AutoDock

# AutoDock
# Virtual Screening:
# Relaxed Complex scheme

Stefano Forli
Alexander L. Perryman

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

THE SCRIPPS RESEARCH INSTITUTE®

3 December 2011 v.0.4

**Relaxed Complex (RC)**

Often the high affinity of a ligand is the result of a sum of bindings to multiple receptor conformations (even conformations occurring rarely).

The relaxed complex method was designed to incorporate protein flexibility into drug design, it can simulate these multivalent binding events (like SAR by NMR), and it has been developed using AutoDock[1-3].

Essentially, the method consists in two phases:

1. molecular dynamics (MD) simulation to sample multiple diverse receptor conformations (tens or hundreds)

2. docking of ligand libraries (i.e. fragments) against the sampled receptor conformations

The docking results are analyzed to derive a binding energy distribution of the results, as following:
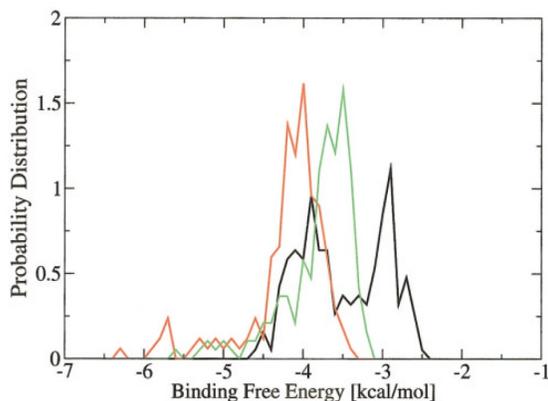


The target used for this turorial is the HIV protease structure, an important target for the AIDS therapy[4]. The structure is a two-fold symmetry homodimer with two flexible flaps. The catalytic activity is performed by two Asp residues processing viral proteins during the infection.

For this tutorial, a RC experiment is approximated using a small ligand library of

**10** compounds and **5** receptor conformations, for a total of 50 dockings.

Fragment-screening hits are usually used to design bigger binders, for example by linking together two fragments binding to two distinct protein '*hot spots*'.

The protein conformations are sampled from a MD simulation of 22 ns for an actual experiment performed in our lab, involving a particular mutant of HIV-protease (courtesy of Dr. A. L. Perryman).

## Bibliography

[1] Lin, J.H., Perryman, A.L., Schames, J.R., and McCammon, J.A. "Computational drug design accommodating receptor flexibility - the Relaxed Complex scheme," J. American Chemical Society.

[2] Lin, J.H., Perryman, A.L., Schames, J.R., and McCammon, J.A. "The Relaxed Complex method: Accommodating receptor flexibility for drug design with an improved scoring scheme," *Biopolymers*, 68 (1): 47-62

[3] Rommie E. Amaro, Riccardo Baron and J. Andrew McCammon, An improved relaxed complex scheme for receptor flexibility in computer-aided drug design", JOURNAL OF COMPUTER-AIDED MOLECULAR DESIGN Volume 22, Number 9, 693-70

[4] Perryman, A.L., Zhang, Q., Soutter, H.H., Rosenfeld, R., McRee, D.E., Olson, A.J., Elder, J.E., and Stout, C.D. "Fragment-based screen against HIV protease" *Chemical Biology & Drug Design*, 75(3): 257-268 (2010) Fragment screening protease

[5] Perryman, A.L.*, Lin, J.H., and McCammon, J.A. "HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: Possible contributions to drug resistance and a potential new target site for drugs," *Protein Science*, 13 (4): 1108-1123 (2004).

# A. Results extraction   (Fox)

Fox will be used to cluster results and extract the energy of every ligand. The **5** target structures used in the exercise are the following:

> xMut_md00230
> xMut_md02390
> xMut_md09800
> xMut_md17230
> xMut_md21960

The numbering correspond to the picoseconds of the ~21 ns MD trajectory at which the structure has been sampled [5]. Therefore, from the first to the last, the conformations <u>deviate</u> more and more from the initial crystal structure.

For each of them, the corresponding docked ligand results must be loaded, processed and results exported.

