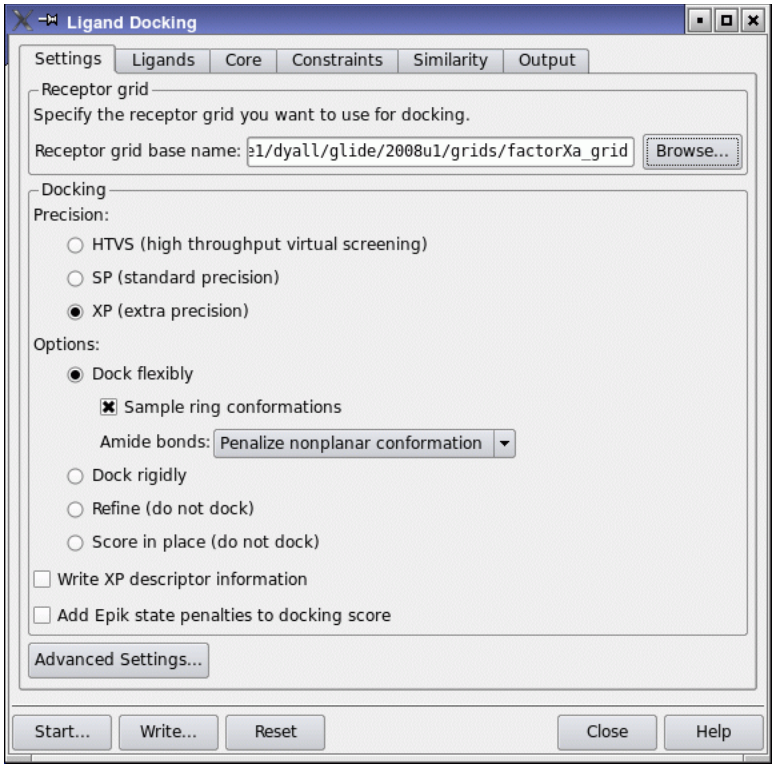


Figure 5.2. The Settings tab of the Ligand Docking panel.

5.2 The Settings Tab

The Settings tab defines the basic options for docking ligands: specifying the grid, selecting the precision, and setting flexibility options. You can also make settings for the selection of initial poses and for the energy minimization of the poses that pass the initial selection in the Settings - Advanced Settings dialog box. To open this dialog box, click Advanced Settings. The settings, both basic and advanced, are described in the following subsections.

The force field used for docking is the OPLS_2001 force field. You cannot change the force field from Maestro. If you want to use a different force field, you must edit the input file.

5.2.1 Specifying the Receptor Grid

To specify the receptor grid for the docking job, click Browse in the Receptor grid section of the Settings tab to open a file selector and choose a grid file (`.grd`) or a compressed grid archive (`.zip`). Note that the grid archive cannot be renamed, otherwise the docking job will

EX2037.0055

fail. The file name, without the extension, is displayed in the Receptor grid base name text box. You can also enter the base name directly into the text box.

5.2.2 Selecting the Docking Precision

There are three choices of docking precision, given under Precision in the Docking section.

- HTVS (high-throughput virtual screening)—High-throughput virtual screening (HTVS) docking is intended for the rapid screening of very large numbers of ligands. HTVS has much more restricted conformational sampling than SP docking, and cannot be used with score-in-place. Advanced settings are not available for HTVS, but are fixed at predetermined values.
- SP (standard precision)—Standard-precision (SP) docking is appropriate for screening ligands of unknown quality in large numbers. Standard precision is the default.
- XP (extra precision). Extra-precision (XP) docking and scoring is a more powerful and discriminating procedure, which takes longer to run than SP. XP is designed to be used on ligand poses that have a high score using SP docking. We recommend that you run your database through SP docking first, then take the top 10% to 30% of your final poses and dock them using XP, so that you perform the more expensive docking simulation on worthwhile poses. For more information on XP docking, see Section 2.3 on page 10.

If you want to dock a set of ligands using a progression of precision, you can use the Virtual Screening Workflow to set up and run the docking jobs. See the *Virtual Screening Workflow* document for details.

5.2.3 Setting Docking Options

Under Options in the Docking section of the Settings tab, you can choose whether ligands are docked flexibly, rigidly, or not at all (score in place), and set options for conformation generation. These options are described below.

Dock flexibly

This is the default option, and directs Glide to generate conformations internally during the docking process; this procedure is known as *flexible docking*. Conformation generation is limited to variation around acyclic torsion bonds and pyramidal nitrogen inversions such as in amines and sulfonamides. In addition, you can generate ring conformations by selecting the option Sample ring conformations (this is the default). Ring conformations are generated with the same technology as in LigPrep—see Section 4.11 of the *LigPrep User Manual* for details. This option is selected by default. Only low-energy conformations are kept.

EX2037.0056

For amide bonds, you can control sampling of the O=C–N–R torsion with the Amide bonds option menu, which has four options:

- Penalize nonplanar conformation—Apply penalties to amide bonds that are not cis or trans, rather than freezing them entirely.
- Vary conformation—Allow the conformation to vary, including nonplanar conformations.
- Retain original conformation—Fix amide bonds in their input conformation.
- Allow trans conformation only—Enforce the trans conformation to within a small angle range (20° by default). Ligands that do not dock in a trans conformation are rejected.

The default is to penalize nonplanar conformations.

Dock rigidly

Rigid docking allows the input conformation of the ligand to be translated and rigidly rotated relative to the receptor binding site, but skips the conformation generation step.

Refine (do not dock)

When this option is selected, Glide does no docking, but rather uses the input ligand coordinates to perform an optimization of the ligand structure in the field of the receptor, and then the ligand is scored. The goal of this docking method is to find the best-scoring pose that is geometrically similar to the input pose. For HTVS and SP, a minimization is performed; for XP, the ligand is regrown in place. This option is not a substitute for a full docking calculation, and requires accurate initial placement of the ligand with respect to the receptor.

Score in place (do not dock)

When this option is selected, Glide does no docking, but rather uses the input ligand coordinates to position the ligands for scoring. It therefore requires accurate initial placement of the ligand with respect to the receptor.

This option is useful to score the native ligand in its cocrystallized or modeled position, or as a post-processing step on Glide-generated poses to obtain individual components of the GlideScore prediction of the binding affinity. It can also be used to check whether the scores of the known binders in their native proteins are similar enough to their scores when cross-docked to the chosen receptor protein. If this is the case, it is reasonable to expect that similar structures would also score well.

Score in place should not be used with Glide XP, as full XP sampling is normally needed to avoid strong XP penalties for ligands that should be able to dock correctly.

Note: You cannot use score in place for the ligand that is defined as the reference ligand for calculation of the RMSD in conformational comparisons.

EX2037.0057

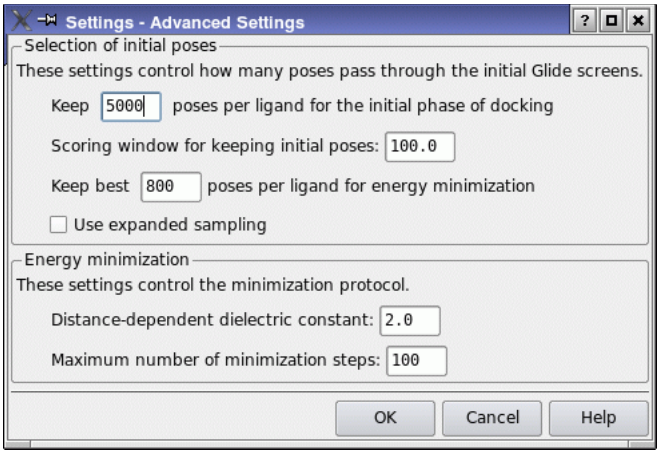


Figure 5.3. The Settings - Advanced Settings dialog box.

5.2.4 Writing XP Descriptor Information

If you want to visualize the various scoring terms from the XP scoring function in the XP Visualizer panel, select Write XP descriptor information. These terms are written to a file that can be read by the XP Visualizer. You can write these terms only for XP docking runs.

Note: To generate XP descriptor information, you must have a license for this feature.

5.2.5 Adding Epik State Penalties to the Docking Score

If the ligands have been prepared using Epik for ionization and tautomerization, the Epik penalties for adopting higher-energy states (including those where metals are present) are added to the docking score when the Add Epik state penalties to docking score option is selected. Ligands that do not have this information are not penalized and will therefore have better scores, so you should ensure that the ligand set is consistent.

If the ligand interacts with a metal (distance less than 3.0 Å), the metal penalties that are computed when Epik is run with the metal binding option are used. If multiple ligand atom-metal interactions are found, the smallest value of the metal-specific penalty is used.

5.2.6 Advanced Options

Further options for controlling the docking process are available in the Settings - Advanced Settings dialog box, which opens when you click Advanced Settings in the Settings tab. You will not normally need to change these options from their defaults.

EX2037.0058

5.2.6.1 Selection of Initial Poses

The options in the Selection of initial poses section of the box control the way poses pass through the filters for the initial geometric and complementarity "fit" between the ligand and receptor molecules. The grids for this stage contain values of a *scoring function* representing how favorable or unfavorable it would be to place ligand atoms of given general types in given elementary cubes of the grid. These cubes have a constant spacing of 1 Å. The "rough score" for a given *pose* (position and orientation) of the ligand relative to the receptor is simply the sum of the appropriate grid scores for each of its atoms. By analogy with energy, favorable scores are negative, and the lower (more negative) the better.

The initial "rough scoring" is done on a coarse grid, on which the possible positions for placing the ligand center are separated by 2 Å, twice the elementary cube spacing, in *x*, *y*, and *z*. The "refinement" step rescores the successful rough-score poses after the particular rigid translational repositioning of –1.0 Å or +1.0 Å in *x*, *y*, and *z* that gives the best possible score. This procedure effectively doubles the resolution of the scoring screen.

Poses that pass these initial screens enter the final stage of the algorithm, which involves evaluation and minimization of a grid approximation to the OPLS-AA nonbonded ligand-receptor interaction energy.

This section contains three text boxes and an option for controlling the initial screening:

Keep *n* initial poses per ligand for the initial phase of docking

This text box sets the maximum number of poses per ligand to pass to the grid refinement calculation. The value must be a positive integer. The default setting depends on the type of docking specified and whether Glide constraints have been applied:

- 5000 poses for flexible docking jobs in general
- 500 poses for flexible docking jobs to which Glide constraints are applied
- 1000 poses for rigid docking jobs (Glide constraints do not change this value)

Scoring window for keeping initial poses

This text box sets the rough-score cutoff for keeping initial poses, relative to the best rough score found so far. The value must be positive. The default window is 100.0 kcal/mol, meaning that to survive, a pose must score within 100.0 kcal/mol of the best pose. Using the default window, for example, if the best pose found so far has a score of -60.0 kcal/mol, poses with a score more positive than 40.0 kcal/mol will be rejected.

Keep best *m* poses per ligand for energy minimization

This text box specifies the number of poses per ligand to be energy minimized on the OPLS-AA nonbonded-interaction grid. The default setting depends on the type of docking specified.

EX2037.0059

- 400 poses for flexible SP docking jobs, 800 poses for flexible XP docking jobs.
- 40 poses for flexible docking jobs to which Glide constraints are applied.
- 100 poses for rigid docking jobs (Glide Constraints do not change this value).

The range for this value is 1 to *n*, where *n* is the value in the Keep *n* initial poses per ligand for the initial phase of docking text box, with the exception that for XP docking, the number can be adjusted between 800 and 1000.

Use expanded sampling

This option bypasses the elimination of poses based on the rough score, so that many more poses are passed on to the subsequent stages. This option is particularly useful for fragment docking to ensure that good poses are found.

5.2.6.2 Energy Minimization Settings

The energy minimization stage of the docking algorithm minimizes the energy of poses that are passed through the Selection of initial poses scoring phase. The Energy minimization section of the Settings - Advanced Settings dialog box contains two options:

Distance-dependent dielectric constant

Glide uses a distance-dependent dielectric model in which the effective dielectric "constant" is the supplied constant multiplied by the distance between the interacting pair of atoms. The default setting is 2.0, and Glide's sampling algorithms are optimized for this value. Although this text box allows you to set the dielectric constant to any real value greater than or equal to 1.0, changing this setting is not recommended. This relatively strong dielectric is sometimes needed to "hold" a hydrogen bond in the energy-grid optimization phase of the algorithm.

Maximum number of conjugate gradient steps

This text box specifies the maximum number of steps taken by the conjugate gradient minimization algorithm. The value must be greater than or equal to 0; the default value is 100. Setting the value to 0 results in a single-point energy calculation on each pose that survives rough-score screening, or on the single initial pose if no screening was done.

5.3 The Ligands Tab

In the Ligands tab you specify the source of ligands to be docked or scored and set size limits for skipping ligands. You can also choose the partial charges to use and change the settings for van der Waals radii scaling of nonpolar ligand atoms.

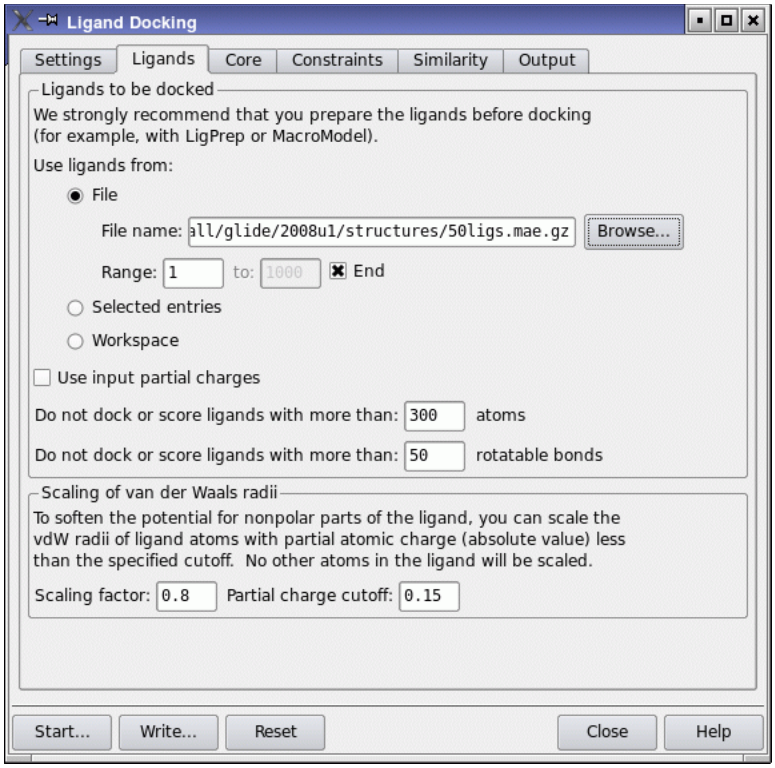


Figure 5.4. The Ligands tab of the Ligand Docking panel.

5.3.1 Specifying the Source of the Ligands

In the Ligands to be docked section you specify the source of ligands to be docked (or scored in place). To specify the source, choose one of the following options under Use ligands from:

File

If you select File as the source of ligands, you can type the file name in the File name text box, or click the Browse button to open a file selector, in which you can choose the format and navigate to the ligand file. The file must be in Maestro, SD, MOL2, or PDB format. If you wish, you can specify a compressed Maestro file (`.mae.gz` or `.maegz`).

By default, all structures in the file are docked. To select a contiguous subset of ligands from the file, enter values in the Range text boxes. To enter the upper limit, you must deselect End. The End option allows you to dock all ligands from a particular ligand to the end of file.

EX2037.0061

The ability to select a subset of ligands is useful if a problem with Glide or with the input ligand causes a Glide job to terminate abnormally. Then you can set the initial ligand number when you restart the job to pick up at the point in the ligand file at which the problem occurred.

Selected entries

Select this option to dock ligands that are selected entries in the Project Table.

Workspace

Select this option to dock the structure in the Workspace. If you select this option, only a single entry is permitted in the Workspace, and this entry is treated as the ligand to be docked. You should therefore ensure that the structure in the Workspace is the ligand to be docked, without the receptor or any other structures.

5.3.2 Selecting the Source of Ligand Partial Charges

Partial charges for the ligand atoms are obtained from the force field. These partial charges do not reflect the influence of the environment on the atomic charges. To compensate, you can use pregenerated partial charges from some other source. To use the partial charges that are supplied with the ligand, select Use input partial charges.

The QM-polarized ligand docking solution provides a way of using quantum mechanical charges from a QSite calculation to dock ligands. For more information, see the document *QM-Polarized Ligand Docking*.

5.3.3 Setting Restrictions on the Type of Ligands

In the text boxes at the end of the Ligands to be docked section, you can set limits on structural feature counts to screen out structures before docking. These text boxes are described below.

Do not dock or score ligands with more than a atoms

This text box sets the maximum number of atoms a ligand structure may have if it is to be docked. Ligand structures in the input file that have more than the specified number of atoms will be skipped. The default is also the maximum, 300 atoms. You can reduce the maximum number of atoms a, if the active-site region is small and enclosed, to speed up a docking calculation on a large ligand database.

