Preparing a Library of Compounds for Glide or Phase

This document contains instructions on how to prepare a virtual screening library of compounds for use by Glide or by Phase. The full preparation procedure uses LigPrep, Epik, and QikProp, but you can run the main part of it with just LigPrep. While this procedure can be varied, the instructions given below provide a good rule-of-thumb.

The steps in this process are as follows:

1. Obtain 2D SD files
2. Prepare structures
3. Filter structures (optional)

At the end of this process, you can use the structure file as input to Glide ligand docking or to Phase database creation or file searching.

1. Obtain 2D SD files

If you do not already have a library of compounds in 2D SD format, you can find links to sources of compounds at the HTScreening.net web site, http://www.htscreening.net/home, or the ZINC web site, http://blaster.docking.org/zinc.

If your computer does not have a file size limit (such as the common 2 GB limit), or if the library contains less than 50,000 compounds, you can continue with the procedure. Otherwise you must split the library into separate files, each containing about 50,000 compounds, and process each file. You can split the file using the sdsubset utility:

```
$SCHRODINGER/utilities/sdsubset -n range full.sdf > subset.sdf
```

Here, *range* specifies the first and last structure to extract, separated by a colon, e.g. 1:50000.

2. Prepare structures

The structures in your 2D SD files must be converted into 3D all-atom structures, using LigPrep. The LigPrep actions performed in this part of the preparation are:

- Retain the specified chiralities
- Generate a maximum of 4 stereoisomers
- Remove cofactors (desalt)
- Neutralize charged acidic or basic groups
- Generate only the lowest-energy ring conformation
- Generate tautomers and ionization states

The output should be generated in Maestro format. Suggested commands to run this job with or without Epik are as follows.

Glide:
```
$SCHRODINGER/ligprep -epik -r 1 -s 4 -isd infile.sdf -omae outfile.mae
$SCHRODINGER/ligprep -i 2 -r 1 -s 4 -isd infile.sdf -omae outfile.mae
```
Phase:
```
$SCHRODINGER/ligprep -epik -i 2 -W e,-ms,4 -r 1 -s 4 -isd infile.sdf -omae outfile.mae
$SCHRODINGER/ligprep -i 2 -r 1 -s 4 -t 4 -isd infile.sdf -omae outfile.mae
```

Here, *infile*.sdf is the name of the 2D SD file, and *outfile*.mae is the name of the Maestro output file. If you are planning to use the file with Phase to search for matches to a pharmacophore hypothesis, you can save the file in SD format with the -osd *filename* option. If you want to distribute your jobs over multiple processors, you can include the options -NJOBS *n* -HOST *multi-cpu-host*:*n*, where *n* is the number of processors.

3. Filter structures (optional)

After preparing the structures, you might want to filter the structures, based on their ADME or toxicology properties, for example. Below are examples of the use of QikProp to calculate ADME properties and filtering based on these properties.

To calculate the QikProp properties, run the following command on the files generated in the last step:

```
$SCHRODINGER/qikprop -nosa filename.mae
```

The output structure file is named *filename*-out.mae, and contains all the properties and descriptors generated by QikProp.

SCHRÖDINGER

Two filtering schemes are provided below, which use the utility `ligfilter`. You can also create your own scheme. The input file for `ligfilter` is the output file from QikProp, *filename*`-out.mae`. The output file is *filter*`.mae`.

- **Lipinski filter**

  This filter passes compounds with molecular weight ≤ 500, octanol/water partition coefficient log $P_{o/w} \leq 5$, 5 or fewer hydrogen bond donors, and 10 or fewer hydrogen bond acceptors:

  ```
  $SCHRODINGER/utilities/ligfilter -e "r_qp_mol_MW <= 500" -e "r_qp_QPlogPo/w <= 5"
      -e "r_qp_donorHB <= 5" -e "r_qp_hb_accptHB <= 10" -o filter.mae filename-out.mae
  ```

- **Coarse filter to remove high risk compounds**

  This filter passes compounds with molecular weight in the range 150–650, octanol/water partition coefficient in the range –3 to 6.5, aqueous solubility logS $\geq$ –7, and polar surface area FISA ≤ 175:

  ```
  $SCHRODINGER/utilities/ligfilter -e "r_qp_mol_MW >= 150" -e "r_qp_mol_MW <= 650"
      -e "r_qp_QPlogPo/w >= -3" -e "r_qp_QPlogPo/w <= 6.5" -e "r_qp_QPlogS >= -7"
      -e "r_qp_FISA <= 175" -o filter.mae filename-out.mae
  ```