**1.** Setting the receptor structure

> - Click on  **Choose structure...**  button
>
> - Browse to the **"receptors"** directory and load one structure

**2.** Import ligand results corresponding to the structure:

> - click on  **Import DLG...**  button
>
> - browse to the **"dockings"** directory and select the results corresponding to the receptor structure
>
> - click on the scan icon 🔍
>
> - once the scan is done, press "**Process**"

**3.** Filtering is not necessary, then in the "Filter & Analysis" tab select "**Lowest energy**", leave all default values and press "**Filter**". In this way, all the ligands will pass the filters (by default, in fact, filter values are equal to current ligand set properties).

**4.** Briefly inspect briefly the ranking of the ligands in the "**Viewer**" tab (i.e. *is the top hit always the same for all targets?*)

**5.** In the "**Export**" tab save the result text files:

      - change "Ligand set" to "**filtered**" (i.e. keep all the results)

      - click on "**Export**"

- change the file type to "**Text file**" and save the file with a progressive name (i.e. **receptor1.txt**, **receptor2.txt**, …)

**6.** Repeat steps **1** to **5** for all the receptor structures.

**7.** With Explorer file manager browse to the location where the text files have been saved. Double-click on the first two result text files to open them (i.e. "**receptor1.txt**", "**receptor2.txt**") and compare the top two hits.

Given the size of the samplings, it is not possible to perform a distribution of the docking results, therefore the hits selection will be performed on the simple ranking comparison in the energy-sorted results

The full list of the hits should be like the following table:

**Combined docking rankings on all protein snapshots**

| md00230 | md02390 | md09800 | md17230 | md21960 |
|---|---|---|---|---|
| ZINC02390501 | ZINC04387707 | ZINC19089585 | ZINC02390501 | ZINC19089585 |
| ZINC02685827 | ZINC02387687 | ZINC02390501 | ZINC04387707 | ZINC04387707 |
| ZINC19089585 | ZINC02685827 | ZINC00572736 | ZINC02387687 | ZINC02685827 |
| ZINC04387707 | ZINC02390501 | ZINC06067294 | ZINC02685827 | ZINC02390501 |
| ZINC00572736 | ZINC19089585 | ZINC15015303 | ZINC19089585 | ZINC00572736 |
| ZINC02387687 | ZINC00572736 | ZINC04387707 | ZINC06067294 | ZINC02387687 |
| ZINC15015303 | ZINC06067294 | ZINC02685827 | ZINC00572736 | ZINC15015303 |
| ZINC06067294 | ZINC15015303 | ZINC02387687 | ZINC15015303 | ZINC06067294 |
| ZINC02753817 | ZINC02753817 | ZINC09355876 | ZINC09355876 | ZINC02753817 |
| ZINC09355876 | ZINC09355876 | ZINC02753817 | ZINC02753817 | ZINC09355876 |

The results suggests the following hit ranking:

        **ZINC02390501** > **ZINC19089585** > **ZINC04387707**

# B. Results visual inspection  (PMV)

The 3 promising results will be visually inspected with PMV.

**8.** Open PMV and load a receptor structure to be used to localize the docking results:

> - on the menubar, click on **File->Read molecule**  (or click with the right-mouse button on "All molecules" in the Dashboard)
>
> - browse to the **"receptors"** directory and load the structure "**xMut_md00230.pdbqt**"
>
> - [OPTIONAL] on the menubar, click on  **3D Graphics-> SetBackGroundColor**  and set background color to gray or dark-cyan; click on "Dismiss" then press "**d**" on the keyboard

**9.** Change the representation of the receptor to make easier to analyze the results. In the Dashboard in the line corresponding to the molecule "**xMut_md02230**":

> - click on the circle button below the "**R**" to show the secondary structure
>
> - click on the circle button below the "**L**"  to hide lines
>
> - click on the triangle button below the "**Cl**", select '**custom color**", and pick an orange color

**10.** Highlight the two catalytic aspartic acid residues in the active site:

> - on the menubar, click on  **Select-> Select from string**
>
> - in the "Residue" field, type "**ASP25,ASP124**", press "**Add**", "**Dismiss**"
>
> - in the Dashboard in the line corresponding to "**Current Selection**", press the green circle under the "**C**  "
>
> - in the Dashboard in the line corresponding to "**Current Selection**" click on the triangle button below the  "**Cl**", select '**cpk**', then '**by atom type**'
>
> - in the Dashboard, deselect all by clicking on

**11**. Import the reference structure of a known binder (**indinavir**):

> **-** on the menubar, click on **File->Read molecule** (or click with the right-mouse button on "All molecules" in the Dashboard)

> - browse to the **"_results"** directory, load "**indinavir_xray.pdb**"

The highlighted ASP residues are essential to the catalytic activity of the HIV protease, then the interaction with them is important for the inhibition activity of the ligands. When binding in the active site, Indinavir engages them preventing the protease activity.

