

Tutorial for LeDock



Email: htzhao@lephar.com

Website: www.lephar.com

1. Introduction

This tutorial is mainly meant for those with little background in computational chemistry and computer science, by focusing on the use of the graphic version of LeDock. Before we get started, please have a visualization tool such as Pymol in place, in order to have an intuitive comparison between docking poses and the native pose in the X-ray structure.

In this tutorial there are a few test cases with targets including kinases, heat shock protein 90 and Influenza virus neuraminidase. In each directory, there are two files: a PDB file from Protein Data Bank and a SYBYL Mol2 file of the ligand. In the following, we will use neuraminidase as an example to show how to dock the ligand (i.e. oseltamivir) into its target.

2. Graphic interface

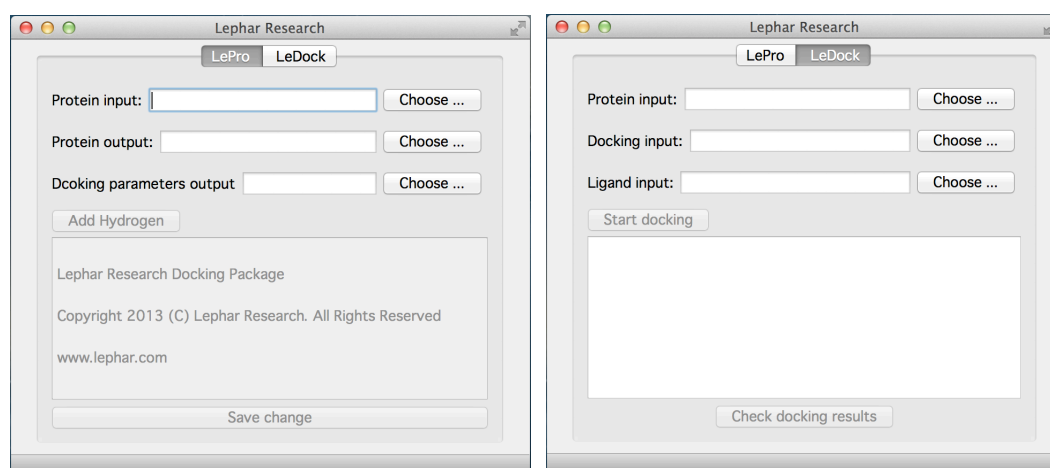


Figure 1. Graphic interface of LeDock on Mac. The interface of Win version is similar to Mac.

Step 1: Start the program by double click “LeDock” for Mac or “LeDock.exe” for Windows. The graphic interface is shown as in Figure 1. It contains two tabs. Tab “LePro” can be used to process the protein (i.e. remove all water, cofactor, ions, metal, ligands) and write an input file for docking. Tab “LeDock” is used for docking.

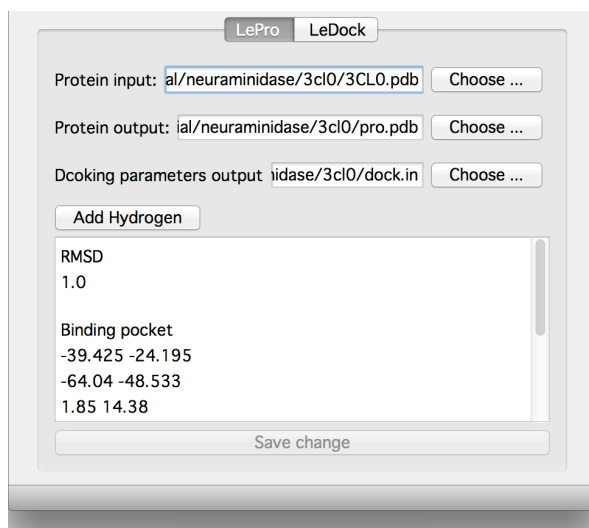


Figure 2.

Step 2: If you already have a protein ready for docking (e.g., explicit hydrogen atoms have been added, the ligand has been removed from the binding site), one can skip this step. Click Tab “LePro” and choose the PDB file of the target 3CL0.pdb, and then click “Add Hydrogen”. Two additional files will be generated “pro.pdb” and “dock.in”. See Figure 2.

pro.pdb: the protein which will be used for docking. In pro.pdb, hydrogen atoms have been added and all water molecules/ligands/ions/cofactors have been removed. In some cases, cofactor, metal or a few water molecules are important for docking. If so, one can manually copy the coordinates of such molecules from the original PDB into pro.pdb and change “HETATM” to “ATOM”. Missing hydrogen atoms for such small molecules shall not significantly influence the docking outcome. However, it is recommended to use other methods/tools to add hydrogen atoms for such molecules. In the tutorial, there is no need to keep any water/cofactor/metal.

dock.in: the input for docking. The binding site is determined as to include any protein atom within 4 Å of any heavy atom of the complexed ligand. It uses the coordinates of “HETATM” of the first small molecule to define the binding site. However, in most cases, the first small molecule in the PDB file is not necessarily the one binding to the targeted binding site. Please make sure the binding site is correct.

It is important to understand the docking parameters in “dock.in”, even though LeDock takes very few parameters for docking.

