



AutoDock Virtual Screening: Raccoon & Fox Tools

Stefano Forli
Ruth Huey

The Scripps Research Institute
Molecular Graphics Laboratory
10550 N. Torrey Pines Rd.
La Jolla, California 92037-1000
USA



3 December 2011 v.0.4

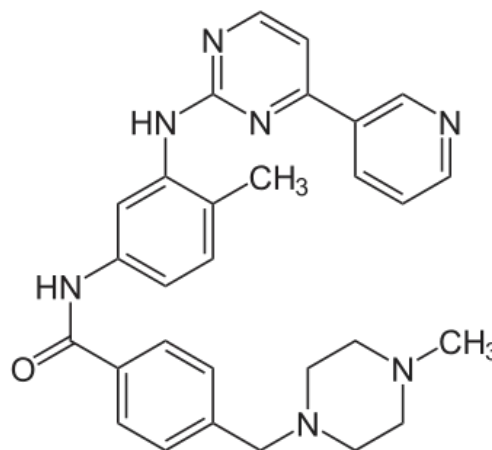
Introduction

Protein

The target used in this virtual screening tutorial is kinase domain of the proto-oncogene tyrosine-protein kinase C-Abl. (PDB id: 1iep). This protein is an important target for the chemotherapy of the cancer, in particular for the treatment of chronic myelogenous leukemia (CML).

Reference ligand

The protein structure used is complexed with the inhibitor Imatinib (Gleevec, Novartis), an approved drug for the treatment of CML and gastrointestinal stromal tumors (GISTs).



Files

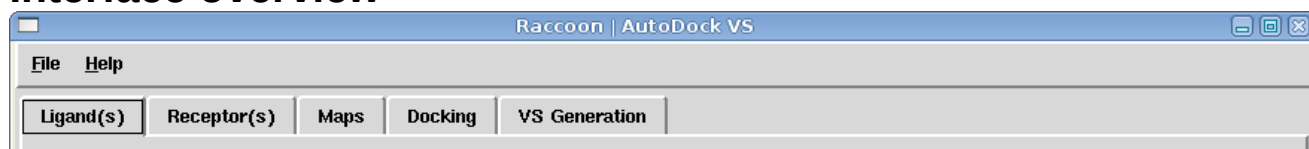
The files used in this tutorial are on the Desktop, under the directory “*VirtualScreening*”. This directory contains the following sub-directories:

- input/ligands** : imatinib + 499 NCI compounds;
- input/receptors** : PDBQT of target protein structure; experimental structure of imatinib bound to C-Abl;
- input/config** : AutoGrid (GPF) and AutoDock (DPG) parameter files;
- vs_results** : pre-calculated docking results to be used during the analysis step.

A. Virtual screening preparation (Raccoon)

To start Raccoon, double-click on the icon “**Start Raccoon**” on the Desktop.

Interface overview



Ligand(s)

Specify and filter input ligand libraries

Receptor(s)

Specify one or more target structures including automatic flexible residues generation.

Maps

Specify grid parameters and map policies for AutoGrid calculations.

Docking

Specify accuracy and speed parameters for AutoDock dockings.

VS Generation

Generate auto-run virtual screening packages

Ligands

1. Import the Imatinib structure to be used as reference compound:

- Click on the **[+] Add ligands...** button.
- Browse to the **"input/ligands"** directory
- Open the file **"imatinib.pdbqt"**

The ligand is in the native AutoDock format (PDBQT) therefore it will be imported immediately.

When a non-PDBQT file is imported, it will be converted automatically using the parameters specified in the **PDBQT generation options** button.

The default options will be used for the next file.

2. Import a small subset of the NCI library (499 compounds) obtained from the ZINC database:

- Click on the **[+] Add ligands...** button
- Browse to the **"input/ligands"** directory and change the file type to "Mol2".
- Open the file **"499_NCI_subset.mol2"**

Raccoon will report the number of molecules found in the multi-structure file and it will ask for a directory where to split it into single PDBQT files. Accept the default location and start the conversion.