Do not dock or score ligands with more than r rotatable bonds

This text box sets the maximum number of rotatable bonds a ligand structure may have if it is to be docked. Ligand structures in the input file that have more than this number of rotatable

EX2037.0062

bonds will be skipped. The default is also the maximum, 50 rotatable bonds. If only relatively small or rigid ligand "hits" are wanted, you can decrease the value of *r*.

5.3.4 Van der Waals Radii Scaling

Glide does not generally allow for flexible receptor docking, except for reorientation of selected hydroxyl groups (see Section 4.5 on page 42).[1] However, successful docking sometimes requires that the ligand and the receptor "give" a bit in order to bind. To model this behavior, Glide can scale the van der Waals radii of nonpolar atoms (where nonpolar is defined by a partial charge threshold you can set), thereby decreasing penalties for close contacts. Scaling is performed for qualifying atoms in the ligand, but not those in the receptor. Ligand atom radii scaling settings can be changed using the options in this section. To scale receptor atom radii, you must choose the appropriate options in the Receptor tab of the Receptor Grid Generation panel prior to grid generation.

Scaling factor

This text box specifies the scaling factor. The default is 0.80. To turn van der Waals radii scaling off, set the scaling factor to 1.0. Full penalties for close contacts of nonpolar ligand atoms will then be used.

Partial charge cutoff

Scaling is performed only on nonpolar atoms, defined as those for which the absolute value of the partial atomic charge is less than or equal to the number in the text box. Since this is an absolute value, the number entered must be positive. The default for ligand atoms is 0.15.

5.4 The Core Tab

In the Core tab, you can specify the "core" of a ligand, and use this core to constrain the docking of other ligands or to calculate the rms deviation from the defined core for all docked ligands. The core is specified in terms of a *reference ligand*.

The constraint that you apply here is a *ligand-based* constraint, which means that the ligands that are subject to the constraint are those that have the same core moiety as the reference ligand. Ligands that do not match the core can be screened out in the first docking stage. In later stages, ligands that match the core pattern but do not meet the rms tolerance for the position of the core can be screened out. Receptor-based constraints can be applied in the Constraints tab, which is described in the next section.

1. Flexible receptor docking can be performed with the Induced Fit Docking solution, which uses Glide with Prime to allow for receptor flexibility. For more information, see the document *Induced Fit Docking*.

EX2037.0063

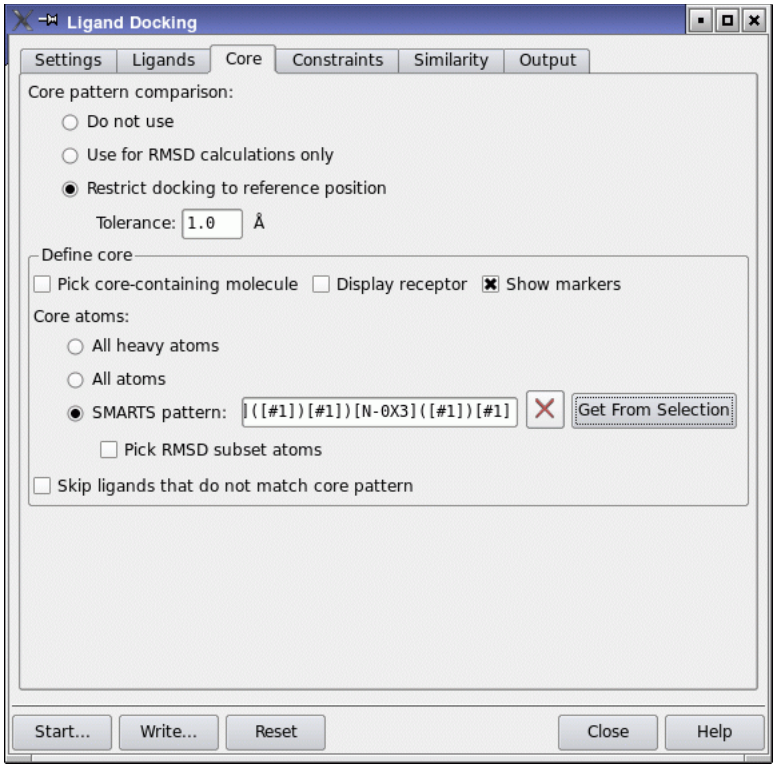


Figure 5.5. The Core tab of the Ligand Docking panel.

You can choose how you want to use the core definition by making a selection under Core pattern comparison:

- Do not use—Do not use a core definition. This is the default.
- Use for RMSD calculations only—Use the core definition for computing the RMSD from the reference ligand, but not for restricting docking.
- Restrict docking to reference position—Restrict the docking of ligands so that the ligand core lies within a given RMSD of the core in the reference ligand. The core is defined in terms of a set of atoms or a SMARTS pattern; if the ligand does not contain the topology defined in the core, it can be skipped. You can enter a tolerance for the RMSD in angstroms for restricting the docking to the reference position in the Tolerance text box.

The Core tab offers a great deal of flexibility in defining the core. You can use the entire molecule (or its heavy atoms) for the RMSD evaluation, or you can define a subset of the molecule, both for the matching and for the RMSD, in terms of SMARTS patterns. Choosing the core

molecule and setting up the core pattern is done in the Define core section of the Core tab. The controls in this section become available when you choose Use for RMSD calculations only or Restrict docking to reference position.

If you plan to use the entire reference ligand for the core (or its non-hydrogen atoms), ensure that Pick core-containing molecule is selected, then pick an atom in the Workspace that belongs to the core molecule. The molecule is marked in purple in the Workspace if Show markers is selected. Pick core-containing molecule is automatically selected when the controls in this section become available. You can then proceed to choose All heavy atoms or All atoms for the core, under Core atoms.

If you want to define the core in terms of a smaller set of atoms, select SMARTS pattern under Core atoms. Pick core-containing molecule is automatically deselected if you make this choice, because you can choose any atoms and picking a molecule is no longer relevant. You can then enter a SMARTS pattern into the text box, or you can select atoms in the Workspace with the Workspace selection tool,

then click Get From Selection to create a SMARTS pattern for the selected atoms. The atoms in the core molecule that match the pattern are marked in the Workspace with green markers if Show markers is selected. Note that it is advisable not to have the receptor displayed when you click Get From Selection, because it can slow the generation of the SMARTS pattern.

You might not want to use all the atoms in the SMARTS pattern for the RMSD calculation. If this is the case, you can choose Pick RMSD subset atoms, and pick atoms in the Workspace that are in the SMARTS pattern to define a subset of atoms for the RMSD calculation. A lock symbol appears next to the atoms you pick. Each pick adds to the set; picking an atom again deletes it. The RMSD atoms are used both when calculating only the RMSD and when restraining the core: it is the RMSD of these atoms that must fall below the prescribed tolerance in the latter case.

Finally, you can select Skip ligands that do not match core pattern if you want to strictly apply the core pattern filter. This option is selected by default. When it is selected, ligands that do not match the core are discarded in the first stage of the funnel. If this option is not selected, ligands that do not match the core pattern are docked anyway, the core constraints are not applied, and the RMSD is only calculated for those ligands that do match the core pattern.

EX2037.0065

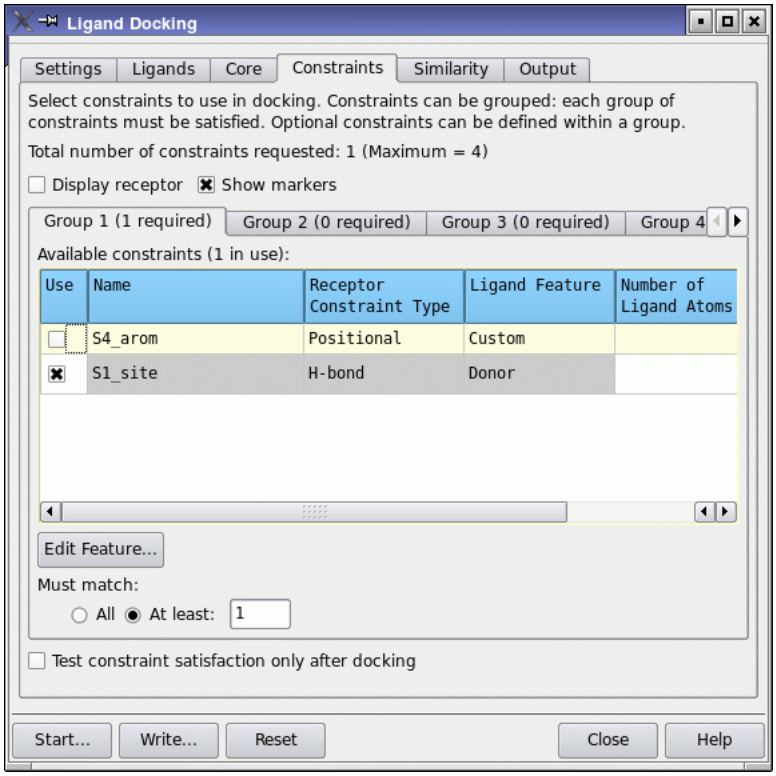


Figure 5.6. The Constraints tab of the Ligand Docking panel.

5.5 The Constraints Tab

The Constraints tab lists all the Glide constraints that are defined for the receptor grid file you specified in the Settings tab, and provides the means to apply these constraints in docking. The available constraints are shown in the Workspace if Show markers is selected and the receptor is displayed. To display the receptor, select Display Receptor. Positional constraints are marked with gray spheres. Atoms for which H-bond/metal constraints are defined are marked with a red asterisk and padlock. Regions for which hydrophobic constraints are defined are marked with translucent gray hydrophobic cells.

Constraints are applied during or after docking by identifying the relevant features in the ligand, and requiring specified atoms in those features to be spatially confined to the constraint region. The ligand features and the atoms in those features that must be constrained are defined in terms of SMARTS patterns. You can customize the features in the Edit Feature dialog box, which is described in Section 5.5.2 on page 61. For positional constraints, no default features are defined, so you must provide a feature definition in the Edit Feature dialog box.

Glide constraints can be used in HTVS mode, but may result in poor pose recovery relative to unconstrained HTVS docking.

Only the constraints you select will be applied in the ligand docking job you are setting up. If there are no constraints selected when you start the docking job, no constraints will be applied.

5.5.1 Setting Constraints

To allow some flexibility in setting constraints, including optional constraints, Glide provides four constraint groups, presented in subtabs labeled Group 1 through Group 4. If you want to set simple constraints you can do so in the Group 1 tab, which is displayed by default, and ignore the remaining groups.

Each subtab contains a table of all the available constraints (described in Table 5.1), an Edit Feature button, and two Must match options. To set constraints within a tab, first click the Use column for each of the desired constraints. An X appears in the check box. If the receptor is displayed and Show markers is selected, the hydrophobic and positional constraint markers turn red when the constraint is selected, and the H-bond and metal constraint markers turn turquoise. The number of constraints you selected is displayed next to the table title.

Table 5.1. Description of the Available constraints table.

Column	Description
Use	Check box to select the constraint for use in docking. Click to select or deselect.
Name	Constraint name. This is the name defined in the Constraints tab of the Receptor Grid Generation panel.
Receptor Constraint Type	Type of receptor constraint. Hydrogen bond constraints are classified into H-bond for hydrogen-bond acceptors and Polar Hydrogen for hydrogen-bond donors (i.e. hydrogen atoms).
Ligand Feature	Name of the feature in the ligand that must match the constraint. The available features are: Acceptor, Charged Acceptor, Neutral Acceptor, Donor, Hydrophobic, and Custom. By default, the feature that matches a receptor polar hydrogen is set to Acceptor; for a receptor H-bond type it is Donor, for a receptor hydrophobic feature it is Hydrophobic, and for a positional constraint it is Custom. Custom is undefined by default, so you must edit this feature to define the patterns that match the desired ligand atoms.
Required Ligand Atoms	In this column, the number of ligand atoms that must occupy a hydrophobic region can be set. The value is only meaningful for hydrophobic constraints.

EX2037.0067

Next, decide whether all the constraints should be applied, or only some of them.

- If you want all of the chosen constraints to be applied, select All under Must match.
- If you want some of the chosen constraints to be applied, select At least under Must match, and enter the number of constraints that must be satisfied in the text box. For example, if you chose three hydrogen-bond acceptors, and you want any two out of the three to be satisfied, you would enter 2 in the text box.

The number of constraints that are required is displayed in the subtab tab, and is added to the total, which is displayed in the upper part of the Constraints tab.

To clear a constraint that you have set, click the Use column again. The X in the check box disappears.

More complex constraints can be applied by using more than one constraint group. If you set constraints in more than one group, each group of constraints is applied to the ligand (that is, a Boolean AND is applied between groups). The general syntax of the constraints is (N1 required from Group 1) AND (N2 required from Group 2) AND (N3 required from Group 3) AND (N4 required from Group 4).

For example, suppose the grid contains a hydrogen bond constraint hbond1 and two hydrophobic constraints hphob1 and hphob2. If you want to enforce the hydrogen bond constraint, and require only one of the two hydrophobic constraints, you would use two groups. In Group 1, you would click the Use check box for hbond1 and select All under Must match. In Group 2, you would click the Use check boxes for hphob1 and hphob2, select At least under Must match, and enter 1 in the text box. Both groups are applied. The resulting constraints can be represented as “hbond1 AND (hphob1 OR hphob2)”.

If your desired constraint specification cannot be put in the general form above, you might be able to achieve your goal by running more than one docking job with a separate constraint specification for each. For example, if you want to apply the constraints “(hbond1 AND hphob1) OR (hbond2 AND hphob2) OR (hbond3 AND hphob3)”, you could run three separate docking jobs, one with (hbond1 AND hphob1) set, one with (hbond2 AND hphob2) set, and one with (hbond3 AND hphob3) set.

The total number of required constraints, summed over all groups, must be four or fewer. The number in parentheses after the group name in the subtab tab is the number of constraints that must be satisfied in this group. If you selected All, it is the number of constraints you chose. If you selected At least, it is the value in the At least text box.

EX2037.0068

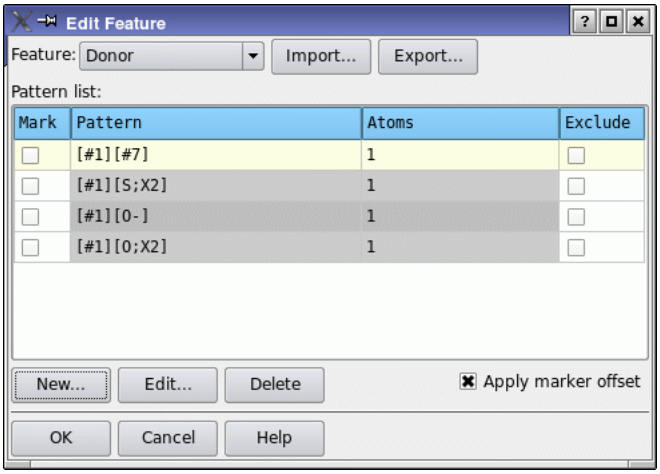


Figure 5.7. The Edit Feature dialog box.

A default ligand feature definition is supplied and assigned to each receptor constraint type, with the exception of positional constraints, for which the Custom feature is undefined. To change either the assignment or the feature definitions, use the Edit Feature dialog box, as described in the next section. You can select the same constraint in more than one group, but the assignment and the feature definitions are the same for each group.

To use positional constraints, you *must* define the ligand feature that it should match. No default feature definition is provided, and the job cannot be run until a feature definition is provided. To define the ligand feature, select the appropriate row in the Available constraints table and click Edit Feature. The feature can then be defined in the Edit Feature dialog box.

5.5.2 Defining Ligand Features

Ligand features are identified by matching to a collection of SMARTS patterns that defines a feature type. There are six feature types: Acceptor, Charged Acceptor, Neutral Acceptor, Donor, Hydrophobic, and Custom. The feature definitions for these six types form a *feature set*, which can be imported and exported. Each constraint can have its own feature definition, so you can have a different definition of a given feature type for each constraint. However, the same feature definition from the same set is used for a given constraint in all groups. For each feature definition, you can add patterns, edit and delete custom patterns, and define patterns for exclusion of functional groups. Feature sets can be imported and exported. These tasks are carried out in the Edit Feature dialog box, which you open by clicking Edit Feature.

To change the feature type assigned to the constraint or to edit the feature definition for a particular constraint, select the row in the Available constraints table, and click Edit Feature.

EX2037.0069

If you only want to change the feature type assigned to the constraint, choose the feature type from the Feature menu, and click OK. The change is reflected in the Ligand Feature column of the Available constraints table, and applies to all groups. If you want to edit the feature definition, follow the instructions in Section 5.5.2.2 and Section 5.5.2.3.

5.5.2.1 Loading and Saving Feature Sets

Built-in feature sets are stored with the distribution, so you do not need to create your own, except for positional constraints.

You can import a feature set for the selected constraint from a file by clicking Import, and navigating to the feature file. When you import a feature set, the definitions of all feature types are replaced, not just the feature type chosen from the Feature menu. The feature definitions are replaced only for the selected constraint, but are replaced for that constraint in all groups.

Feature sets can be saved to a file by clicking Export, and specifying the file location in the file selector that is displayed.