RMSD

1.0

RMSD cutoff in Å for clustering of poses after docking, in order to reduce the redundancy of docking poses. The default value is 1.0 Å, which is suitable for high-throughput virtual screening of a large chemical library. For medicinal chemists, it is recommended to use smaller values such as 0.5 Å to keep more poses.

Binding pocket

-39.425 -24.195

-64.04 -48.533

1.85 14.38

The six numbers above define a binding pocket where ligands will be docked in. This parameter is foremost. The first two numbers (-39.425 and -24.195) defines on the X dimension the starting and ending point. Similarly, the second and the third two numbers define Y and Z dimension, respectively. If you have observed the below information by using LePro

Binding pocket

xmin xmax

ymin ymax

zmin zmax

Which simply means that LePro failed to detect any pocket information. Thus, it is important for you to specify the correct binding site. If you ever used AutoDock Vina, you can easily get the required information by the following equations

$$\text{xmin} = \text{center_x} - \text{size_x}/2$$

$$\text{xmax} = \text{center_x} + \text{size_x}/2$$

$$\text{ymin} = \text{center_y} - \text{size_y}/2$$

$$\text{ymax} = \text{center_y} + \text{size_y}/2$$

$$\text{zmin} = \text{center_z} - \text{size_z}/2$$

$$\text{zmax} = \text{center_z} + \text{size_z}/2$$

Modify it directly in the text window, and please remember click on the “Save change” button afterwards.

Number of binding poses

20

Which defines the number of runs. After clustering by a non-zero cutoff, there could be fewer poses. For larger molecules with many rotatable bonds, the number of runs shall be increased to search the configuration space more exhaustively.

That's all for the docking parameters!

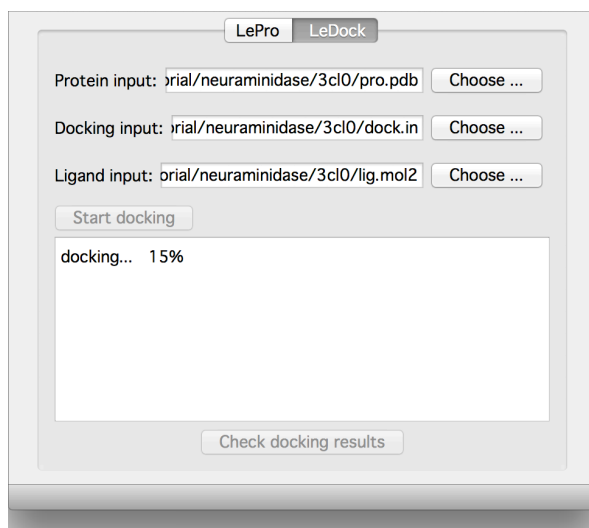


Figure 3.

Step 3: Click Tab “LeDock”. By default, the paths for protein and docking parameter are those we just generated. If not, please choose the right path. Choose the small molecule (lig.mol2 in directory “neuraminidase/3cl0”) and then click “Start docking”. Now, it is almost finished.

Step 4: After docking is finished, click “Check docking results” to check the predicted binding affinity. The binding poses were written in the file “lig.dok” in PDB format. It can be opened by Pymol or VMD (choose PDB format) for visualization of poses.

Important Notes:

1. It is not necessary to use the “LePro” to preprocess the protein. One can use a protein structure preprocessed by any third-party software as long as the format follows the CHARMM definition. In some cases, water molecules/metal/cofactors play an important role in docking, and if so, please keep such molecules in the protein and change “HETATM” to “ATOM”. Unfortunately, LePro always automatically deletes all such small molecules.
2. Please make sure the definition of binding site is correct.

3. Advanced (docking with water)

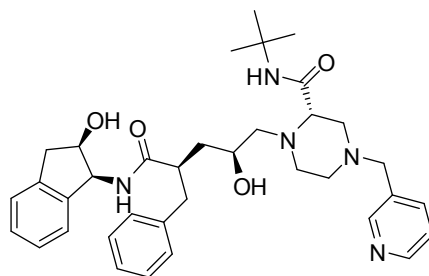


Figure 4. 2D representation of Drug Indinavir targeting HIV-1 protease.

In this section, we will dock Indinavir (Fig. 4) into HIV-1 protease. Given the number of rotatable bonds (>10), it is rather challenging.

As shown in many X-ray structures of HIV-protease complexed with small molecule inhibitors, there is a highly conserved water molecule, which forms two hydrogen bonds with the protein as acceptors, and another two hydrogen bonds with the ligand as donors. This water molecule plays an important role in protein-ligand recognition. Thus, we keep this water molecule in the protein for docking (pro.w.pdb). Please also notice that only one of the catalytic aspartyl diad (Asp25 and Asp125) was deprotonated, as evidenced by NMR study. Besides, Indinavir binds to HIV-1 protease in two distinct binding modes (see crystal01.pdb and crystal02.pdb).

By using “pro.w.pdb” as input, both distinct binding modes can successfully reproduced (Fig. 5). Docking result was saved in “lig.w.pdb”. Since this is a very challenging case, to increase the number of poses to be generated will increase the probability of getting the correct pose.