When ligands are added, they are automatically filtered with default filter parameters.

3. Click on the **Filter ligands** button and check how different filters affect the list of accepted ligands. Use default values for this generation.

Receptor

4. Import C-Abl receptor structure:

- Click on the button **Add receptor file**
- Browse to the **"input/receptors"** directory

-Open the file “**1iep_receptor.pdbqt**”

Grid maps

5. Specify the grid box placement and size:

-Click on **Load GPF template** the button

-Browse to the “**input/config**” directory

-Open the file “**kinase_binding_site_template.gpf**”

6. [OPTIONAL] Change the map policy to save time. In this way Raccoon will run AutoGrid to cache the maps and copy them to every ligand directory:

-In the “**Run AutoGrid...**” field select the “**Now (and cache the maps)**” option.

-Be sure that AutoGrid program is found, otherwise specify it.

[NOTE: AVOID THIS STEP UNLESS THE HARDISK HAS A LARGE ENOUGH SPACE AVAILABLE]

Docking parameters

7. Use the AutoDock predefined search parameters:

-Click on the “**[select docking setup]**” menu and select “From template...” option.

-Click on **Load DPF template** , browse to the “**input/config**” directory and specify the file “**imatinib_1iep_receptor.dpf**”, then set: “**ga_run 100**”.

[OPTIONAL: generate a default DPF file by clicking on **default**]

VS generation

8. Revise the virtual screening information to be sure that all required parameters are green, then specify the output directory where the files for the docking will be generated

-Click on the “**Set directory...**” menu, browse to the VirtualScreening directory and specify a new directory called “**abl_virtual_screening**”

9. Start the generation by pressing **“Generate”**

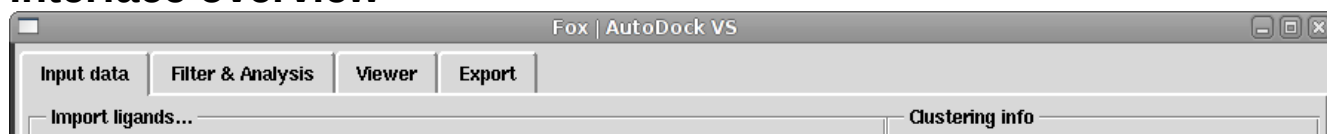
10. Run the docking

When the generation process is done, use the system File Explorer to open the folder containing the results, locate and double-click on the file **“RunVs.bat”**. This batch script will open a terminal and run the entire sequence of docking jobs.

B. Virtual screening analysis (Fox)

To start Fox, double-click on the shortcut “**Start Fox**” on the Desktop

Interface overview



Input data

Specify which result files will be analyzed together with the receptor structure. Perform the ligand pose extraction.

Filter & Analysis

Filter the raw results to refine the quality by using statistical properties of the docking and interactions with the receptor.

Viewer

Allow to visualize and select the top results inside the receptor structure.

Export

Export selected results in various formats.

Input data

All the AutoDock DLG files will be imported and ordered per ligand (if multiple DLG files are available per ligand). Dockings will be re-clustered to extract result poses, calculating interactions with the receptor, and saved in a PDBQT+ structure file, containing statistical and interactions data.

1. Select the receptor structure


- Click on **Choose structure...** button
- Browse to the **“receptors”** directory
- Open the file **“1iep_receptor.pdbqt”**

The receptor will be loaded in the session and used to calculate interactions.

2. Load the grid box in the session:

- Click on **Set grid box...**
- Browse to the **“config”** directory
- Load the file **“kinase_binding_site_template.gpf”**

3. Import the docking files:

- click on **Import DLG...**
- click on **Choose directory**
- select the **“kinase_screening_RESULTS”** directory
- check **“include subdirectories”** option
- click on the **“Options”** button to show extra customizable parameters. Leave the defaults for this exercise.
- click on the scan icon 
- once the scan is done, press **“Process”**

When the process is completed, close the window. Ligands should be now listed in the Input tab spreadsheet.