5.5.2.2 Adding, Editing, and Deleting Patterns

The patterns that define a feature are displayed in the Pattern list table (see Table 5.2) when you choose the feature type from the Feature option menu.

Table 5.2. Description of Pattern list table.

Column	Description
Mark	Column of check boxes. Selecting a check box marks the pattern on any structures that are displayed in the Workspace.
Pattern	Pattern definition. The definitions are SMARTS strings.
Atoms	The list of atoms that may be constrained, numbered according to the SMARTS string.
Exclude	Column of check boxes. If a check box is selected, the atoms in other patterns are matched only if they do not match this pattern. This is essentially a NOT operator. Excluded patterns are processed before other patterns.

If the patterns in a given feature definition do not cover all the functional groups that you want to include in the definition, you can add extra patterns. Matching of patterns to ligand structures is done in the order specified in the Pattern list table. You cannot change the order of the patterns once they are in the table, so you must add new patterns at the appropriate point. If you want to move a pattern, you must delete it and add it again.

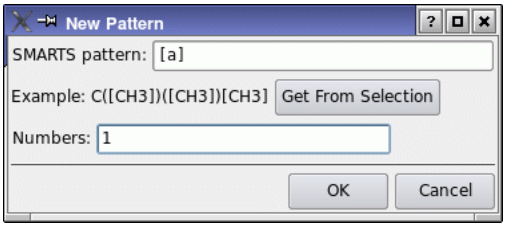


Figure 5.8. The New Pattern dialog box.

To add a new SMARTS pattern, click the table row above which you want the pattern to be inserted, then click New. In the New Pattern dialog box (Figure 5.8), you can provide a SMARTS pattern and define the atoms that must satisfy the constraint: the acceptor or donor atoms, for example.

There are two ways to provide a SMARTS string. The first is to type the string into the SMARTS pattern text box. The second is to select atoms in the Workspace, then click Get from selection. Maestro generates a SMARTS string for the selected atoms and places it in the SMARTS pattern text box. You can then edit it if you like. You should make sure that you have a suitable molecule, such as a known active ligand, in the Workspace before you open the Edit Feature dialog box. This second procedure can be slow if large molecules, such as a receptor, are displayed in the Workspace.

To specify the atoms in the ligand that must satisfy the constraint, enter the atom numbers as a comma-separated list in the Numbers text box. The atom numbers are given relative to the SMARTS pattern: atom 1 is the first atom in the SMARTS pattern, and so on.

For positional, metal, and H-bond constraints, you can specify multiple atoms, but only one of the specified atoms is constrained. Normally, only one atom should be specified, to ensure that the correct atom is constrained. However, if the functional group contains multiple acceptors or donors, for example, you can specify all of them if you want any one of them to be constrained. This is not necessary for groups such as carboxylates, where the local symmetry of the functional group is used and either of the oxygen atoms in the carboxylate can be selected for the constraint, even though only one of them is specified.

For hydrophobic constraints, the non-hydrogen atoms of the hydrophobic group should be specified. The number of atoms that is constrained is specified in the Required Ligand Atoms column of the Available constraints table.

When you click OK, a new row is added to the Pattern list table.

To edit a pattern, select the table row for the pattern, then click Edit. In the Edit Pattern dialog box, you can modify the SMARTS pattern or get a new pattern from the Workspace selection,

EX2037.0071

and change the atoms in the pattern that must satisfy the constraint. The Edit Pattern dialog box is identical to the New Pattern dialog box except for the title.

To delete a pattern, select the table row for the pattern, then click Delete.

5.5.2.3 Excluding Functional Groups from a Feature

If you want to ensure that certain functional groups are not matched, you can select the check box in the Exclude column for the pattern for that group. For example, you might want to exclude a carboxylic acid group from being considered as a hydrogen bond donor, because it will be ionized under physiological conditions. Excluded patterns are processed first, regardless of their position in the Pattern list table. Thus, an excluded pattern prevents the atoms that match it from being matched by any other pattern.

5.5.2.4 Visualizing Patterns

If you want to see a pattern for a given ligand or group of ligands, you can select the check box in the Mark column for the pattern. Any occurrences of the pattern are marked in the Workspace. You can display markers for more than one pattern, but the markers do not distinguish between patterns. If you want to see the atoms and bonds as well as the markers, select Apply marker offset.

5.5.3 Choosing When To Apply Constraints

Constraints can be applied during the docking process, or only after the ligands have been docked. When testing for constraint satisfaction after docking, constraints are used as a post-docking filter and have no influence on the docking or scoring. You can therefore compare results from such a run directly with an unconstrained docking run. To apply constraints only after docking, select Test constraint satisfaction only after docking.

5.5.4 Using Multiple Constraints

When you use multiple constraints with complex optional constraint specifications and many ways of satisfying the constraints, it is possible to lose poses because Glide cannot track all possible ways a given pose can satisfy the constraints, due to memory limitations. Under these circumstances, Glide truncates the list of poses or the selections of constraint-satisfying atoms per pose, or both. The result is that poses that would ultimately satisfy the constraints are lost.

There are several ways in which you can alleviate this problem:

- Select Test constraint satisfaction only after docking in the Constraints folder when setting up the docking job. (If you run from the command line, add the keyword `finalonly` after `constraints` in the *jobname*`.inp` file.) This avoids the memory explosion needed

EX2037.0072

to track complicated constraint combinations through the Glide funnel, but still ensures that the final docked poses satisfy the required constraints.

- Use more restrictive SMARTS patterns to define the kinds of ligand atoms that can satisfy the constraints. This can reduce the memory requirements for tracking constraint satisfaction in a given ligand to a point where all constraint-satisfying poses can be kept. If you need to allow more general chemistry to satisfy your constraints, you can run multiple Glide jobs to cover the various possibilities.
- Split your job into several jobs with tighter restrictions on which constraints must be satisfied, while still spanning all the desired possibilities. For instance, suppose you want your ligands to make hydrogen bonds to each of three protein hydroxyl groups, but don't care which molecule is the donor and which is the acceptor for each group. A single job to cover all the possibilities would have three groups of two constraints each, each group consisting of H-bond constraints to the hydrogen and oxygen of one of the protein hydroxyls, and would require one of the two constraints in each group to be satisfied. This job is likely to run into the memory problems and lose poses, particularly for ligands that have multiple H-bonding groups (e.g. sugars). Instead, you could run eight Glide jobs, one for each specific combination of donors and acceptors among the three protein hydroxyls. Each job would specify a single group of three constraints, all required, and would be less likely to lose poses due to memory restrictions.

5.6 The Similarity Tab

You can use the options in the Similarity tab to set up and incorporate atom pair similarity scoring in your docking job. Similarity algorithms provide a mechanism for quantifying how alike or unlike two molecules are. Similarity can be used to select molecules that are similar to known actives from a large molecular database, based on the "similarity property principle", which states that closely related chemical structures should share similar biological activity and physiochemical properties [2]. Similarity can also be used to find molecules in a database screen that dock and score well yet are *not* similar to any of a series of known actives. In this way, the docking algorithm can be directed away from ligands that are structurally related to what is already known, and toward structures with new chemical features.

Glide uses an atom-pair similarity scoring algorithm [3–8]. In atom-pair similarity, the two molecules being compared are first processed to generate sets of atom pairs. Each non-hydrogen atom is represented by a similarity type based on the connectivity, bond orders, number of pi electrons, and formal charges of the molecule. For each pair of similarity types, the shortest path (the path with the smallest number of connections) is determined. The unique combination of

type(atom A) + connectivity distance + type(atom B)

EX2037.0073

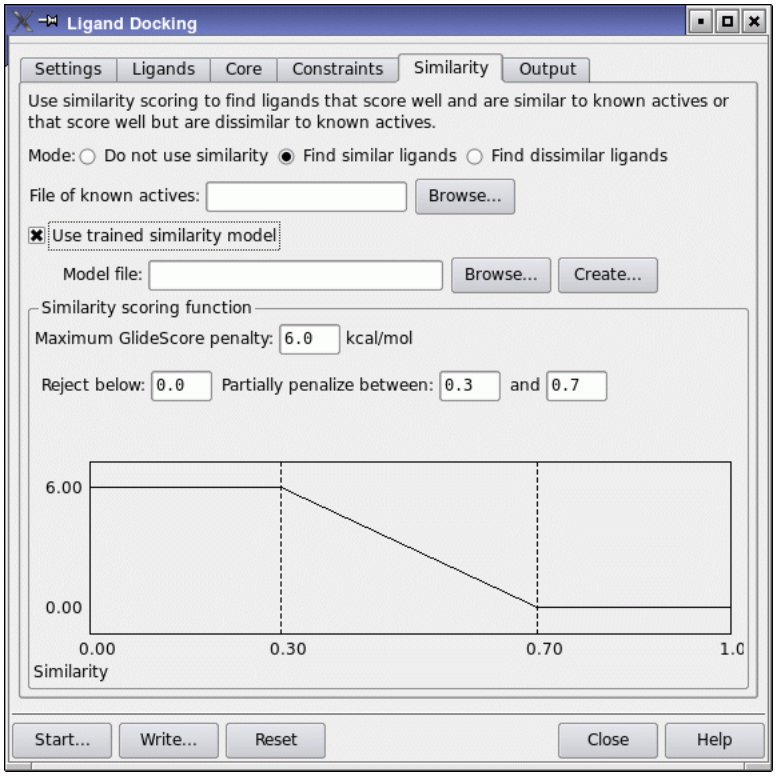


Figure 5.9. The Similarity tab of the Ligand Docking panel.

defines one atom pair. All atom pairs for a given molecule constitute the atom pair list for that molecule. The similarity between two molecules is a function of the number of atom pairs that appear in both lists. The similarity function is normalized so that the result is a number between 0.0 (no atom pairs in common) and 1.0 (identical atom pair lists).

Similarity or dissimilarity scoring is combined with the results of Glide docking to produce a composite score. By default, no similarity scoring is performed.

5.6.1 Setting up Similarity Scoring

To set up similarity scoring, you must select a mode and provide a file of known actives. The possible modes are:

- Do not use similarity

  Docks and scores ligands without calculating or using similarity scores. This option is selected by default.

- Find similar ligands

  Penalizes ligands that are different from known actives. In the results of the docking job, ligands that score well and are similar to known actives are ranked high.

- Find dissimilar ligands

  Penalizes ligands that are similar to known actives. In the results of the docking job, ligands that score well and are different from known actives are ranked high.

To specify the file of known active ligands, type the file name in the File of known actives text box, or click the Browse button to open a file selector, in which you can choose the format and navigate to the file. The file must be in Maestro, SD, or MOL2 format.

5.6.2 Creating a Trained Similarity Model

In a *trained similarity model*, contributions from atom pair matches to the similarity score are weighted unequally. The weights are determined from a training calculation that uses a set of active ligands and a set of inactive ligands to maximize active-active similarity scores and minimize active-inactive similarity scores. Thus, atom pair matches to inactive ligands are penalized in the similarity score, while atom pair matches to active ligands are rewarded. The weights are stored in a file named *filename*`.wgt`. This file is referred to as the *trained similarity model file*, or just *model file*.

To make use of a trained similarity model, select Use trained similarity model. The controls for specifying the model are activated. If you already have a model file, you can enter the name in the Model file text box, or click Browse and navigate to the file in a file selector.

To generate a trained similarity model, click Create. The Create Similarity Model panel opens, in which you can set up a job to generate a set of optimized weights for a trained model.

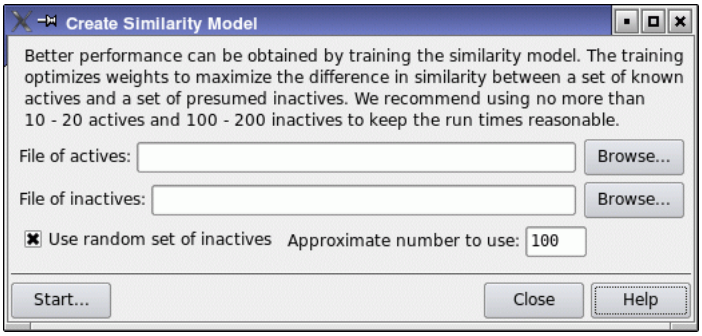


Figure 5.10. The Create Similarity Model panel.

EX2037.0075

To generate the weights file, Glide requires a file of active ligands and a file of presumed inactives. You can specify these files in the File of actives and File of inactives text boxes, or click the Browse button to open a file selector, in which you can choose the format and navigate to the file. The file must be in Maestro, SD, or MOL2 format. Information on choosing the files and ligands is given below:

- File of actives

  You can use the same file you specified in the Similarity tab for the file of actives. The file must contain at least two active ligand structures. Best results are obtained with at least 10, but more than 20 are unnecessary and may increase calculation time. To take full advantage of the weighting algorithm, include actives with a wide variety of structures.

- File of inactives

  In addition to known inactives, this file may include untested compounds which are very unlikely to be active. The file of inactives must contain at least one ligand-like structure, but more will improve results. Using 5 to 10 times the number of actives is optimal.

Typical numbers of structures might be 15 structures in the actives file and 150 structures in the inactives file.

If your inactives file is much larger than about 200 structures, or 10 times the number of actives, using all the structures in the inactives file can lead to convergence problems in the training algorithm. To avoid this difficulty you can use a random subset of inactives, by selecting Use random set of inactives and entering an approximate number of inactive ligands in the Approx. number to use text box.

When you have specified the actives and inactives to use, click Start. The Create Similarity Model - Start dialog box opens, in which you can specify a job name, a user name, and a host, then click Start to launch the job. When the job finishes, the weights are written to the trained similarity model file *jobname*`.wgt`, and the file name appears in the Model file text box.

5.6.3 Similarity Scoring Function

Similarity information is incorporated into docking results according to a scoring function that assigns a GlideScore penalty based on a ligand's similarity score, which has been normalized to a value between 0.00 and 1.00. The general shape of the scoring function depends on your choice of similarity mode. If you have selected Find similar ligands, penalties decrease with increasing similarity. If you have selected Find dissimilar ligands, penalties increase with increasing similarity.

The scoring function is controlled by parameters that you can change if necessary. The Similarity scoring function section of the Similarity tab contains the four parameter settings for the

EX2037.0076

similarity scoring function. Beneath these options, the GlideScore penalties are plotted against similarity score, which ranges from 0.0 to 1.0. The plot is updated to reflect a change in the value of a parameter when you press ENTER or click in the next parameter text box.

The scoring function options are described below. In describing features that differ according to similarity mode (Find similar ligands or Find dissimilar ligands), the default is assumed to be Find similar.

Maximum GlideScore penalty *p*

This is the maximum penalty for ligands with low similarity scores (or for ligands with high similarity scores, if Find dissimilar ligands is selected). Changing this number adjusts the relative importance of the similarity score and the basic GlideScore to the combined result. Increasing the maximum penalty increases the importance of similarity relative to docking; decreasing the maximum penalty has the reverse effect. The default value of *p*, 6.00 kcal/mol, gives similarity score and GlideScore roughly equal weighting, which is generally optimal.

Reject below *r*

Ligands with similarity scores below the threshold *r* will not be docked. The default value of *r* is 0.00. Since similarity scores cannot be negative, the default is to dock all ligands regardless of similarity.

If the mode is Find dissimilar ligands, the option is Reject Above: ligands with similarity higher than *r* will not be docked. The default is 1.00, a score that cannot be exceeded, and therefore corresponds to docking all ligands.

Partially penalize between *m* and *n*

Ligands with similarity scores lower than *m* receive the maximum penalty. Those with similarity higher than *n* receive no penalty. Similarity scores between *m* and *n* receive a penalty that decreases linearly from the maximum value at *m* to zero at *n*, as shown in the scoring function plot. The default value of *m* is 0.30. The default value of *n* is 0.70.

In Find dissimilar ligands mode, the default values of *m* and *n* remain the same, but no penalty is applied to ligands with similarity scores of *m* or lower. Scores between *m* or *n* receive increasing penalties, reaching the maximum penalty at *n* and above.

EX2037.0077

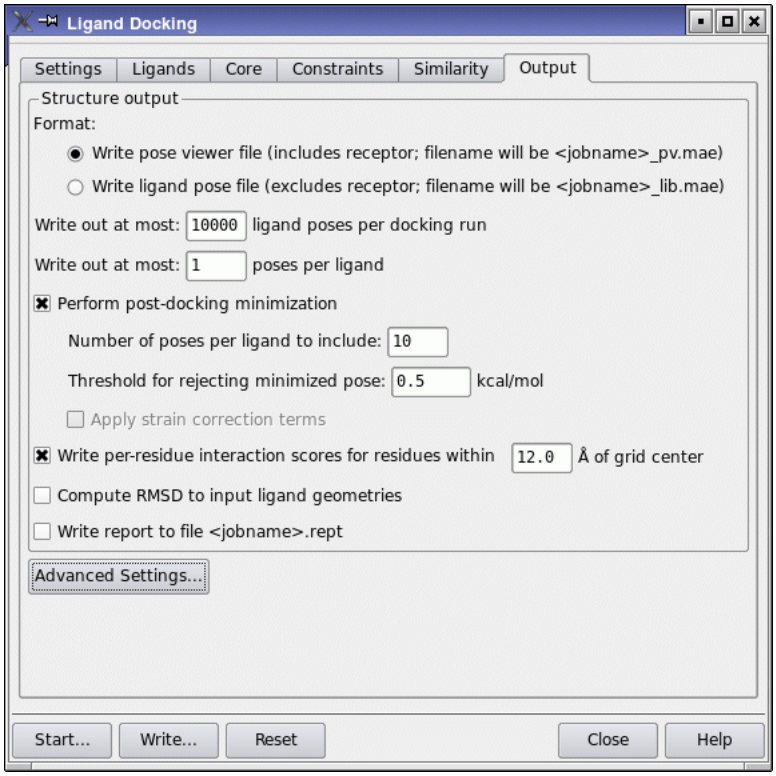


Figure 5.11. The Output tab of the Glide panel.