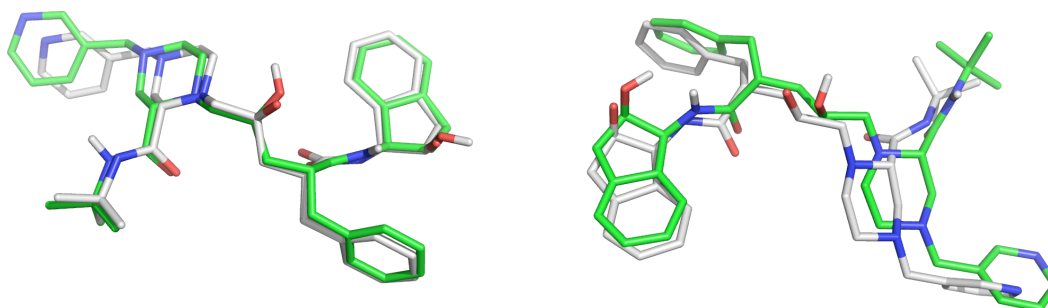
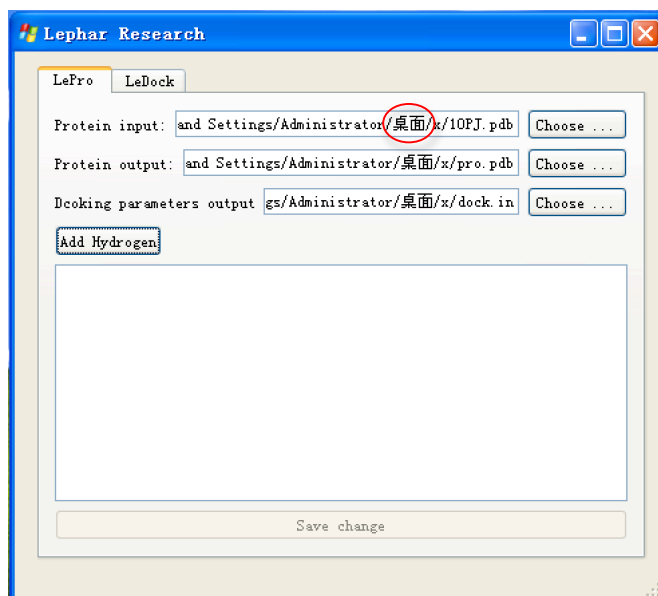


Figure 5. Superposition of predicted binding modes (carbon colored in green) on the two native poses (carbon colored in grey), respectively.

4. Troubleshooting

1. After clicking “Add Hydrogen” or “Start docking”, there is no output.

Please check whether the path contains Chinese character (see below). If so, copy the files to a place where the absolute path contains no Chinese character.



2. Microsoft C++ runtime error on Win7.

This can be fixed with CMD trick. Open cmd as administrator, and then input `bcdedit.exe /set IncreaseUserVA 2800`

For details, please refer to <https://www.youtube.com/watch?v=H7ZcBsSCISc>

3. The ligand was docked outside the binding site

Please check the setting of the binding site.

4. The ligand was docked in the binding site, but the poses are completely wrong and the predicted binding affinity is extremely unfavorable.

1) Check whether the SYBYL atom typing of the ligand is correct.

2) Check whether the PDB format (such as atom names) of the protein is CHARMM compatible.

Citation

LeDock (<http://lephar.com>) is based on a combination of simulated annealing and evolutionary optimization of the ligand pose (position and orientation) and its rotatable bonds, using a physics/knowledge-based scoring scheme derived from years of prospective virtual screening campaigns.¹⁻⁶

1. Zhao, H.; Huang, D. Hydrogen bonding penalty upon ligand binding. *PLoS One* **2011**, 6, e19923.
2. Zhao, H.; Huang, D.; Caflisch, A. Discovery of tyrosine kinase inhibitors by docking into an inactive kinase conformation generated by molecular dynamics. *ChemMedChem* **2012**, 7, 1983-90.
3. Zhao, H. T.; Dong, J.; Lafleur, K.; Nevado, C.; Caflisch, A. Discovery of a Novel Chemotype of Tyrosine Kinase Inhibitors by Fragment-Based Docking and Molecular Dynamics. *ACS Med. Chem. Lett.* **2012**, 3, 834-838.
4. Zhao, H.; Caflisch, A. Discovery of ZAP70 inhibitors by high-throughput docking into a conformation of its kinase domain generated by molecular dynamics. *Bioorg. Med. Chem. Lett.* **2013**, 23, 5721-6.
5. Zhao, H.; Caflisch, A. Discovery of dual ZAP70 and Syk kinases inhibitors by docking into a rare C-helix-out conformation of Syk. *Bioorg. Med. Chem. Lett.* **2014**, 24, 1523-7.
6. Zhao, H.; Gartenmann, L.; Dong, J.; Spiliotopoulos, D.; Caflisch, A. Discovery of BRD4 bromodomain inhibitors by fragment-based high-throughput docking. *Bioorg. Med. Chem. Lett.* **2014**, 24, 2493-6.