Filter & Analysis

Switch to the “**Filter & Analysis**” tab. Result poses need to be filtered in order to remove low score ligands and increase the quality of results.

It is possible to specify which pose is going to be considered the result, either the “*Lowest energy*” or the “*Largest cluster*”. Alternatively, it is possible to restrict the analysis only on ligands in which the lowest energy pose is the most statistically relevant (“*Unique*”).

Depending on the systems considered, either energy or cluster criteria have provided better results.


The ligand poses are filtered in the following order:

1. **energy**
2. **clustering**
3. **ligand efficiency**
4. **interactions (if selected)**

For this reason the energy filter is crucial to define the set of ligands that will be visually analyzed eventually, although the energy range is also function of the ligand set (drug-sized VS fragment ligands)

The Worst/Best values of every filter are adapted automatically to the current ligand set.

The selection of the results is often an *iterative process* requiring to go back and forth from the visual analysis to the filtering phase. Different filter sets can be saved to visually inspect the results in the next steps.

3. Every time a filter value is changed, press “**Filter**” to check how many ligands passed the filtering test. Different filter sets can be saved for further inspection by clicking on the [] icon.

Test different filter combinations separately or combining them together. Use the **Default** button to restore the initial values.

4. Try to test the following conditions:

- verify that a minimum energy cutoff of about -8.0 Kcal/mol discards most of the ligands.
- select all ligands with a minimum **ligand efficiency** of -0.5 (very high value) and see if imatinib is in the list of ligands in the Viewer tab. (**why?**)

5. Go back to the **Filter & analysis** tab, set the following minimum values and be sure that **60** ligands pass all the filters:

Pose : Unique
Energy : -7.0 Kcal/mol
Cluster size : 51 %
Ligand eff. : -0.40

6. Save the filter set [] with a convenient name, i.e.:

“First filter ---E(-7.0,-15.4) LE(-0.4,-0.67) Clust(51.0) Pose(unique)”

then switch to the **Viewer** tab.

7. The **Viewer** tab is arranged in the *Information part* (on the left) and the *3D Viewer*. Most of the viewer parts can be re-arranged by dragging the edges of the panels.

The 3D viewer representations are conceptually arranged in this way:


LIGAND


- current ligand (currently loaded ligand)
- secondary ligands

RECEPTOR

- special residues (inside the grid-box)
- all residues (the entire protein)


In the 3D viewer there is the protein surface and a representation of the grid box used for the screening. The binding site is mostly buried inside the protein surface, so to easily analyze the results the protein representations must be adapted:

- click on the settings icon 
- switch to the “Target” tab and change the MSMS transparency value to 10

8. Visualize the results by switching in the Viewer tab. It is possible to have an overview of the entire ligand set by clicking on the  **Hits view** button:

- click on **Hits view**->“**Show centroids**”
- inspect the color-coded spheres (blue=better score, red=worst score) and note any clusters
- select imatinib in the results list and compare the centroids with the ligand
- click on **Hits view**->“**Hide centroids**”


9. Load the reference structure of the imatinib and compare it with the docked one to see if docking protocol accuracy was enough to reproduce the experimental structure.

- Click on the icon  in 3D viewer toolbar and select “load”,
- Browse to the “**ligands**” directory and load “*xray_ligand.pdbqt*”
- To hide/show the ligand, click on the same icon again and choose hide/show

Inspect the interactions the ligand is establishing by clicking on the “**Interactions**” button and dragging the panel down.

10. Use the reference ligand structure to compare the other ligand docked poses, select ligands that look interesting accordingly to your judgement by clicking in the orange bar under the column “S” (select). If a ligand looks particularly interesting, use the right-click button to set it to high priority.