5.7 The Output Tab

The options in the Output tab control the final output of ligand poses that pass successfully through Glide's various scoring stages. The features of this tab are options in a section called Structure output and an Advanced Settings button. Clicking the button opens the Output - Advanced Settings dialog box.

5.7.1 Structure Output Options

In flexible docking runs, Glide appends docked poses to a file named *jobname*_`raw.mae`. The "raw" in the filename indicates that these poses are not sorted by GlideScore. Once the Glide job has docked all the ligands, it runs the `glide_sort` utility script to sort the raw poses and write them to *jobname*_`pv.mae` or *jobname*_`lib.mae`. The *jobname*_`raw.mae` file is then deleted.

EX2037.0078

Rigid docking runs sort the poses internally, then, at the end of the job, the sorted poses are written to either *jobname*_pv.mae or *jobname*_lib.mae. The glide_sort utility is not automatically used in rigid docking jobs.

For score-in-place jobs, the structure output file contains the input structures, since they are not altered.

The output files are returned in compressed format, i.e. as .maegz files rather than as .mae files.

The glide_sort utility can sort any pose file, whether _raw.mae, _pv.mae, or _lib.mae, using default or alternate criteria. For more information on how to use glide_sort, see Section 7.4.1 on page 101.

The options in the Structure output section are described below.

Format

The final list of poses that pass Glide's criteria are written to a multi-structure Maestro-format (.mae) pose file. You can select one of two options:

- Write pose viewer file (includes receptor; filename will be <jobname>_pv.mae)

  This version of the pose file, which includes the receptor, is intended for use with the Pose Viewer tool in the Maestro Project Table.

- Write ligand pose file (excludes receptor; filename will be <jobname>_lib.mae)

  This option includes only the ligand structures in the output pose file. This version of the pose file cannot be used in Pose Viewer mode, but may be appropriate if the output poses are intended for input to a subsequent Glide job, or for some other purpose.

Write out at most *n* ligand poses per docking run

This text box limits the total number of the predicted best-binding poses written to the sorted pose file. 10 000 poses is the default option.

Write out at most *m* poses per ligand

This text box limits the number of poses per ligand written to the sorted pose file. The default choice of 1 pose per ligand is intended for database screening applications. A larger choice may be appropriate for lead-optimization studies or whenever several "reasonable" poses are wanted; for example, to generate a variety of docked poses for study with Liaison or another post-docking program. However, if you have a small binding pocket or ligands with few rotatable bonds, choosing a larger number might simply retain poor poses.

EX2037.0079

Compute RMSD to input ligand geometries

This option enables the computation of the RMSD between the docked poses and the corresponding input ligand geometries. The value is recorded as a Maestro property named glide rmsd to input (internal name `r_i_glide_rmsd_to_input`).

Write report to file <jobname>.rept

A report file that lists critical information about the scoring of the retained poses is written to *jobname*`.rept` if this option is selected. Like the pose viewer file or ligand pose file, the report file lists the poses in rank order by the selected final-scoring function. The report file is not written for distributed jobs, unless you run them from the command line with `para_glide`.

For score-in-place jobs, instead of the `.rept` file, a `.scor` file is produced, containing scoring information for the input structures.

5.7.2 Post-Docking Minimization

The docking process relies on rapid generation of ligand conformers and use of a grid to represent the receptor. The ligand poses generated during docking are rarely exactly at a local minimum, and post-minimization can improve the geometry of the poses. The Grid Minimization step of the Glide funnel (see Figure 2.2 on page 7) does not perform a full force-field minimization and excludes interactions beyond 1,5 interactions. The full force-field minimization performed by post-docking minimization penalizes highly strained ligand geometries and eliminates poses with eclipsing interactions, many intraligand close contacts, and so on.

The Perform post-docking minimization option specifies a full force-field minimization of top-ranked poses following the final docking. This minimization optimizes bond lengths and angles as well as torsional angles, and rescores the poses using the scaled Coulomb-van der Waals term and the GlideScore. You can perform the post-docking minimization for both flexible and rigid docking.

When you perform post-docking minimization, you can select the number of poses per ligand to minimize. Testing has indicated that the rescoring of poses after post-docking minimization generally finds a lower GlideScore than is reported in the top few poses from docking. Thus, it is strongly recommended to apply the minimization to a number of poses. You can enter the number in the Number of poses per ligand to include text box. This number must not be less than the number of poses per ligand that is written to the output file. The default is 10 for XP and 5 for SP or HTVS. If the binding pocket is small, or the ligands are fairly rigid, the post-docking minimization will not improve the results for poses that do not fit or are already poorly aligned to the receptor.

EX2037.0080

For HTVS and SP docking, the minimized poses replace the original poses. For XP docking, the pose is kept if the GlideScore is no higher than that of the original pose by less than a certain threshold, otherwise the original pose is kept. This threshold is specified in the Threshold for rejecting minimized pose text box.

The time taken for post-docking minimization is less than 1% of the total docking time for SP and XP docking, and can be around 10% for HTVS.

In addition to minimizing, you can choose to calculate and apply a strain correction when doing the final scoring, by selecting Apply strain correction terms. These terms are evaluated by optimizing each ligand pose as a free ligand, first with constraints on all torsions, then without these constraints. The difference is used to compute a penalty term for unreasonably high strain: the strain correction is only added if it is above a threshold, and the excess strain above this threshold is scaled before adding it to the GlideScore. You can set the threshold and the scaling factor in the Output - Advanced Settings dialog box.

5.7.3 Per-Residue Interaction Scores

To examine interactions between a ligand and individual, nearby residues of the receptor, you can write out per-residue interaction scores for residues within a selected distance of the grid center. The Coulomb, van der Waals, and hydrogen bonding scores and the minimum distances are calculated between the ligand and the specified residues. These values are written as structure-level properties for each ligand to the output file, as well as to the log file. To enable writing, select Write per-residue interaction scores for residues within *N* Å of grid center and enter a value in the text box.

Per-residue interactions can be visualized in the Workspace when viewing poses. See Section 6.1 on page 77 for more information.

5.7.4 Advanced Settings

In the Output - Advanced Settings dialog box, you can set options to screen out poses that either have too high an energy or are too similar to other poses, and change the parameters associated with strain correction. The thresholds for rejection of poses are set into two sections, Filter and Clustering; the strain correction thresholds are set in the Strain Correction section.

The Filter section contains a single option, Reject poses with Coulomb-vdW energy greater than *x* kcal/mol. If the pose has a coulomb-vdW score greater (more positive) than this value, the pose is rejected. The default value is 0.0 kcal/mol. This means that poses that interact favorably with the protein site, however weakly, are retained, whereas poses that interact unfavorably are rejected. Change *x* to a negative value to reject poses with weakly favorable interactions, or to a positive value to keep poses with mildly unfavorable interactions.

EX2037.0081

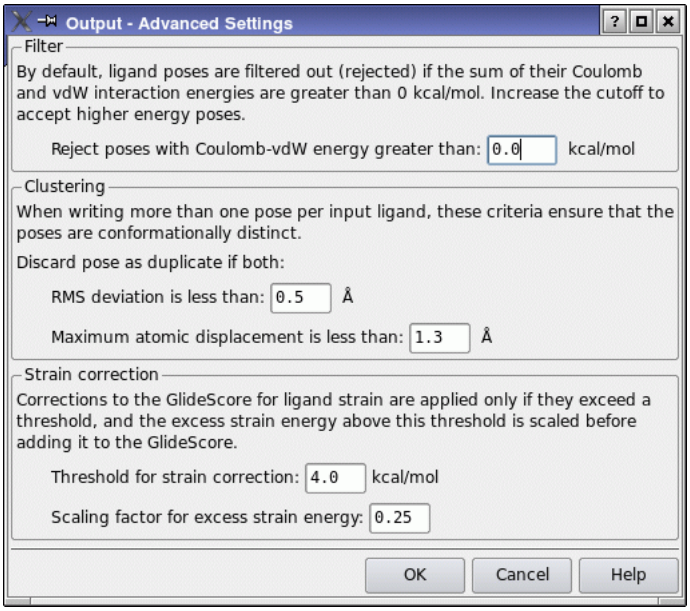


Figure 5.12. The Output - Advanced Settings dialog box.

The clustering options determine which ligand poses are sufficiently alike to be considered duplicates. To determine whether a pose is a duplicate, Glide uses a sophisticated algorithm for recognizing "number reorderings" that need to be taken into account—for example, to accommodate trivial variations such as a 180º rotation of a phenyl group. Ligand poses are compared to those previously selected for inclusion in the reported output, and are discarded as duplicates if both of the following criteria are met. This ensures that reported poses are conformationally distinct. The criteria are:

- RMS deviation is less than *d* Å

  For a pose to be considered a duplicate of one already scored the heavy-atom (nonhydrogen) RMS deviation must be less than *r* Å. The default RMS deviation threshold for *r* is 0.5 Å; choose any value greater than 0.0 Å.

- Maximum atomic displacement is less than *d* Å

  For a pose to be considered a duplicate the maximum atomic displacement must be less than *d* Å. The default is 1.3 Å; choose any value greater than 0.0 Å.

The GlideScore is corrected for excessive strain of the ligand by adding to it a fraction of the strain energy that exceeds a threshold value. This correction is made only if you have selected Apply strain correction terms in the Output tab. The threshold value is specified in the Threshold for strain correction text box, and the fraction of the excess strain energy, expressed

EX2037.0082

as a decimal number, is specified in the Scaling factor for excess strain energy text box. The default value of the threshold is 4.0 and of the scaling factor is 0.25.

5.8 Docking Output Properties

In addition to the docked poses, HTVS and SP docking runs generate a number of Maestro properties, including terms in the scoring function; various indexes; and ligand efficiency metrics, which allow discrimination on the basis of size for ligands that have similar GlideScore values. These properties are listed in Table 5.3. See page 8 for a description of the GlideScore terms and energy components.

Table 5.3. Maestro properties generated by HTVS and SP docking runs.

Property	Description
glide lignum	Ligand number
docking score	Docking score, including all additional terms.
glide gscore	GlideScore
glide lipo	Lipophilic contact plus phobic attractive term in the GlideScore.
glide hbond	Hydrogen-bonding term in the GlideScore.
glide metal	Metal-binding term in the GlideScore.
glide rewards	Various reward or penalty terms
glide evdw	Van der Waals energy.
glide ecoul	Coulomb energy.
glide erotb	Penalty for freezing rotatable bonds in the GlideScore.
glide esite	Term in the GlideScore for polar interactions in the active site.
glide emodel	Model energy, Emodel.
glide energy	Modified Coulomb-van der Waals interaction energy
glide einternal	Internal torsional energy
ligand efficiency	GlideScore / (number of heavy atoms)
ligand efficiency sa	GlideScore/ (number of heavy atoms)$^{2/3}$. This efficiency metric approximates the effect of surface area.
ligand efficiency ln	GlideScore/(1 + ln(number of heavy atoms))

EX2037.0083

Table 5.3. Maestro properties generated by HTVS and SP docking runs. (Continued)

Property	Description
glide confnum	Conformer index for the pose
glide posenum	Pose index of the pose
glide RMSD to input	RMSD between the docked poses and the input geometries. Only present if Compute RMSD to input ligand geometries was selected in the Output tab.

EX2037.0084

Visualizing Docking Results

Several tools are provided to help visualize the results of docking runs. The View Poses facility in the Project Table panel enables you to display the ligand poses with the receptor in the Workspace, along with hydrogen bonds, bad and ugly contacts, and per-residue interaction information. This facility is described in the next sections.

For Glide SP and XP docking runs, you can visualize the contributions to the XP docking score, provided that descriptor information was requested in the docking run. Visualization in the Workspace is done using the Glide XP Visualizer panel, which is described in Section 6.2.

6.1 Viewing Poses

The Project Table panel offers a special facility for viewing poses from a pose viewer file. This facility allows you to step through poses with the receptor fixed in the Workspace. You can also define and display hydrogen bonds, contacts, and per-residue interactions, and count hydrogen bonds and contacts. If you have flexible hydroxyl groups ("rotatable groups") in the receptor, these are shown in the correct orientation for each pose. These capabilities are available from the View Poses submenu of the Entry menu.

To use this facility you must select a single entry group. The group must contain the receptor as the first entry in the group, followed by the ligands. This is the normal situation when you import a pose viewer file into the project.

To start viewing poses, choose Setup from the View Poses submenu. The receptor is locked in the Workspace, and the first ligand entry is included. You can step through the poses using the RIGHT ARROW and LEFT ARROW keys, or the ePlayer controls on the toolbar. For example, you can use the ePlayer to play through the poses automatically, and perform actions on the poses with a script. When you have finished viewing poses, exclude the receptor from the Workspace. For more information on the ePlayer, see Section 8.6 of the *Maestro User Manual*.

In addition to displaying the poses, you can display hydrogen bonds, contacts, and per-residue interactions (if they have been calculated). To do so, choose the appropriate item from the View Poses submenu. The H-bonds and contacts are turned on for each ligand in the entry group. You can also set the parameters that define H-bonds and contacts from this submenu. The Define H-bonds and Define Contacts items open the Measurements panel at the relevant tabs. If you want to write out the H-bonds or contacts, you can do so from the Measurements panel.

EX2037.0085

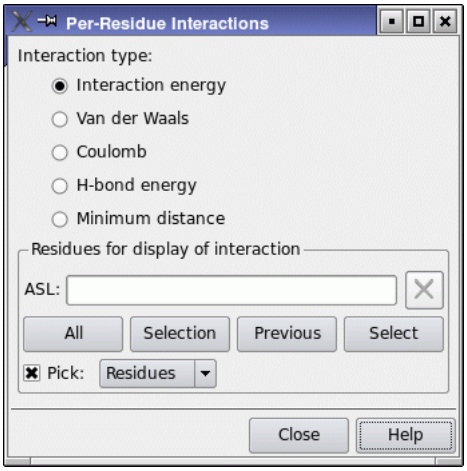


Figure 6.1. The Per-Residue Interactions panel.

For more information on the Measurements panel, see Section 9.2 of the *Maestro User Manual*.

Display of per-residue interactions can only be done if you selected Write per-residue interaction scores for residues within *N* Å of grid center in the Output tab of the Ligand Docking panel. The display is turned on by choosing Display Per-Residue Interactions from the View Poses submenu. The per-residue interactions are represented by coloring the residues by the value of the interaction, using a color ramp. To select the residues that are colored and which interactions are visualized, choose Define Per-Residue Interactions from the View Poses submenu, and make settings in the Per-Residue Interactions panel. In this panel you can select the interaction type, and pick residues for which the interaction is displayed.

If you have defined a set of receptor rotatable groups (hydroxyls in Ser, Thr, and Tyr residues) in the Rotatable Groups tab of the Receptor Grid Generation panel (see Section 4.5 on page 42), you can export the receptor with different orientations of these groups. The coordinates of the receptor in the pose viewer file are written for the original orientation, but the receptor includes information on the alternative orientations. This option allows you to write a file with the coordinates for alternative orientations, which you can then use in other applications.

To do so, you should first include one of the poses that has this orientation in the Workspace. Next, from the Entry menu, choose View Poses, then Export Rotatable Groups. In the Export Rotatable Groups dialog box you can choose whether to export just the receptor, the receptor and the included pose, or the receptor and all poses for which the receptor has that particular orientation of the rotatable groups.

EX2037.0086

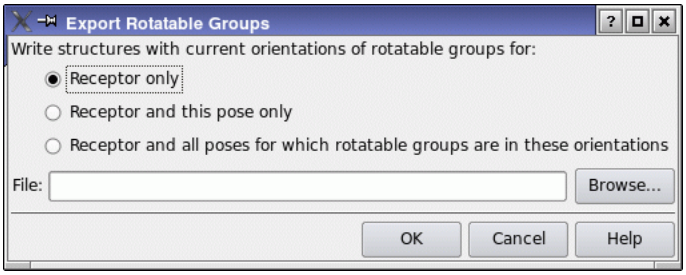


Figure 6.2. The Export Rotatable Groups dialog box.

