Tutorial of molecular generation and drug modification by MolAICal and deep learning model

Qifeng Bai

Email: molaical@yeah.net

Homepage: https://molaical.github.io

School of Basic Medical Sciences

Lanzhou University

Lanzhou, Gansu 730000, P. R. China

1. Introduction

In this tutorial, two examples are introduced: one is drug-like and FDA-like molecular generation that can be used for drug virtual screening, the other is scaffolds and groups modification based on a known ligand, which can be used to modify or design new drugs. For more detailed MolAICal, please read this paper (https://doi.org/10.1093/bib/bbaa161).

2. Materials

2.1. Software requirement
1) MolAICal: https://molaical.github.io

2.2. Example files
1) All the necessary tutorial files are downloaded from:

https://github.com/MolAICal/tutorials/tree/master/016-AImolecularGeneration

3. Procedure

3.1 Generate drug-like and FDA-like dataset

1) Generating FDA-like molecules.
#> molaical.exe -dock AI -s FDAMol -n 30 -nf 10 -nc 3 -g on -b on -v off

It will generate a file named tmpGenVS.dat that contains 30 molecular SMILES strings, and 3 folders named tmpGenMols1, tmpGenMols2, and tmpGenMols3. Users can generate much more molecules according to their needs. The generated 3D molecules are in PDBQT format by default. If users need to change the molecular format, for instance, they can use the below command for format conversion:
① Adding hydrogen (option)
#> molaical.exe -dock addh -i 1.pdbqt
② Changing "pdbqt" to "pdb" format
#> molaical.exe -dock pdbqt2pdb -i 1.pdbqt

2) Generating drug-like molecules.
#> molaical.exe -dock AI -s ZINCMol -n 30 -nf 10 -nc 3 -g on -b on -v off

It will generate a file named tmpGenVS.dat that contains 30 molecular SMILES strings, and 3 folders named tmpGenMols1, tmpGenMols2, and tmpGenMols3. Users can generate much more molecules according to their needs. The generated 3D molecules are in PDBQT format by default. If users need to change the molecular format, for instance, they can use the below command for format conversion:
① Adding hydrogen (option)
#> molaical.exe -dock addh -i 1.pdbqt
② Changing "pdbqt" to "pdb" format

3.2 Generate new similar molecules based on a known ligand

Webserver could supply a convenient way for the researchers to design drugs without any special software and hardware requirements via the browsers. MolAICal supplies this function via webserver which URL can be obtained by the below command:

#> molaical.exe -model ligdream

1) Input URL https://heisenberg.ucam.edu:5000 into the browser, you can choose the way for drawing your molecule or input your SMILES string (see Figure 1). Finally, click "Submit and Running".

Figure 1

2) It will turn to another result web page (see Figure 2). Users can download the modified molecules

in mol2 format and check them with graphical tools such Pymol, UCSF Chimera, etc (For more detailed information, please check the manual of WADDAICA: https://heisenberg.ucam.edu:5000).

The current job status information:

Job ID: f7f6a1f9-7b11-4ae3-858c-a7494eb8499b

Status: finished

Created at: Fri, 08 Oct 2021 15:24:23

Notice: This page will refresh at intervals!!! Please save below url if you donot want to wait for results. You can load this url for checking the final results after you are offline for long time.

https://heisenberg.ucam.edu:5000/static/upload/210-26-124-238/d113b6f8-284b-11ec-9239-7f4ceea7db04.html

Download generated files

The generated molecular file:

Download the generated molecules with formats of SMILES and Mol2.

Note: SMILES.dat contains the generated molecules with SMILES format. The molecules with mol2 style use the number as prefix.

View the generated molecules:

Notice: to improve the display speed, 10 molecules are loaded at the most. If you want to check more molecules, please load them manually (see documentation).
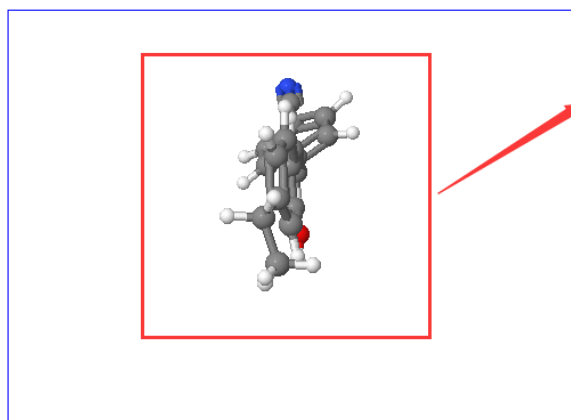
Show generated ligand in web

Figure 2