Tutorial of ADMET prediction of drugs and molecular properties calculation by MolAICal

Qifeng Bai
Email: molaical@yeah.net
Homepage: https://molaical.github.io/baiqf.html
School of Basic Medical Sciences
Lanzhou University
Lanzhou, Gansu 730000, P. R. China

1. Introduction

About 40% of drug candidates have failed in the past due to toxicity and unacceptable efficacy. In this case, the predictions of drug ADMET (**absorption, distribution, metabolism, excretion, and toxicity**) play an important role in the success of a drug candidate. In this tutorial, MolAICal is introduced to calculate ADMET and molecular properties of drugs by implement the interface package of FP-ADMET [1] tool and improved adme-pred-py under the permitted GNU General Public License v3.0 and AGPL-3.0 license. 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 or https://molaical.gitee.io

Notice: make sure installation of MolAICal correctly!

2.2. Example files

1) All the necessary tutorial files are downloaded from:
https://github.com/MolAICal/tutorials/tree/master/018-admet

Part I. ADMET calculation of drugs

1. Install FP-ADMET model

Currently, only **Linux version** of MolAICal is supported in Part I tutorial.

1) Click and open the website: **DownloadModel**
Then go to the folder "AImodels→ADMET→FPADMET→linux64" in the open website, download a file named "fpadmet.tar.gz"

2) Move "fpadmet.tar.gz" to the folder "MolAICal-xxx/mtools". Where "MolAICal-xxx" is your decompressed root directory of MolAICal. The folder "mtools" is the targeted directory.

3) Decompress file

```
#> tar -xzvf fpadmet.tar.gz
```

And go to folder named fpadmet:

```
#> cd fpadmet
```

4) install fpadmet model

```
#> chmod +x install.sh
#> ./install.sh
```

Until now, fpadmet model is installed completely.

2. Calculate ADMET of drugs

2.1. Input file format for ADMET calculation

Users should prepare the right format for the input file. One molecule occupies one line of the input file. For one line, the first column contains molecular SMILES string, the second column includes the molecular name. And the molecular name should have one space with molecular SMILES string at least (see Figure 1).

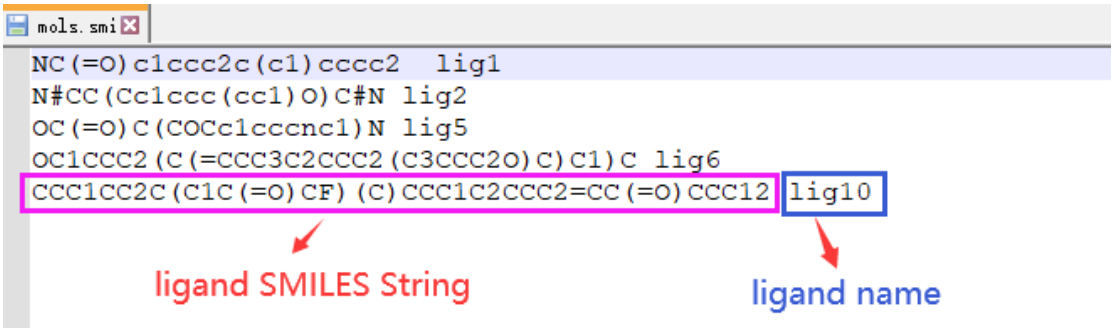

Figure 1. Content of input file

2.2. Calculate all items of ADMET from a file containing molecular SMILES

Go to 018-admet, and use the below command:

#> molaical.exe -model fpadmet -s all -i mols.smi -o results.dat

Interpretations of results: It will generate a file named "results.dat" that contains all items of ADMET of drug. Open file "results.dat", it will show 58 item results for one drug. For example, predicting Blood–brain-barrier penetration (item 2) and pKa dissociation constant (item 50) are shown below :

```
##############################
2: Blood-brain-barrier penetration
Predicted Confidence Credibility
lig1 Yes 0.93 0.32

……

50: pKa dissociation constant
Predicted quantile_0.025 quantile_0.975
lig1 5.19 -1.60 13.06
##############################
```

"lig1" is a drug or ligand name. The label "Yes" of compound "lig1" indicates that this compound is predicted to be suitable for blood brain barrier penetration. A confidence value of 0.93 indicates

that the classifier is quite certain, and the prediction is possible to be a single label. A relatively low value of credibility (0.32) suggests that the compounds like "lig1" are not sufficiently represented in the training set, and users should deal with the prediction with caution. For the regression example of item 50, a 95% prediction interval (predictions at 0.025 and 97.5 percentiles for pKa) is computed and gives a range for the predictions on individual observation. As the original paper on FP-ADMET said: narrow prediction intervals indicate a lower uncertainty associated with the prediction [1]. For more detailed interpretations of predicted results, please check the files named "ADMETModels-doc.pdf", "FP-ADMET.pdf" and "FP-ADMET-SI.pdf" in folder "doc" along with this tutorial document.

2.3. Calculating a specified item of ADMET based on a file that contains SMILES strings

Go to 018-admet, and use the below command:

```
#> molaical.exe -model fpadmet -s single -i mols.smi -n 1
```

Interpretations of results: It will generate a file named "1.dat" due to selecting task item 1. If task item 2 is chosen, it will produce a file named "2.dat". If task item 3 is chosen, it will produce a file named "3.dat"… So the prefix name of generated result file is the same as the selected task item. Open file "1.dat", the content of "1.dat" is like below:

```
##############################
Predicted Confidence Credibility
lig1 Negative 0.77 0.31
lig2 Negative 0.90 0.59
lig5 Negative 0.80 0.35
lig6 Negative 0.91 0.64
lig10 Negative 0.96 0.86
##############################
```

It only shows the calculated results of "1: Anticommensal Effect on Human Gut Microbiota" item of ADMET for all molecules. For understanding all items of ADMET, please check the MolAICal manual or the files named "ADMETModels-doc.pdf", "FP-ADMET.pdf" and "FP-ADMET-SI.pdf" in folder "doc" along with this tutorial document

Part II. Calculations of ADME and molecular properties by MolAICal

In this part, **Linux or Windows version** of MolAICal is supported for this part tutorial.

In this part tutorial, it will show the usages of MolAICal for the calculations of ADME and molecular properties.

1) If "-s" is 0, it will show all calculated molecular properties information

```
#> molaical.exe -model admepro -s 0 -i "O=C(C)Oc1ccccc1C(=O)O"
```

2) If "-s" is 1, it will show the brief general ADME information of a ligand

```
#> molaical.exe -model admepro -s 1 -i "O=C(C)Oc1ccccc1C(=O)O"
```

For example, the output is shown below:

```
# 1. Pharmacokinetics:
GI absorption: High
BBB permeant: Yes

# 2. Lipinski Rule of 5 Violations:
No violations found

# 3. Medicinal Chemistry:
PAINS filter: False
Brenk filter: True
```

3) If "-s" is 2, it will show basic molecular properties, for example: molecular weight, Number of Hydrogen Bond Donors, etc.

```
#> molaical.exe -model admepro -s 2 -i "O=C(C)Oc1ccccc1C(=O)O"
```

For more detailed items of molecular properties, please check MolAICal manual.

Reference

1. Venkatraman V. FP-ADMET: a compendium of fingerprint-based ADMET prediction models. J Cheminform. 2021;13(1):75.