6.2 The Glide XP Visualizer

The Glide XP Visualizer panel provides a way to visualize and analyze the results of a Glide XP docking run. The main functions of this panel are:

- To display the Glide XP results from a pose viewer file (*jobname*_pv.mae) in a table of XP terms for each ligand.
- To provide 3D visualizations for XP terms. Information for these visualizations is read from the descriptor file (*jobname*.xpdes), which is generated by Glide XP and must be located in the same directory as the pose viewer file.
- To allow selective evaluation of ligands (and groups of ligands) within the table. This helps you analyze ligands separately during the screening process.

Before you can use the Glide XP Visualizer, you must generate the descriptor information. This information is not included in a normal XP run. To generate it, select Write XP descriptor information in the Settings tab of the Ligand Docking panel. You should also select Write pose viewer file in the Output tab of the Ligand Docking panel to write the required pose viewer file.

Note: Generation of XP descriptor information requires a separate license.

To open the Glide XP Visualizer panel, choose XP Visualizer from the Glide submenu of the Applications menu.

When the panel opens, click Open to select the descriptor file. The corresponding pose viewer file must have the same base name. The file name and receptor name are displayed in the File and Receptor text boxes.

After the data has been loaded, the Select Activity Property dialog box opens so that you can select an activity property to display in the table. Choose a property from the list and click OK if you want to display the activities in the ligands table, or click Cancel if you do not want to display an activity property. If you do not choose a property, the Activity column is present, but

EX2037.0087

shows '--' for the activity. You can change an activity value by clicking in the table cell. When you do so, the Edit Activity dialog box opens, and you can enter a new value.

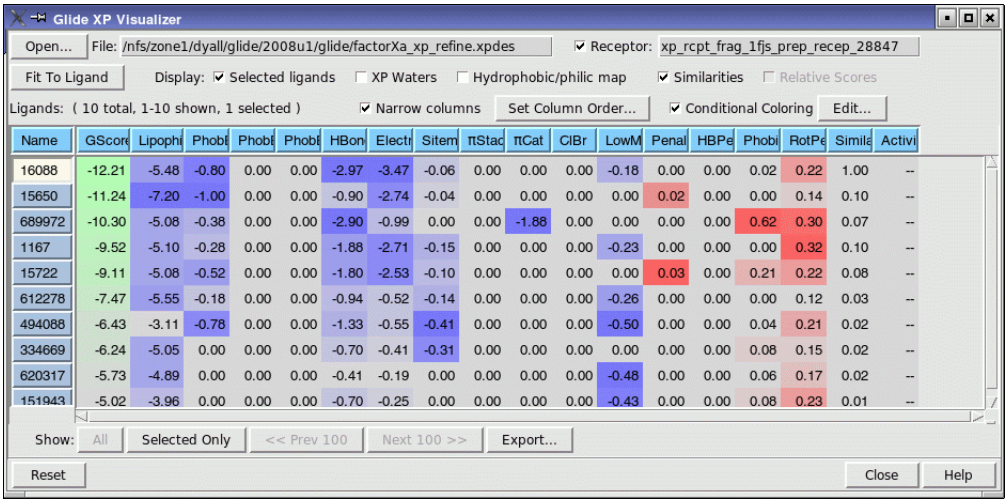
6.2.1 The Ligands Table

When you open a file, the ligands are loaded into the table, with the name of the ligand in the far left column, along with the XP terms in the rest of the columns. If the same ligand name is listed twice in the pose viewer file, then a *-N* is added to the ligand name, where *N* is an integer starting with 2 that is incremented for each successive instance of a ligand name. By default the first ligand is selected.

To sort the rows of the table by the values in a column, click on the column heading. There are three sort settings: ascending (+), descending(-), and the original pose viewer order. These settings can be changed by successively clicking on the column heading. One column can be sorted at a time. The sorting mechanisms work on the ligands shown in the table.

To change the order of the columns in the table, click Set Column Order, which opens the Set Column Order dialog box. In this dialog box, you can select the columns from a list and move them up and down in the list. When you click OK, the columns are reordered with the top column in the list at the left of the table. You can also order the columns by decreasing values of the variance, which is listed in the list along with the column name.

The background color of the cells can be chosen to reflect the values of the XP terms. To color the cells, select Conditional Coloring. To set up the coloring scheme, click Edit, and make changes in the Edit Conditional Coloring dialog box:



Name	GScore	Lipophi	Phobi	Phobi	Phobi	HBon	Electr	Sitem	πStac	πCat	ClBr	LowM	Penal	HBPe	Phobi	RotPe	Simila	Activi
16088	-12.21	-5.48	-0.80	0.00	0.00	-2.97	-3.47	-0.06	0.00	0.00	0.00	-0.18	0.00	0.00	0.02	0.22	1.00	--
15650	-11.24	-7.20	-1.00	0.00	0.00	-0.90	-2.74	-0.04	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.14	0.10	--
689972	-10.30	-5.08	-0.38	0.00	0.00	-2.90	-0.99	0.00	0.00	-1.88	0.00	0.00	0.00	0.00	0.62	0.30	0.07	--
1167	-9.52	-5.10	-0.28	0.00	0.00	-1.88	-2.71	-0.15	0.00	0.00	0.00	-0.23	0.00	0.00	0.00	0.32	0.10	--
15722	-9.11	-5.08	-0.52	0.00	0.00	-1.80	-2.53	-0.10	0.00	0.00	0.00	0.00	0.03	0.00	0.21	0.22	0.08	--
612278	-7.47	-5.55	-0.18	0.00	0.00	-0.94	-0.52	-0.14	0.00	0.00	0.00	-0.26	0.00	0.00	0.00	0.12	0.03	--
494088	-6.43	-3.11	-0.78	0.00	0.00	-1.33	-0.55	-0.41	0.00	0.00	0.00	-0.50	0.00	0.00	0.04	0.21	0.02	--
334669	-6.24	-5.05	0.00	0.00	0.00	-0.70	-0.41	-0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.15	0.02	--
620317	-5.73	-4.89	0.00	0.00	0.00	-0.41	-0.19	0.00	0.00	0.00	0.00	-0.48	0.00	0.00	0.06	0.17	0.02	--
151943	-5.02	-3.96	0.00	0.00	0.00	-0.70	-0.25	0.00	0.00	0.00	0.00	-0.43	0.00	0.00	0.08	0.23	0.01	--

Figure 6.3. The Glide XP Visualizer panel.

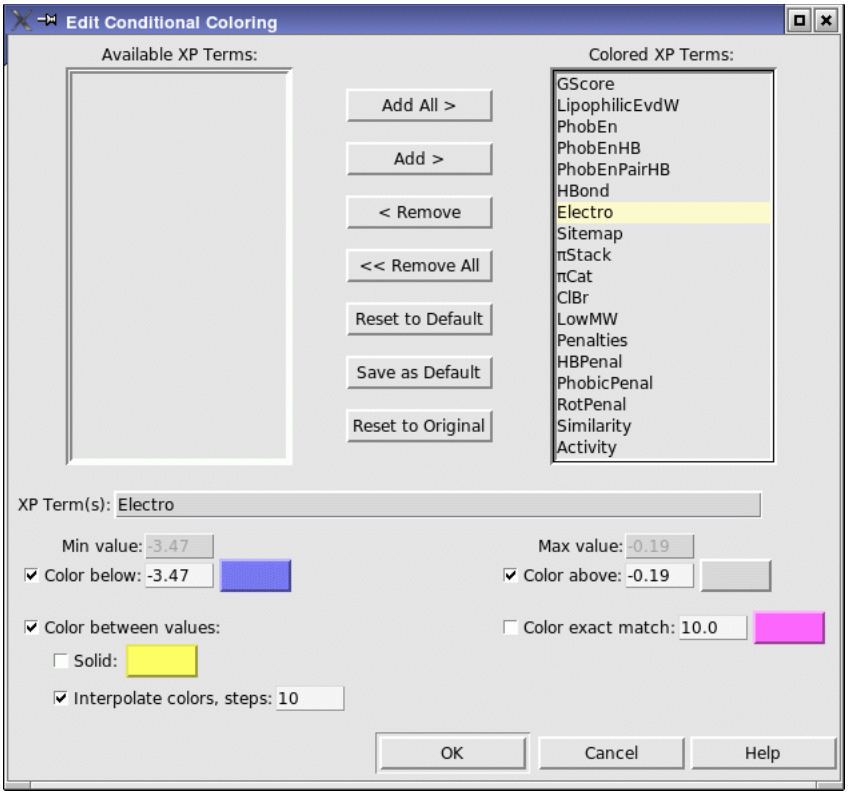


Figure 6.4. The Edit Conditional Coloring dialog box.

- To remove coloring for XP Terms, select them in the Colored XP Terms list, and click Remove. The terms are transferred to the Available XP Terms list, and the color scheme is no longer displayed. (The color scheme is still associated with the terms, so it can be reinstated).
- To reinstate coloring for XP Terms, select them in the Available XP Terms list, and click Add. The terms are transferred to the Colored XP Terms list, and the color scheme is displayed.
- To change the color scheme for a term, select it in the Colored XP Terms list, and use the controls in the lower half of the panel to set up the scheme. You can choose to apply a color to terms that are below or above a given value, to color between these thresholds either with a single color or a color ramp, and to color particular values.

EX2037.0089

Clicking on cells in the table produces two different functions depending on the cell type:

- XP Term Cells—The underlining of the text in the XP term cells signifies that a visualization exists for that XP term value. By clicking on one of these cells, the cell is highlighted and the visualization is displayed in the Workspace. These visualizations depend on the XP term and are described below.
- Ligand Name Cells—You can select multiple ligands by clicking in the Name column, using shift-click and control-click. If Selected ligands is selected in the Display section, then Maestro shows the selected ligands and the view changes each time the selection is changed.

You can step through the ligands in the table using the LEFT ARROW and RIGHT ARROW keys.

The table columns can be widened so that the entire heading is displayed by deselecting Narrow columns. The column descriptions are displayed in tooltips. If you want to change the column order, click Set Column Order and use the tools in the Set Column Order dialog box to move the columns in the list. You can select multiple columns and move them to the top or bottom of the list, or move them up or down in the list.

In addition to selecting ligands in the table, you can limit what is shown in the table by using the Show buttons. By default, only the first 100 ligands are shown in the table. You can show the next 100 ligands or the previous 100 ligands using the corresponding buttons below the table. Clicking Selected Only when there are multiple ligands selected shows only the selected ligands in the table. The All button is used to show all ligands in the table.

The Ligands table also lists a set of fragments that maximize the values of the XP descriptors, if these have been docked along with the ligands. These fragments are a special set of fragments that are docked using the Virtual Screening Workflow, by selecting Compute maximum values by docking fragments in the Docking Options tab. The fragments are displayed in the Workspace with purple carbons, to distinguish them from the ligands. You can display the scores for the ligands relative to the maximum value for each descriptor by selecting Relative Scores. Positive values of the relative score mean that the ligand score does not achieve the maximum as obtained from the fragment set, and indicates room for improvement.

6.2.2 Controlling the Display

The normal Workspace controls to rotate, translate, and zoom in on the structure can be used. The Glide XP Visualizer panel provides some additional control over the display.

To change the view so that the ligands that are displayed occupy the entire Workspace, click the Fit to Ligand button.

EX2037.0090

To control what is displayed in the Workspace, you can use the Display options:

- Receptor—Display or undisplay the receptor. By default, the receptor is displayed. Only the residues within 5 Å of the ligand are displayed.
- Selected ligands—Display the ligands that are selected in the table. If this option is selected, the Workspace is updated whenever the selection of ligands is changed.

  Note: If multiple ligands are selected, changing the view might take a considerable amount of time. Unless it is necessary to view many ligands at once, it is recommended that this option is used infrequently.

- XP Waters—Waters are shown as small red spheres in the visualizations and can be displayed or hidden using this option. By default, waters are not displayed.
- Hydrophobic/philic map—Hydrophobic and hydrophilic surfaces can be shown in the visualizations by selecting this option. The first time this option is selected, the HPPMap job is run, and the button flashes until the job finishes. The surfaces are stored for subsequent use. The job takes 1–2 minutes.

6.2.3 Exporting XP Terms

You can export the information shown in the table to a CSV file. To do so, click Export. A file chooser opens, in which you can navigate to a location and name and save the file. If you want to export data for a selection of ligands, select them in the table and click Selected Only before clicking Export.

6.2.4 XP Terms and Their Visualizations

The XP terms and a description of their visualizations is given in Table 6.1. In the visualizations, hydrogen atoms are not generally displayed, and the ligand carbon atoms are colored green.

In addition to the XP terms, you can display similarities between the ligands in XP descriptor space, by selecting Similarities. The similarity is only displayed if a single ligand is selected in the table.

EX2037.0091

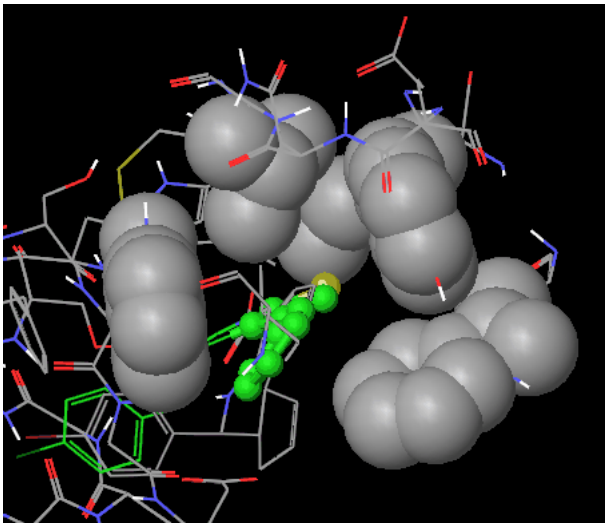


Figure 6.5. Visualization of hydrophobic enclosure terms

Table 6.1. Glide XP terms and their visualizations

XP Term	Description	Visualization
GScore	Total GlideScore; sum of XP terms	None
LipophilicEvdW	Lipophilic term derived from hydrophobic grid potential and fraction of the total protein-ligand vdW energy	None
PhobEn	Hydrophobic enclosure reward	Hydrophobic atoms on the protein that are necessary for recognition of hydrophobic enclosure are displayed in CPK representation in gray. Hydrophobic atoms on the ligand necessary for hydrophobic enclosure are displayed in green in ball and stick representation.
PhobEnHB	Reward for hydrophobically packed H-bond	H-bonds are displayed as pink dotted lines. Hydrophobic atoms on the protein that enclose hydrophobic groups on the ligand are displayed in gray in CPK representation. Hydrophobic atoms on the ligand are displayed in ball and stick representation.

Table 6.1. Glide XP terms and their visualizations (Continued)

XP Term	Description	Visualization
PhobEnPairHB	Reward for hydrophobically packed correlated H-bonds	H-bond are displayed as pink dotted lines. Hydrophobic atoms on the protein that enclose hydrophobic groups on the ligand are displayed in gray in CPK representation. Hydrophobic atoms on the ligand are displayed in ball and stick representation.
HBond	ChemScore H-bond pair term	H-bonds are displayed as yellow dotted lines.
Electro	Electrostatic rewards	Relevant atoms in the ligand are displayed in ball-and-stick representation.
SiteMap	SiteMap ligand/receptor non-H-bonding polar/hydrophobic and hydrophobic/hydrophilic complementarity terms	Hydrophobic (orange) and hydrophilic (turquoise) surfaces are displayed, representing the field of the receptor. These are the same as the Maestro Hppmap surfaces.
π Stack	Pi-pi stacking reward	The relevant aromatic groups in the protein are displayed in CPK representation and in the ligand in ball and stick representation.
π Cat	Reward for pi-cation interactions	The relevant aromatic groups in the protein are displayed in CPK representation and in the ligand in ball and stick representation.
ClBr	Reward for Cl or Br in a hydrophobic environment that pack against Asp or Glu	None
LowMW	Reward for ligands with low molecular weight	None
Penalties	Polar atom burial and desolvation penalties, and penalty for intra-ligand contacts	Atoms are labeled with the penalties. The labels and their descriptions and the associated visualizations are as follows: Charge+*value*—Charge penalty. Penalty for putting a ligand charge in a region without water and with the charge not in a salt bridge. Ball and stick representation for ligand atoms. Water+*value*—Water-protein and water-ligand penalties. Penalties for desolvating (via presence of ligand) a polar group on the protein or ligand and not making an H bond to the polar group. Ball and stick representation for ligand and protein atoms

EX2037.0093

Table 6.1. Glide XP terms and their visualizations (Continued)

XP Term	Description	Visualization
		DDpen+*value*—Donor-donor penalty. Penalty for non-H-bonded donors on the ligand and protein too close together and with semi-collinear NH--HN. Blk+*value*—Penalty for burial of charged group on the protein by the ligand with no H bonds made to the charged group by the ligand or protein. Ball and stick representation for protein atoms Polar+*value*—Penalty for desolvation of a polar ligand atom in a hydrophobic protein environment. Similar to water-ligand penalty but specific to phobic environments. Ball and stick representation for protein atoms Contact+*value*—Penalty for intraligand contacts. Ball and stick representation for ligand atoms PhoBl+*value*—Phobic ligand-protein blockage. Penalty for putting a phobic group of the ligand against polar (donor/acceptor) groups of the protein in a protein region which normally would give a favorable phobic packing score (PhobEn). Ball and stick representation for protein atoms Twisted amide linkages in the ligand are also displayed in ball-and-stick, but not labeled.
HBPenal	Penalty for ligands with large hydrophobic contacts and low H-bond scores	None
PhobicPenal	Penalty for exposed hydrophobic ligand groups	The molecular surface of the protein around the ligand is displayed as a semi-transparent surface. Hydrophobic ligand groups that are exposed are rendered in ball-and-stick representation
RotPenal	Rotatable bond penalty	Rotatable bonds in the ligand are displayed in tube representation.