11. Select again imatinib and show residue labels to identify residues interacting with the ligand:

- click on the settings button 
- switch to the “**Target**” tab and set “**Label**” to “**inside the box**”
- annotate any interesting interactions between imatinib and the receptor
[HINT: hydrogen bond donor with Glu286?]
- to remove labels, set “**Label**” to “[off]”

12. Go back to the “*Filter & Analysis*” tab and add a hydrogen bond interaction filter with Glu286:

- load the first filter set saved “**First filter ---E(-7.0,-15.4) LE(-0.4,-0.67) Clust(51.0) Pose(unique)**”
- activate a filter in the Interactions panel by clicking on the checkbox on the left
- double-click on the text entry and select chain A and residue Glu286; there is no need to specify also the atoms;
- select the interaction type “HB(d)”
- press “Filter”
- save the filtering set with a new name as “HB Glu286”
- note the progressive reduction of ligands passing each filter

13. Now **14** ligands have passed the filters. Use the reference ligand structure to compare the other ligand docked poses, select ligands that look interesting accordingly to your judgement by clicking in the orange bar under the column “S” (select). If a ligand looks particularly interesting, use the right-click button to set it to high priority.

[HINTS: ZINC01596698, ZINC01665940...]

14. Use directly imatinib docking to find ligands binding with the same interactions:

- load imatinib by clicking on its name on the list
- click on “Find similar...” near the Interactions panel
- use the following settings:

Pose	:	“Lowest energy
Energy	:	off
L. efficiency	:	off
HB donor	:	ON
HB acceptor	:	ON
van der Waals	:	off → WHY?
Residue only	:	ON
- change the name of the filter set by adding “*All interactions*” to the current name
- press “Filter”

No ligand should pass this test, because no ligand in this library is large enough to fit in the elongated binding site as imatinib does. Switch to the “**Filter & analysis**” panel to see how the filters have been set.

15. Repeat the similarity search by using different values:

- load imatinib by clicking on its name on the list
- click on “Find similar...” near the Interactions panel
- use the following settings:

Pose	:	“Lowest energy”
Energy	:	off
L. efficiency	:	off
HB donor	:	off
HB acceptor	:	ON
van der Waals	:	off
Residue only	:	ON
- change the name of the filter set by adding “*Less interactions*” to the current name
- press “Filter”

Now **4** more ligands should be present together with imatinib. Inspect the results and note the increased 'quality' of the results. Select interesting ligands that are not yet selected.

16. Go back (last time!) to the “*Filter & Analysis*” and restore all the default values in the filters. Turn also the interaction filters off (“*Disable all*”), and set the following filters

- set the pose selection to “Lowest energy” (**WHY?**)
- set “Ligand efficiency” to -0.55
- press “Filter”
- save the filter set with the name “Fragments”

17. In the Viewer, inspect the **15** results and compare them with the reference structure of imatinib. Look for hits with interactions similar to imatinib, and hits that explore novel binding regions. Use the green arrows ◀ ▶ near the energy score panel (or keyboard arrows *left-right*) to switch between lowest energy (LE) and largest cluster (LC) poses.

Fragments -versus- full ligands

The energy of the results of the “Fragments” set is low in absolute value, but rather high for the size of the molecules (i.e. high efficiency). The ligand efficiency is an index of how 'efficiently' the atoms contribute to the overall interaction of the molecule. For this reason, the absolute value of the energy (score) is not always the right criterion for selecting hits.

The fragment results should be inspected and saved separately in the list of selected molecules.

18. Switch to the “**Export**” tab and save the results selected during the analysis. For each filter set it is possible to choose which set of ligand is going to be saved:

- any selected ligands ('**selected**')
- only high-priority selected ligands ('**selected (HiPriority)**')
- ligands for which a special comment has been saved ('**annotated**')
- any ligands selected or not that satisfied the criteria specified in the filter set ('**filtered**').

For each filter set to be saved, export as comma-separated values (CSV) or text (TXT) files:

- select the filter set in “Filter set”
- select the ligand set in “Ligand set”
- press the “Export button, select the file type (TXT) and specify a filename.

CSV files are useful for virtual screenings because allow to store a compact results list that can be later opened with spreadsheets like Microsoft Excel or OpenOffice/Libreoffice Calc. Text files can be opened with any text editor.

FOR QUESTIONS, SUGGESTIONS:
forli@scripps.edu