**12**. The molecular shape of indinavir will be used to help evaluating the results:

> - on the line of "**indinavir_xray**" click on the red button below "**L** " to hide lines

> - click on the green button below "MS " to show the MSMS surface

> - move to the "**Tools**" tab, and in the Geometry Tools tree scroll to '**Indinavir_xray->MSMS**", click on "**MSMS**"

> - in the "**Appearance Tools"** scroll the "opac:" wheel to about 0.5

> - move back to the Dashboard and double-click on the "**indinavir_xray**" entry to hide it

To simplify the visualization of the results, for every docked ligand all lowest energy and largest poses of the different dockings have been combined in a single file.

**13.** Load the first hit (ZINC02390501):

> - on the menubar, click on **File->Read molecule** (or click with the right-mouse button on "All molecules" in the Dashboard)

- browse to the **"_results"** directory, load the ligand "**ZINC02390501_VS_combined**" and press "**OK**"

**14.** Select all the ligand poses and change the representation style. In the Dashboard:

- for each of the models click on the square button below "S ⬚ " to select

- on the line "**Current Selection**" click on the red button below "**L** ⚬ " to hide lines

- on the line "**Current Selection**" click on the green button below "**B** ⚬ " to show sticks

- on the line "**Current Selection**" click on the triangle button below the "**Cl** ▦ ", select '**stick**', select '**balls**', then '**by atom type**'

- deselect all by clicking on ✎

**15.** Visually inspect the arrangement of the docking poses and their interaction with the catalytic residues:

- compare the overlap between fragment dockings and the molecular shape of indinavir

- which portion of the ligand varies its placement the <u>most</u>?

- which portion of the ligand varies its placement the <u>less</u>?

- how the <u>symmetry</u> of the receptor affects result analysis?

From the overall docking analysis, ZINC02390501 can be considered a good hit.

**16.** Hide docked poses of ZINC02390501 by double-clicking on the Dashboard entry names ("..._model1", "...model2", …). The entry names will become gray.

**17.** Load the second hit (ZINC19089585) and repeat step **10** and **11**.

**18.** Visually inspect the arrangement of the docking poses and their interaction with the catalytic residues:
- how is the spreading of the docked poses?

- is again the receptor symmetry helpful in analyzing the results?

From the overall docking analysis, also ZINC19089585 looks like a promising

hit.

**19.** Hide again all poses of ZINC19089585 (step **13**), then repeat step **10** and **11** for the third hit ZINC04387707.

> \- compare the overlap between fragment dockings and the molecular shape of indinavir
>
> \- compare the overlap with the position of water 301 (load the file "**_results\water301.pdb**")

Docking results show that ZINC04387707 does not interact directly with the catalytic Asp residues. Although, all docking poses are focused in a precise region of the protein, below the flaps. This region is corresponding to the position of the structural water 301 [2] that is very important for ligand binding.
For this reason, ZINC04387707 could be used as a starting point to build a binder that is able to interact with the moving loop and possibly displace the water.

**20.** As a negative control, load the worst result of the series in the table at step **7**. Hide all the poses of the previous ligands (step **13**), then load the file "**ZINC09355876_VS_COMBINED_NEGATIVE_CONTROL.pdbqt**" and repeat step **10** and **11.**

**21.** Notice the spread of the poses around the active site, the lack of a consistent binding mode and interactions with the catalytic residues in most of the cases.

To recapitulate, when performing a RC scheme experiment, it is important to look for the following patterns:

> \- good **score distribution** among target conformations
>
> \- interactions with **important residues**
>
> \- **consistent binding** (or low-spread) poses
>
> \- any matches with **known binders** (ligands, water, …)