Running Glide from the Command Line

This chapter contains information on running Glide from the command line. Although you will usually set up Glide jobs using the controls and settings in the Maestro GUI, you sometimes might want to submit jobs from the command line for the following reasons:

- The command-line scripts can run all full-featured jobs written using the Glide panels in Maestro, and also allow you to override specific run-time values that are not accessible through the Maestro interface.
- Command-line scripts allow you to run Glide jobs when you want.
- Input files and scripts can be modified and jobs can be re-run without reconfiguring and reloading job settings in Maestro. The quickest way to create the input files needed to run a job from the command line is to set up the job in its Maestro panel, click the panel's Write button, and modify the files that are generated.

Glide jobs are ultimately run using the Impact program. There are several commands that can be used to run Glide jobs:

- `glide`—This command reads a Glide input file, processes it and runs Impact. It is the preferred method for running Glide, and is the command used when you run Glide from Maestro. You can run both single-processor and distributed Glide docking jobs with this command.
- `impact`—This command runs Impact directly, and uses an Impact input file. You can produce an Impact (DICE) input file (`.inp`) by running the `glide` command, or you can use Impact input files from previous releases. You can run single-processor docking jobs as well as grid generation jobs with this command.
- `para_glide`—This command is used for distributed Glide docking jobs, and also uses an Impact input file.

Glide also has a three command-line utilities, kept in the `$SCHRODINGER/utilities` directory: `glide_sort`, `glide_merge`, and `glide_rescore`, which are also described in this chapter.

For information about general-purpose utilities, see Appendix D of the *Maestro User Manual*.

EX2037.0095

7.1 Running Jobs

To run jobs, the `SCHRODINGER` environment variable must be set to the Schrödinger software installation directory. You can define `SCHRODINGER` as follows:

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

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

Command-line applications are located in the main Schrödinger directory; command-line *utilities* are located in the directory `$SCHRODINGER/utilities`. For usage summary information on any application or utility, use the `-h` (help) option:

```
$SCHRODINGER/command -h
$SCHRODINGER/utilities/utilityname -h
```

Schrödinger applications and utilities run under Schrödinger's Job Control facility. To run jobs on remote hosts, you must set up a hosts file. For information on this file and other aspects of configuring hosts and Job Control, see Chapter 6 of the *Installation Guide*. You can force Impact to run outside Job Control, but you are then responsible for managing all the files.

Jobs are automatically run in the background. You need not add an `&` at the end of the commands to have them run and immediately display the command prompt. The `-WAIT` option forces the shell to wait until the job is finished, so you can embed such commands in other scripts.

Once your jobs are launched, you can monitor their progress using the Monitor panel in Maestro. The command `$SCHRODINGER/jobcontrol` can also be used. It has many options, but the two most useful options are:

```
$SCHRODINGER/jobcontrol -list
```

which shows the status of all your jobs, and:

```
$SCHRODINGER/jobcontrol -kill jobid
```

to terminate any job and its subjobs, if any exist.

For a summary of `jobcontrol` options, use:

```
$SCHRODINGER/jobcontrol -h
```

For more information, see the *Job Control Guide*.

EX2037.0096

7.2 Job Files and Directories

Location of Files and Working Directory

For Glide jobs, Maestro normally writes input files to the directory from which you launched Maestro (the *Maestro working directory*). Impact also normally writes its output files to the same location, though Impact input files and the Glide interface allow you to specify an arbitrary location for grid files.

File Name Conventions

A typical job has one command-script file (*jobname*`.in` or *jobname*`.inp`), one or more structure files (*jobname*`.mae`, *jobname*`.pdb`, or *jobname*`.sdf`), and after execution, several output files (e.g., *jobname*`_out.mae` for structure files and *jobname*`.out` for textual data).

If a file already has the name of an output file, in many cases Impact renames the old file with a numerical extension (*filename*`.out.01`, *filename*`.out.02`, and so on) for archival purposes. The new job's output is then written to the base name (*filename*`.out`). If you do not need the old files, you can remove them.

Some files, such as *jobname*`.log` files, are newly written each time Impact runs a calculation. Likewise, old *jobname*`_pv.mae` files are overwritten.

In addition, *jobname*`_out.mae` files are *not* produced by default for Glide jobs—Glide writes intermediate Maestro-format structure output to *jobname*`_raw.mae` files, which are incremented.

Table 7.1 contains descriptions of the various files generated during a Glide docking run. For more information, see the Maestro online help or the *Impact Command Reference Manual*.

Table 7.1. Input and output files for Glide docking jobs.

File	Description
jobname`.in`	Glide input file, required as input to the `glide` command. Glide input files are formatted plain-text files written in the Glide simplified input format. Maestro creates Glide input files for job submission. You can create or edit them manually with a text editor.
jobname`.inp`	Impact input file, required as input to the `impact` command. Impact input files are formatted plain-text files written in the Impact input file language, DICE. Maestro can write Impact input files when you click Write in a Glide panel, but by default writes Glide input files. You can create or edit them manually with a text editor.
jobname`.mae`	Maestro format file containing input structures for docking.

EX2037.0097

Table 7.1. Input and output files for Glide docking jobs. (Continued)

File	Description
jobname.log	Impact log file. If specified, a .log file captures standard output and standard error messages in text form. This file is overwritten during subsequent runs.
jobname.out	Impact output file, containing output from standard output but not standard error. Output files are appended with numerical extensions when the input file is used again. Up to 99 output files are retained.
jobname.rept	Plain text file containing a table of ranked poses, scores, and score components.
*jobname*_lib.mae	Glide ligand structure output file, in Maestro format.
*jobname*_pv.mae	Glide pose-viewer file, in Maestro format. Contains the receptor structure, followed by all output ligand poses ranked by their score.
*jobname*_raw.mae	Glide's intermediate structure files, containing unranked ligand poses.

7.3 Receptor Grid Generation and Ligand Docking

Using Maestro is the best way to write Glide input files, even if you intend to run them from the command line, because the files are syntactically correct. The three possible commands for running Glide jobs are described in the subsections below.

7.3.1 The glide Command

The glide command is the preferred way to run Glide jobs. This command can run both single processor and distributed docking jobs, as well as grid generation jobs. The syntax of the command is as follows:

$SCHRODINGER/glide [*options*] *input-file*

The glide command supports the standard Job Control options, which are described in Table 2.1 of the *Job Control Guide*, and the -WAIT and -LOCAL options described in Table 2.2 of the *Job Control Guide*. The options for the glide command are given in Table 7.2.

Table 7.2. Options for the glide command.

Option	Description
-v	Display the version number and exit.
-h[elp]	Print usage message and exit.
-no_cleanup	Do not clean up after the job finishes.
-strict	Terminate upon first subjob failure.

EX2037.0098

Table 7.2. Options for the glide command.

Option	Description
`-write_inp`	Write the Impact `.inp` file and exit (do not run jobs).
`-NJOBS` *njobs*	Number of subjobs to create. If greater than 1, a driver job is submitted that manages the subjobs. The value of *njobs* must be greater than the number of CPUs requested for the job; if it is not, it is set to the number of CPUs. This option is only available for docking jobs. Default: 1.
`-DRIVERHOST` *host*	Run the driver job on *host*, if multiple subjobs are requested. Default: first host specified with the `-HOST` option.
`-NOJOBID`	Run job outside of Job Control. If *njobs* > 1, only applies to the driver job.
`-LOCAL`	Run in local directory (default for localhost jobs).
`-NOLOCAL`	Run in scratch directory (default for remote jobs).
`-RESTART`	Enables driver script to restart. Command must be run from the host on which the driver was run.
`-SUBLOCAL`	Do not create a temporary directory for each subjob. NOTE: will not work if using compressed grids.
`-OVERWRITE`	Remove previous job files before running.

The input file is a Glide simplified input file, whose default extension is `.in`. This file contains keyword-value pairs that define the parameters of the calculation. These keywords are listed in the tables below. Table 7.3 contains keywords for grid generation jobs, and Table 7.4 contains keywords for ligand docking jobs. The keywords for grid generation and ligand docking are grouped according to the tabs (and associated Advanced Settings dialog boxes) in the Receptor Grid Generation and Ligand Docking panels in which they are set. Defaults are available for many of the keywords, so you do not need to include all the keywords in the input file.

Boolean values in the input file can be represented by `True` or `False`, `Yes` or `No`, `1` or `0`. The last two pairs are converted internally to `True` or `False`. The values are case-insensitive. A Boolean value is denoted by *boolean* in the tables below.

Table 7.3. Glide input file keywords for grid generation jobs.

Keyword	Description
`GRIDFILE` *filename*	Filename of the grid `.zip` file, if compressed, or the `.grd` file.
`DIELECTRIC` *value*	Dielectric constant.
`CV_CUTOFF` *value*	Coulomb-van der Waals cutoff.

EX2037.0099

Table 7.3. Glide input file keywords for grid generation jobs.

Keyword	Description
`GRIDLIG` *boolean*	Cocrystallized ligand has been identified if set to `True`. Default: `False`.
Receptor keywords	
`RECEP_FILE` *filename*	Receptor filename.
`RECEP_VSCALE` *value*	General van der Waals radius scaling factor. Default: 1.0.
`RECEP_CCUT` *value*	General van der Waals radius scaling partial charge cutoff. Default: 0.25.
`GLIDE_RECEP_MAESCALE` *boolean*	Read per-atom scale factors from the input structure file if set to `True`. Default: `False`.
`GLIDE_RECEP_ASLSCALE` *boolean*	Specify per-atom scale factors by ASL if set to `True`. Default: `False`.
`ASLSTRINGS` *list*	Comma-separated list of ASL expressions, one expression for each set of residues to be scaled. Only set if `GLIDE_RECEP_ASLSCALE` is set to `True`.
`GLIDERECEPTORSCALECHARGES` *list*	Comma-separated list of charge scaling factors, one factor for each set of residues to be scaled.
`GLIDERECEPTORSCALERADII` *list*	Comma-separated list of van der Waals scaling factors, one factor for each set of residues to be scaled.
`REC_MAECHARGES` *boolean*	Use partial charges from the input structure if set to `True`. Default: `False`.
Site keywords	
`INNERBOX` *x,y,z*	Comma-separated list of the *x*, *y*, and *z* dimensions of the inner (ligand centroid bounding) box in angstroms.
`ACTXRANGE` *value*	Outer box *x* dimension. Must be set to the same as for the `OUTERBOX` keyword.
`ACTYRANGE` *value*	Outer box *y* dimension. Must be set to the same as for the `OUTERBOX` keyword.
`ACTZRANGE` *value*	Outer box *z* dimension. Must be set to the same as for the `OUTERBOX` keyword.
`OUTERBOX` *x,y,z*	Comma-separated list of the *x*, *y*, and *z* dimensions of the outer (enclosing) box in angstroms.
`GRID_CENTER` *x,y,z*	Comma-separated list of the *x*, *y*, and *z* coordinates of the center of the grid.

Table 7.3. Glide input file keywords for grid generation jobs.

Keyword	Description
Constraints keywords	
`GLIDECONS` *boolean*	Define constraints if set to `True`. Default: `False`.
`GLIDE_NTOTALCONS` *ncons*	Total number of constraints defined (Positional+H-bond/ Metal+Hydrophobic).
`GLIDE_NUMPHOBCONS` *nhpc*	Number of hydrophobic constraints defined.
`GLIDE_NUMPOSITCONS` *npc*	Number of positional constraints defined.
`GLIDECONSATOMS` *list*	Comma-separated list of atom numbers that define H-bond/ Metal constraints.
`GLIDECONSNAMES` *list*	Comma-separated list of constraint names.

Table 7.4. Glide input file keywords for ligand docking jobs.

Keyword	Description
`GRIDFILE` *filename*	Filename of the grid `.zip` file, if compressed, or the `.grd` file.
Settings keywords	
`PRECISION {XP|SP|HTVS}`	Docking precision mode. Set in the Settings tab.
`DOCKING_METHOD {confgen|rigid|mininplace|inplace}`	Docking method: `confgen`—Dock flexibly. `rigid`—Dock rigidly. `mininplace`—Refine (do not dock) `inplace`—Score in place (do not dock)
`SAMPLE_RINGS` *boolean*	Sample rings if set to `True` and `DOCKING_METHOD` is set to `confgen`.
`AMIDE_MODE {penal|free|fixed|trans}`	Amide bond sampling mode. Only set when `DOCKING_METHOD` is set to `confgen`. `penal`—penalize nonplanar conformation `free`—vary conformation `fixed`—retain original conformation `trans`—allow trans conformation only
`EPIK_PENALTIES` *boolean*	Apply penalties for ionization or tautomeric states calculated by Epik. Default: `False`.
`EXPANDED_SAMPLING` *boolean*	Expand the sampling by bypassing the elimination of poses in the rough scoring stage. Useful for fragment docking. Default: `False`.
`WRITE_XP_DESC` *boolean*	Write XP descriptor information if set to `True`. Default: `False`.

EX2037.0101

Table 7.4. Glide input file keywords for ligand docking jobs.

Keyword	Description
`FLEXTORS` *boolean*	Set to `False` for `rigid`, `refineinput` and `inplace` jobs.
`MAXKEEP` *nposes*	Number of poses per ligand to keep in initial phase of docking. Default: 5000.
`SCORING_CUTOFF` *cutoff*	Scoring window for keeping initial poses. Default: 100.0.
`MAXREF` *nposes*	Number of poses to keep per ligand for energy minimization. Maestro sets this number to 800 for XP. Default: 400.
`DIELECTRIC` *constant*	Distance-dependent dielectric constant. Default: 2.0.
`FORCEFIELD {OPLS2001|OPLS2005}`	Specify the force field to use. Glide is parametrized against the OPLS_2001 force field. Default: `OPLS2001`.
`MAX_ITERATIONS` *niter*	Maximum number of conjugate gradient steps. Default:100.
Ligands keywords	
`GLIDECONSUSEMET` *boolean*	Use a table of individual metal radii in the restraining potential for metal-ligand constraints, rather than a single global value. Default: `False`.
`GLIDESCORUSEMET` *boolean*	Use a table of individual metal radii in the scoring function, rather than a single global value. Default: `False`.
`LIGANDFILE` *filename*	File name of ligands file. File format can be Maestro, SD, MOL2 or PDB.
`LIG_MAECHARGES` *boolean*	Take charges from input structures if set to `True`. Default: `False`.
`MAXATOMS` *maxatoms*	Maximum number of atoms per ligand. Default (and maximum): 300.
`MAXROTBONDS` *maxrotbonds*	Maximum number of rotatable bonds per ligand. Default (and maximum): 50
`LIG_VSCALE` *factor*	Scaling factor for van der Waals radii scaling. Default: 0.8.
`LIG_CCUT` *cutoff*	Partial charge cutoff for van der Waals radii scaling. Default: 0.15.
`RINGCONFCUT` *cutoff*	Energy window for retention of ring conformers. Default: 2.5 kcal mol^{-1}.
Core keywords	
`USE_REF_LIGAND` *boolean*	Use core pattern for comparison or restraint if set to `True`. Default: `False`.
`REF_LIGAND_FILE` *filename*	File name for reference ligand. File format can be Maestro, SD, MOL2 or PDB.

Table 7.4. Glide input file keywords for ligand docking jobs.

Keyword	Description
`REF_LIGAND_SOURCE {ct|extfile}`	Define source of reference ligand. `ct`— `extfile`—Take reference ligand from the external file defined by the `REF_LIGAND_FILE` keyword.
`CORE_DEFINITION {all|smarts|allheavy}`	Specify how the core is defined. `all`—use all atoms in the reference ligand `allheavy`—use all nonhydrogen atoms in the reference ligand `smarts`—use the SMARTS pattern defined by the `CORE_ATOM` keyword.
`CORE_ATOM` *list*	List of core atoms to use for the RMSD calculation. Only set when `CORE_DEFINITION` is set to `smarts`.
`CORE_SMARTS` *pattern*	Smarts pattern for core. Only set when `CORE_DEFINITION` is set to `smarts`. Set in the Core tab.
`CORE_RESTRAIN` *boolean*	Restrict docking to reference position if set to `True`. Default: `False`.
`CORE_POS_MAX_RMSD` *value*	Tolerance on the RMSD to use when restricting docking to the reference position.
`CORE_FILTER` *boolean*	Skip ligands that do not match the core pattern if set to `True`. Default: `False`.
Constraints keywords	
`HAVEGLIDECONSFEAT` *boolean*	A feature file is already available if set to `True`. Maestro writes the feature file and sets this keyword to `True`. Other applications such as VSW set this keyword to `False` because the feature file is generated as part of the workflow. Default: `False`.
`GLIDECONS` *boolean*	Use constraints if set to `True`. Default: `False`.
`GLIDE_CONS_FINALONLY` *boolean*	Test constraint satisfaction only after docking if set to `True`. Default: `False`.
`GLIDE_CONS_FEAT_FILE` *filename*	Specify the name of the constraints feature file. Default: *jobname*`.feat`.
`GLIDEUSECONSFEAT` *boolean*	Use the constraints feature file for the application of constraints if set to `True`. If set to `False`, the constraint information is in the Impact `.inp` file. This keyword should be set to `True` for all jobs.
Similarity keywords	
`SIMIL_WGT_SCORE` *boolean*	Use trained similarity model if set to `True`. Default: `False`.
`SIMIL_PENALTY`	Maximum GlideScore penalty. Default: 6.0 kcal/mol.

EX2037.0103

Table 7.4. Glide input file keywords for ligand docking jobs.

Keyword	Description
`SIMIL_LOW` *lowsimil*	Similarity value for the start of the penalty ramp. Default: 0.3.
`SIMIL_HIGH` *hisimil*	Similarity value for the end of the penalty ramp. Default: 0.7.
`SIMIL_REJECT` *cutoff*	Rejection cutoff. Ligands with similarity below this cutoff are rejected. Default: 0.0.
`SIMIL_MODE {none|simil| dissimil}`	Similarity mode. `none`—do not use similarity `simil`—find similar ligands `dissimil`—find dissimilar ligands
`SIMIL_WGT_FILE` *filename*	Name of trained similarity model file (`.wgt`).
`SIMIL_SCORE_FILE` *filename*	Name of file containing known actives. The default is a Maestro file, compressed or uncompressed. For SD or mol2 files, use the `SIMIL_SCORE_FMT` keyword to select the file type.
`SIMIL_SCORE_FMT {sd| mol2}`	Format of the known actives file if it is not a Maestro file.
Output keywords	
`COMPRESS_POSES {TRUE|FALSE}`	Compress the output pose files. Default: `TRUE`.
`CV_CUTOFF` *cutoff*	Reject poses with Coulomb-van der Waals energy greater than *cutoff* kcal/mol. Default: 0.0.
`HBOND_CUTOFF` *cutoff*	Reject poses with H-bond score greater than *cutoff*. Default: 0.0.
`METAL_CUTOFF` *cutoff*	Reject poses with metal score greater than *cutoff*. Default: 10.0.
`POSE_RMSD`	RMS deviation used in clustering to discard poses, in angstroms. Default: 0.5.
`POSE_OUTTYPE {poseviewer|ligandlib}`	Type of output file to produce. Default: `poseviewer`. `poseviewer`—Write pose viewer file (`_pv.mae`); includes receptor as first structure. `ligandlib`—Write ligand pose file (`_lib.mae`); does not contain receptor.
`POSE_DISPLACEMENT`	Maximum atomic displacement used in clustering to discard poses, in angstroms. Default: 1.3.
`NREPORT` *maxposes*	Maximum number of poses to write out from the docking run. Default: 10000.
`POSES_PER_LIG` *maxperlig*	Maximum number of poses to write per ligand. Default: 1.
`POSTDOCK` *boolean*	Use post docking minimization if set to `True`. Default: `True`.

EX2037.0104

Table 7.4. Glide input file keywords for ligand docking jobs.

Keyword	Description
`POSTDOCK_NPOSE`	Number of poses to use in post-docking minimization. Maestro sets this number to 10 for XP. Default: 5.
`POSTDOCKSTRAIN` *boolean*	Apply strain correction terms if set to `True`. Default: `False`.
`WRITE_RES_INTERACTION` *boolean*	Generate per-residue interaction terms if set to `True`. Default: `False`.
`RADIUS_RES_INTERACTION` *r*	Radius for per-residue interaction terms, in angstroms. Only set if `WRITE_RES_INTERACTION` is set to `True`. Default: 12.0
`WRITEREPT` *boolean*	Write report text file (`.rept` for docking and `.scor` for score-in-place). Default: `False`.
`CALC_INPUT_RMS` *boolean*	Calculate RMSD between each pose and its input structure. Default: `False`.

7.3.2 The impact Command

The `impact` options that you can specify when initiating Glide receptor grid generation and ligand docking jobs from the command line are described in Table 7.5. This list also includes `impact` keywords with more general applications. The `impact` command also supports the standard Job Control options, which are described in Section 2.3 of the *Job Control Guide*.The syntax of the command is as follows:

```
$SCHRODINGER/impact [options] [[-i] input-file]
```

If your protein has more than 8000 atoms and you are making grid files (but not if you are just docking ligands), you will also need to include the `-s huge` option in the command. If you have a large number of ligands, you can use `para_glide` to distribute the docking job over multiple processors.

The `impact` command is also used to run other applications. For more information on the input file syntax, see the *Impact Command Reference Manual*.

7.3.3 The para_glide Command

The para_glide application divides the ligands into batches of ligand structures and submits these batches to multiple processors for docking. The syntax of the `para_glide` command is:

```
$SCHRODINGER/para_glide -i input-file [options]
```


EX2037.0105

Table 7.5. `Impact` *command options*

Option	Description
`-h`	Print usage summary and exit
`-i` *input-file*	Impact input file, conventionally ending in `.inp`. If the *input-file* argument does not end in `.inp`, Impact looks first for *input-file* as specified. If that file doesn't exist, it then looks for *input-file*`.inp`. If `-i` is omitted, then *input-file* **must** end in `.inp` and must be the last argument in the command line.
`-o` *output-file*	File for writing output and log messages. If this option is omitted, the log file is named *jobname*`.log`, where *jobname* is taken from the Impact input file name.
`-s` *size*	Use specific "size" version of the Impact executable. Acceptable values for *size* are `medium` or `huge`. If omitted, `medium` is assumed in most cases; it is valid for up to 8000 atoms or 8000 bonds. The `huge` option accommodates up to 40000 atoms or 40000 bonds.
`-v`	Print Impact version information and exit
`-LOCAL`	Force remote jobs to write files to the job submission directory, rather than on the remote host. Only active when `-HOST` is used.
`-NO_REDIRECT`	Do not run `impact` under Job Control. If you use this option, you must ensure that your input and output files are in the appropriate location, because none of the functions of Job Control are performed.
`-WAIT`	Do not return control to the shell until the job finishes. This is useful in command scripts in which the Impact job must finish before further action is taken.

The options used by `para_glide` are given in Table 7.6. The standard Job Control options, as described in Section 2.3 of the *Job Control Guide*, are also accepted. You should use the `-HOST` option to specify the hosts over which the subjobs are to be distributed. All other options, such as options for the `impact` command, are passed to the subjobs.

The `para_glide` application splits the Glide job specified in the *input-file* into smaller subjobs for distributed execution. The ligands from *firstlig* to *lastlig* are separated into *njobs* equal-sized batches. By default, *njobs* is set to 1, *firstlig* is set to 1 and *lastlig* is set to 0, which is interpreted as the final ligand in the ligands file.

For example, to submit a job with 50 subjobs to be run on 10 processors, you could use the following command:

```
$SCHRODINGER/para_glide -n 50 -HOST mycluster:10 -i myjob.inp
```

This job would run on the host `mycluster` using 10 processors.

Table 7.6. Options for the para_glide command.

Option	Description
`-f` *firstlig*	First ligand to include. Default: 1.
`-l` *lastlig*	Last ligand to include. Default: 0, meaning the last ligand in the file.
`-n` *njobs*	Number of subjobs to prepare. Default: 1.
`-j` *jobname*	Name of job. Determines names of `para_glide` log, rept, and final pose files. The default is taken from the input file name.
`-job` *jobnum*	Subjob number to prepare. A value of 0 means "prepare all subjobs".
`-nosplit`	Do not split input ligand file by subjob. This option should not be used if the job is submitted to a grid, because separate ligand files must be distributed to the hosts on the grid.
`-strict`	Terminate `para_glide` job upon first subjob failure.
`-force`	Run postprocessing script to generate final pose output, even if subjobs fail.
`-h`	Print help message and quit.
`-v`	Print version number and quit.
`-w`	Write input files, including files for subjobs, but do not launch jobs.

Each use of `para_glide` creates two scripts: *jobname*`_report.sh` and *jobname*`_status.sh`. The *jobname*`_report.sh` script collects the output (poses) from subjobs created by `para_glide`, and produces a single pose file and a single report file that summarizes the best poses in the entire job suite. The subjob results are stored in subdirectories of the working directory, named *jobname_firstlig_lastlig*, where *firstlig* and *lastlig* are the indexes of the first and last ligands in the subjob. The *jobname*`_status` script can be run at any time to obtain a summary of the disposition of each job: whether it finished normally, died, was terminated, stopped, and so on, using the conventional Schrödinger job control terms. (See Chapter 3 of the *Job Control Guide* for information about job monitoring and job control.)

The `-job` option is useful for preparing only a single subjob. A value of 0 is equivalent to not specifying `-job` at all, and all subjobs will be prepared. Negative values are not permitted.

By default, a new ligand file is written for each subjob, containing only the ligands for that subjob. If the `-nosplit` option is given, all jobs use the same input ligand file, and each job reads out of it just the ligands pertinent to that job. Splitting the ligand file is likely to be more efficient if there is a large number of ligands, but requires extra disk space to store the ligand files.

EX2037.0107

7.3.4 Restarting Glide Docking Jobs

Although Glide does not have an automatic restart feature for jobs that fail, the Job Control facility can automatically restart failed Glide subjobs that were run with the `glide` command. The number of times a failed subjob is retried is set by the `SCHRODINGER_MAX_RETRIES` environment variable, whose default is 3. This environment variable is described in Table B.1 of the *Job Control Guide*. If you still have failed subjobs when the master job finishes, you can restart the job to run the subjobs that failed. To do so, use the command that you used to run the job originally with the `-RESTART` option added. This restart mechanism only applies to distributed jobs.

For jobs run with `impact` or `para_glide`, you can reclaim the work that has been done to date and rerun jobs for the ligands that have not yet been docked. The results to the point of failure are stored in a `_raw.mae` file, which can be used in conjunction with the results of a new docking run that completes the task to generate the final output file, using `glide_sort`. There are several scenarios to consider.

If you ran a single-processor job:

1. Examine the log file to obtain the number of the ligand being docked when docking failed.
2. Copy the log file to a safe location.
3. Start a `para_glide` job, specifying the first ligand to be docked as the one that failed.

   You can run a `para_glide` job on a single processor. The results will be in a subdirectory created by `para_glide`.

If you ran a para_glide job and isolated subjobs failed or didn't run:

1. Change to the subdirectory for each subjob that failed or didn't run in turn, and start an `impact` job, with the files for the subjob as input. The output is placed in the same subdirectory as the job was started from.

If you ran a para_glide job and a contiguous block of subjobs failed or didn't run:

1. Determine which subjobs ran from the log files.
2. Run a `para_glide` job for the block of subjobs that didn't run.

For subjobs that failed but produced some results, you can rerun them from the point of failure with `para_glide`. This procedure creates another subdirectory for the results.

When all the restarted jobs finish, run a `glide_sort` job from the parent directory, specifying all the `_raw.mae` files as input.

EX2037.0108

7.4 Glide Utilities

7.4.1 glide_sort

Re-ranks Glide poses by custom criteria or combines job outputs into one file. The "best" pose is defined by the property used for sorting. If you set any of the custom scoring function terms, sorting is done only on this custom function instead of the standard sorting options.

Syntax:

`$SCHRODINGER/utilities/glide_sort` *mode* `[`*options*`]` *Glide-pose-files*

The input file extensions can be `.mae` for uncompressed files, or `.maegz` or `.mae.gz` for compressed files.

At least one of these modes is required:

`-o` *output-file*	Write the best-scoring poses to *output-file*. The file extension can be `.mae` for uncompressed output, or `.maegz` or `.mae.gz` for compressed output.
`-r` *report-file*	Create a report of the best scores in *report-file*.
`-R`	Write a report of the best scores to standard output.

The options for `glide_sort` are given in Table 7.7, categorized by function. If multiple `-use_prop[_a|d]` options are given, a multi-key sort is used, with the first property specified as the primary key, the second property as the secondary key, and so on.

Table 7.7. Options for the glide_sort utility

Option	Description
`-f` *filelist*	Specify Glide pose files in the plain text file *filelist*, one file name per line.
Sorting Options:	
`-use_dscore`	Default. Sort poses based on the "docking score" in Glide output poses. "Docking score" is a placeholder for the property on which you would like to sort poses. Initially it is equal to GlideScore, but `glide_rescore` can be used to replace it with other values (see Section 7.4.3.)
`-use_gscore`	Sort poses based on GlideScore. Overrides `-use_dscore`.
`-use_cvdw`	Sort poses based on Coulomb-van der Waals energy, E(CvdW). Overrides `-use_dscore`.
`-use_emodel`	Sort poses based on model energy score, Emodel. Overrides `-use_dscore`.

EX2037.0109

Table 7.7. Options for the glide_sort utility (Continued)

Option	Description
`-use_prop` *prop* | `-use_prop_a` *prop*	Sort poses based on the specified user-defined numerical property, from lowest to highest values (ascending order). Overrides `-use_dscore`.
`-use_prop_d` *prop*	Sort poses based on the specified user-defined numerical property, from highest to lowest values (descending order). Overrides `-use_dscore`.
`-nosort`	Don't sort the poses.
Output Options:	
`-n` *nreport*	Retain only the *nreport* lowest-scoring poses.
`-norecep`	Don't include the receptor structure in output files.
`-noxpdes`	Do not create an output XP descriptor (`.xpdes`) file.
`-best`	Keep only the single best pose for each ligand in each input pose file.
`-best-by-lignum`	Keep only the single best pose for each ligand with a given *lignum* value.
`-best-by-title`	Keep only the single best pose for each ligand with a given title.
`-best-by-prop` *prop*	Keep only the single best pose for each ligand with a given value of the specified property *prop*.
`-h`	Print help message and exit.
`-v`	Print version number and exit.
Custom Scoring Function Options:	
`-gscore` *coef*	GlideScore coefficient for custom scoring function. Default: 0.0.
`-cvdw` *coef*	E(CvdW) coefficient for custom scoring function. Default: 0.0.
`-internal` *coef*	E(internal) coefficient for custom scoring function. Default: 0.0.
`-emodel` *coef*	Emodel coefficient for custom scoring function. Default: 0.0.
`-offset` *coef*	Energy offset for custom scoring function. Default: 0.0.
Filter Cutoff Options:	
`-hbond_cut` *cutoff*	Filter cutoff for H-bond energy. Default: 0.0.
`-cvdw_cut` *cutoff*	Filter cutoff for E(Cvdw). Default: 0.0.
`-metal_cut` *cutoff*	Filter cutoff for metal-ligation energy. Default: 10.0.
`-emodel_cut` *cutoff*	Filter cutoff for Emodel (not used by default).
`-nofilter`	Don't use filter cutoffs at all.

7.4.2 glide_merge

This utility merges results from a list of pose files (such as from `para_glide` subjobs) into a single pose file, sorted by GlideScore. The pose files must be already sorted by GlideScore. The syntax of the command is as follows:

`$SCHRODINGER/utilities/glide_merge` *mode* [*options*] *Glide-pose-files*

The input file extensions can be `.mae` for uncompressed files, or `.maegz` or `.mae.gz` for compressed files.

At least one of these modes is required:

`-o` *output-file*	Write the best-scoring poses to *output-file*. The file extension can be `.mae` for uncompressed output, or `.maegz` or `.mae.gz` for compressed output. If `.xpdes` files exist for the input pose files, an output `.xpdes` file is also created with the same base name.
`-r` *report-file*	Create a report of the best scores in *report-file*.
`-R`	Write a report of the best scores to standard output.

The options for `glide_merge` are given in Table 7.8.

Table 7.8. Options for the glide_merge utility

Option	Description
`-b` *batch-size*	Number of files to merge at a time. Equal to the number of file handles open at once. If the batch size is less than the number of pose files to merge, then the merges will be done iteratively. Default: 128.
`-c` *cutoff*	GlideScore cutoff. Poses with GlideScore greater than or equal to this cut-off are not included in the merged output. Default: 10000.0.
`-f` *file-list*	File containing list of files to merge, one file name per line. This option can be used as a supplement to or replacement for specifying the pose files in the argument list.
`-h`	Print help message and exit.
`-n` *nreport*	Retain only the *nreport* lowest-scoring poses.
`-noxpdes`	Do not create an output XP descriptor (`.xpdes`) file. By default, a matching XP descriptor file is created if the input XP descriptor files exist.
`-reset_lignum`	Reset the ligand numbers, starting from 1 for the first ligand in the output file. The ligands will have unique numbers in the output file. By default, the ligand numbers are taken from the input pose files.

EX2037.0111

7.4.3 glide_rescore

This utility replaces the docking score properties in Glide pose output files with different values, so that the `glide_sort` "best-by-title" option can be used to combine different screens. See Section 7.4.1. The syntax of the command is as follows:

`$SCHRODINGER/utilities/glide_rescore [`*options*`]` *pv-or-lib-files*

The options are given in Table 7.9. Only one of `-rank`, `-offset`, and `-average` can be used at a time. If none is specified, `-rank` is assumed. Using `-top` or `-every` implies `-average` mode.

Table 7.9. Options for the glide_rescore utility.

Option	Description
`-rank`	Replace "docking_score" with ligand rank (default mode).
`-offset` *value*	Replace "docking_score" with GlideScore plus this offset.
`-average`	Calculate the average GlideScore over all the poses; no output besides this average is produced.
`-top` *number*	Average only the top *number* poses.
`-every` *number*	Print running averages every multiple of *number* poses.
`-o` *output-file*	Output to this file name, instead of default name (*input-file*`.rank.mae` or *input-file*`.offset.mae`).
`-h`	Print help message and quit.
`-v`	Print version number and quit.

7.5 Examining Results From the Command Line

While Maestro provides facilities in the Project Table and the Pose Viewer interface to visualize high-scoring poses, you can also see the numerical results for these poses in the *jobname*`.rept` output file. A score-in-place calculation writes a *jobname*`.scor` report file instead, and no structural output for the Pose Viewer. The results are stored as Maestro properties in the Pose Viewer file (*jobname*`_pv.mae`) or ligand database file (*jobname*`_lib.mae`), and can be displayed in the Maestro Project Table.

7.6 Customizing Glide Calculations

The controls in the Glide panels provide a limited number of options for customizing the docking calculations. You can make more extensive customizations by adding commands and keywords to the Impact input file. The syntax of this input file is described in detail in the *Impact Command Reference Manual*. You cannot make these customizations with the Glide input file.

7.6.1 Changing the Glide Atom Typing

Glide uses SMARTS patterns to define the atom types ("ptypes") and hybrid types used in grid generation and docking. By default, these patterns are read from `$SCHRODINGER/impact-v`*version*`/data/opls/ptype.def`. If you want to change the atom types, you can copy and edit this file, and then add the appropriate keywords under the `PARAMETER` subtask in the input file. There are two keywords: `ptype`, which specifies the file to be used and must be followed by the file name, and `prntptypes`, which prints out the ptypes and the hybrid types.

To customize a grid generation job, an additional parameter statement should be included at the beginning of the `DOCK` task:

`parameter prntptypes ptype "`*filename*`"`

For docking jobs, during the first `DOCK` task when reading the grid files, include:

`parameter prntptypes ptype "`*filename*`"`

The quotes are required to preserve the case of the file name.

7.7 Docking Log Messages

When a ligand is docked, its fate is reported in messages in the log file. These messages are described below.

- `DOCKING RESULTS FOR LIGAND` *n*—Docking succeeded and at least one pose was written to the `_raw.mae` file.
- `NO POSES STORED`—Nothing was written to the `_raw.mae` file for this ligand. This usually only happens in constrained jobs, if none of the minimized poses satisfies constraints.
- `NO POSES SURVIVED ROUGH-SCORE`—There are actually several messages similar to this (one of which mentions "refinement" as well as "rough-score"). For all of them, poses have been eliminated somewhere in the rough-score part of the funnel.

EX2037.0113

- `NO GOOD POSES FOUND`—This message should occur in conjunction with one of the two previous messages.
- `PROBLEM IN CONFGEN`—Problems occured in the conformation generator.
- `INVALID ATOMIC NUMBER`—An atom in the ligand has an atomic number for which no atom type can be assigned.
- `TOO MANY ROTATABLE BONDS`—The user-specified limit on the number of rotatable bonds was exceeded.
- `LIGAND TOO BIG`—The ligand has more atoms than either the absolute maximum allowed (200 atoms) or the maximum set by the user.
- `FAILED INTERCONSTRAINT DIST`—More than one constraint was required and the atoms that could satisfy each constraint individually are in positions that cannot possibly satisfy the constraints simultaneously.
- `TOO MANY ROTAMERS`—More conformations were generated for a given rotatable group than the constraint algorithm can handle. (The limit is 6.)
- `HYDROPHOBIC/PERIPHERAL ASSIGNMENT FAILED`—Assignment of hydrophobic regions on the ligand for application of hydrophobic constraints failed.
- `ATOMIC COORDINATES OVERLAP`—Ligand atoms are too close to each other. This message might indicate poor ligand preparation.
- `STRUCTURE INCOMPATIBLE WITH FORCE FIELD`—The ligand cannot be used with the force field selected because the bonding patterns could not be assigned.
- `STRUCTURE HAS TOO MANY ATOMS`—The ligand has more atoms than Impact can handle. (The limit is 300.)
- `SKIP FROM SIMILARITY`—The ligand was skipped because it did not meet the similarity criterion.
- `INVALID ATOM(S)`—The ligand has atoms for which no force field parameters are available.

References

1. Eldridge, M. D.; Murray, C. W.; Auton., T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997,** *11,* 425–445.

2. Maggiora, G. M.; Johnson, M. A. *Concepts and Applications of Molecular Similarity.* Wiley & Sons: New York, 1990, pp 99–117.

3. Moreau, G.; Broto, P. The Autocorrelation of a Topological Structure: A New Molecular Descriptor. *Nouv. J. Chim.* **1980,** *4,* 359–360.

4. Carhart, R. E.; Smith, D. H.; Venkataraghavan, R. Atom Pairs as Molecular Features in Structure-Activity Studies: Definitions and Applications. *J. Chem. Inf. Comput. Sci.* **1985,** *25,* 65–73.

5. Wagener, M.; Sadowski, J.; Gasteiger, J. Autocorrelation of Molecular Surface Properties for Modeling Corticosteroid Binding Globulin and Cytosolic Ah Receptor Activity by Neural Networks. *J. Am. Chem. Soc.* **1995,** *117,* 7769–7775.

6. Sheridan, R. P.; Miller, M. D.; Underwood, D. J.; Kearsley, S. K. Chemical Similarity Using Geometric Atom Pair Descriptors. *J. Chem. Inf. Comput. Sci.* **1996,** *36,* 128–136.

7. Kearsley, S. K.; Sallmack, S.; Fluder, E. M.; Andose, J. D.; Mosley, R. T.; Sheridan, R. P. Chemical Similarity Using Physicochemical Property Descriptors. *J. Chem. Inf. Comput. Sci.* **1996,** *36,* 118–127.

8. Pearlman, R. S.; Smith, K. M. Novel Software Tools for Chemical Diversity. *Perspectives Drug Discovery Design* **1998,** *9,* 339–353.

EX2037.0115

Getting Help

Schrödinger software is distributed with documentation in PDF format. If the documentation is not installed in `$SCHRODINGER/docs` on a computer that you have access to, you should install it or ask your system administrator to install it.

For help installing and setting up licenses for Schrödinger software and installing documentation, see the *Installation Guide*. For information on running jobs, see the *Job Control Guide*.

Maestro has automatic, context-sensitive help (Auto-Help and Balloon Help, or tooltips), and an online help system. To get help, follow the steps below.

- Check the Auto-Help text box, which is located at the foot of the main window. If help is available for the task you are performing, it is automatically displayed there. Auto-Help contains a single line of information. For more detailed information, use the online help.
- If you want information about a GUI element, such as a button or option, there may be Balloon Help for the item. Pause the cursor over the element. If the Balloon Help does not appear, check that Show Balloon Help is selected in the Maestro menu of the main window. If there is Balloon Help for the element, it appears within a few seconds.
- For information about a panel or the tab that is displayed in a panel, click the Help button in the panel, or press F1. The help topic is displayed in your browser.
- For other information in the online help, open the default help topic by choosing Online Help from the Help menu on the main menu bar or by pressing CTRL+H. This topic is displayed in your browser. You can navigate to topics in the navigation bar.

The Help menu also provides access to the manuals (including a full text search), the FAQ pages, the New Features pages, and several other topics.

If you do not find the information you need in the Maestro help system, check the following sources:

- *Maestro User Manual*, for detailed information on using Maestro
- *Maestro Command Reference Manual*, for information on Maestro commands
- *Maestro Overview*, for an overview of the main features of Maestro
- *Maestro Tutorial*, for a tutorial introduction to basic Maestro features
- *Glide Quick Start Guide*, for a tutorial introduction to Glide
- *Protein Preparation Guide*, for information on protein preparation for Glide
- *Impact Command Reference Manual*, for information on Impact commands

EX2037.0117

- Glide Frequently Asked Questions pages, at https://www.schrodinger.com/Glide_FAQ.html
- Known Issues pages, available on the Support Center.

The manuals are also available in PDF format from the Schrödinger Support Center. Local copies of the FAQs and Known Issues pages can be viewed by opening the file `Suite_2009_Index.html`, which is in the `docs` directory of the software installation, and following the links to the relevant index pages.

Information on available scripts can be found on the Script Center. Information on available software updates can be obtained by choosing Check for Updates from the Maestro menu.

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
- Glide purchaser (company, research institution, or individual)
- Primary Glide user
- Computer platform type
- Operating system with version number
- Glide version number
- Maestro version number
- mmshare version number

On UNIX you can obtain the machine and system information listed above by entering the following command at a shell prompt:

```
$SCHRODINGER/utilities/postmortem
```

This command generates a file named *username-host*`-schrodinger.tar.gz`, which you should send to help@schrodinger.com. If you have a job that failed, enter the following command:

```
$SCHRODINGER/utilities/postmortem jobid
```


EX2037.0118

where *jobid* is the job ID of the failed job, which you can find in the Monitor panel. This command archives job information as well as the machine and system information, and includes input and output files (but not structure files). If you have sensitive data in the job launch directory, you should move those files to another location first. The archive is named *jobid*-`archive.tar.gz`, and should be sent to help@schrodinger.com instead.

If Maestro fails, an error report that contains the relevant information is written to the current working directory. The report is named `maestro_error.txt`, and should be sent to help@schrodinger.com. A message giving the location of this file is written to the terminal window.

More information on the `postmortem` command can be found in Appendix A of the *Job Control Guide*.

On Windows, machine and system information is stored on your desktop in the file `schrodinger_machid.txt`. If you have installed software versions for more than one release, there will be multiple copies of this file, named `schrodinger_machid-`*N*`.txt`, where *N* is a number. In this case you should check that you send the correct version of the file (which will usually be the latest version).

If Maestro fails to start, send email to help@schrodinger.com describing the circumstances, and attach the file `maestro_error.txt`, from the directory `%USERPROFILE%\Local Settings\Application Data\maestro90\appcrash`. If Maestro fails after startup, attach this file and the file `maestro.EXE.dmp`, from the same directory, to your email message.

EX2037.0119

EX2037.0120

Index

A

acceptor constraints 59
adding hydrogens 22
amide bonds, rotation around 49
annealing 8
`applyhtreat` utility 23

B

batch processing of ligands 97
binding affinity 8
bonds
  amide, rotation 49
  covalent, to receptor 19
  rotatable 6

C

ChemScore 8
close contacts 17
conformational search 5
conformations
  core 6
  receptor 13, 27
  ring 48
constraints 10
  acceptor 59
  application of 59
  changing name of 36, 39
  donor 59
  feature definitions 61
  feature file 95
  feature type 62
  hydrophobic 39
  ligand core 55
  markers 59
  metal atom 38
  number of atoms matching 63
  optional 60
  point at which tested 64
  positional 36
Constraints tab, Ligand Docking panel 58
Constraints tab, Receptor Grid Generation panel 34
  H-bond/Metal subtab 37
  Hydrophobic subtab 40
  Positional subtab 35
conventions, document vii
core conformations 6
Core tab, Ligand Docking panel 56
core, ligand 6, 56
Coulomb energy 9

D

database screening 9
diameter, ligand 7
dielectric constant, setting 52
directory
  installation 1
  Maestro working 2, 89
  utilities 88
distributed processing 46, 87, 98
docking
  extra-precision 11
  flexible 48
  rigid 6, 49
  subjobs 98
docking mode 11, 48
docking precision 48
donor constraints 59

E

Edit Conditional Coloring dialog box 81
efficiency metrics 75
electrostatic grids 8
Emodel 5, 8
enclosing box 38
energy
  Coulomb 9
  van der Waals 9
energy grid score 8
energy, internal strain 8
enrichment factor 9
environment variable, `SCHRODINGER` 1
Export Rotatable Groups dialog box 79
extra-precision (XP) mode 10, 48

F

feature definitions 61
  adding SMARTS patterns to 63
  excluding groups from 64
feature type, selecting for constraint 62
file name conventions 89
flexible docking 8, 48

EX2037.0121

flexible ligands 6
flexible receptor 12, 42, 55
force field 47
format conversion, to Maestro 23
functional groups, excluding from feature definitions 64

G

Glide XP Visualizer panel 80
`glide_rescore` utility 104
`glide_sort` utility 101
GlideScore 8
greedy scoring 8
grids, receptor 8

H

help option (`-h`) 88
high-throughput virtual screening (HTVS) docking mode 48
hydrogen bonds
  constraint types 59
  defining constraints 38
  geometric criteria for 38
  receptor atom sites 37
  resolving clashes 18
hydrogen treatment 22
hydrophobic map 39
hydrophobic region
  definition of 40
  size of 39
hydroxyl groups, rotation of 42

I

identical chains 15
`impact` command 97
Induced Fit protocol 12
interactions, viewing per-residue 78
internal strain energy 8
  penalty for 73
  reducing with postminimization 72

J

job control commands 88
job incorporation 46

L

license, LigPrep 20
ligands
  constraint features 61
  diameter 7
  efficiency metrics 75
  file format 53
  flexible 6
  pose, definition 5
  rigid 6
Ligands tab, Ligand Docking panel 53
LigPrep 20–22

M

Maestro working directory 89
Maestro, starting 1
metal-ligand interaction
  constraints 38
  state penalties 50
metalloproteins, charge and protonation states . 20
metals
  adjusting charges 16
  coordination of ligands to 20
  covalent bonds to protein 16
  use of Epik state penalties 50
minimization
  of docked poses 72
  of rough-scored poses 52
multimeric protein structures 15
multiple processors, running on 97

O

OPLS-AA 8
Output tab, Ligand Docking panel 70
overview of protein preparation 15

P

`para_glide` utility 97
parallel processing 87
partial charges
  ligand 54
  receptor 29
  threshold for nonpolar atoms 29
penalties
  buried polar groups 9

EX2037.0122

state (Epik) ... 50
strain ... 74
Per-Residue Interactions panel ... 78
poses ... 5
bypassing initial elimination ... 52
number written out ... 71
post-docking minimization ... 72
refining input ... 49
RMSD from input geometry ... 72
viewing ... 77
Prime ... 12
product installation ... 109
protein
adjustment of structure ... 16
misprotonation of ... 18
protein preparation, overview ... 15
protonation state ... 20

R

receptor
constraint sites ... 10
exporting with alternative hydroxyl orientation ... 78
flexibility ... 12, 27, 42, 55
grids ... 8
Receptor tab, Receptor Grid Generation panel . 28
reference ligand ... 55
refining input poses ... 49
rewards
hydrophobic ... 9
polar interactions ... 9
rigid docking ... 6, 49
ring conformations ... 48
rings, 5– and 6–membered ... 6
RMSD
from input geometry ... 72
from reference ligand ... 56
restriction on ligand core ... 56
rotamer groups, ligand ... 6
rotatable bonds ... 6

S

sampling method
SP mode ... 6
XP mode ... 11
Schrödinger contact information ... 110
score
energy grid ... 8
GlideScore ... 8
model energy ... 8
per-residue interaction ... 73, 78
score-in-place ... 49
scoring function ... 8
scoring input poses ... 49
Settings tab, Ligand Docking panel ... 47
similarities, XP descriptor ... 83
similarity scoring ... 65
Similarity tab, Ligand Docking panel ... 66
Site tab, Receptor Grid Generation panel ... 31
SiteMap binding site
selecting for grid generation ... 27
specifying box size and center for ... 31
SMARTS patterns, use for feature definitions .. 63
SP mode ... 48
standard-precision (SP) docking mode ... 11, 48
strain correction ... 73
parameters for ... 74
strain energy ... 8
structures
format conversion ... 23
requirements for docking ... 19
subjobs, Glide docking ... 98
subset test ... 8
symmetry-equivalent atoms ... 37

U

utilities
`applyhtreat` ... 23
`glide_rescore` ... 104
`glide_sort` ... 101
help option (`-h`) ... 88
location of ... 88
`para_glide` ... 97

V

van der Waals grids ... 8
van der Waals radii, scaling of ... 12, 27

W

web interface ... 3
working directory, Maestro ... 89

EX2037.0123

X

XP descriptors, writing 50
XP mode..... 10, 48
XP Visualizer 79
table color schemes..... 80
table navigation..... 82

120 West 45th Street, 29th Floor
New York, NY 10036

101 SW Main Street, Suite 1300
Portland, OR 97204

8910 University Center Lane, Suite 270
San Diego, CA 92122

Zeppelinstraβe 13
81669 München, Germany

Dynamostraβe 13
68165 Mannheim, Germany

Quatro House, Frimley Road
Camberley GU16 7ER, United Kingdom

SCHRÖDINGER®