Feb 23, 2023

# High Throughput Ligand Interaction Profiler

DOI
dx.doi.org/10.17504/protocols.io.4r3l27njqg1y/v1

Akshay Uttarkar[1], Shreya Girish[1], Vidya Niranjan[1]

[1]R V College of Engineering

Vidya Niranjan
R V College of Engineering





**DOI: https://dx.doi.org/10.17504/protocols.io.4r3l27njqg1y/v1**

**Protocol Citation:** Akshay Uttarkar, Shreya Girish, Vidya Niranjan 2023. High Throughput Ligand Interaction Profiler . **protocols.io https://dx.doi.org/10.17504/protocols.io.4r3l27njqg1y/v1**



**Protocol status:** Working
**We use this protocol and it's working**

**Created:** January 15, 2023

**Last Modified:** February 23, 2023

**Protocol Integer ID:** 75343

**Keywords:** high throughput ligand interaction profiler, ligand interaction study, large chemical libraries for potential drug candidate, ligand, protein interaction, key residues for ligand, large chemical library, drug discovery, interacting residue, potential drug candidate, protein, google colab, binding site

## Abstract

High Throughput Ligand Interaction Profiler (HT-LIP) is a web-based tool that runs on Google Colab, which allows users to predict ligand-protein interactions. HT-LIP can be used to screen large chemical libraries for potential drug candidates and can also provide  insights into potential ligand-protein interactions and identifies the interacting residues. HT-LIP can be useful in identifying potential binding sites and key residues for ligand-protein interactions, which can aid in drug discovery and protein-ligand interaction studies.

## Troubleshooting

## Prepare the Google Drive

1 In your Google Drive home directory, create a new folder called “data”. Inside this folder, add 3 files.
(a) protein: in PDB format
(b) ligand: in SDF format
(c) reference ligand: in PDB format

For example, the data folder created here, have files with the following names :
(a) protein: 5Y15.pdb
(b) ligand: compounds_for_ml.sdf
(c) reference: reference_ligand.pdb

## Connect to Google Drive

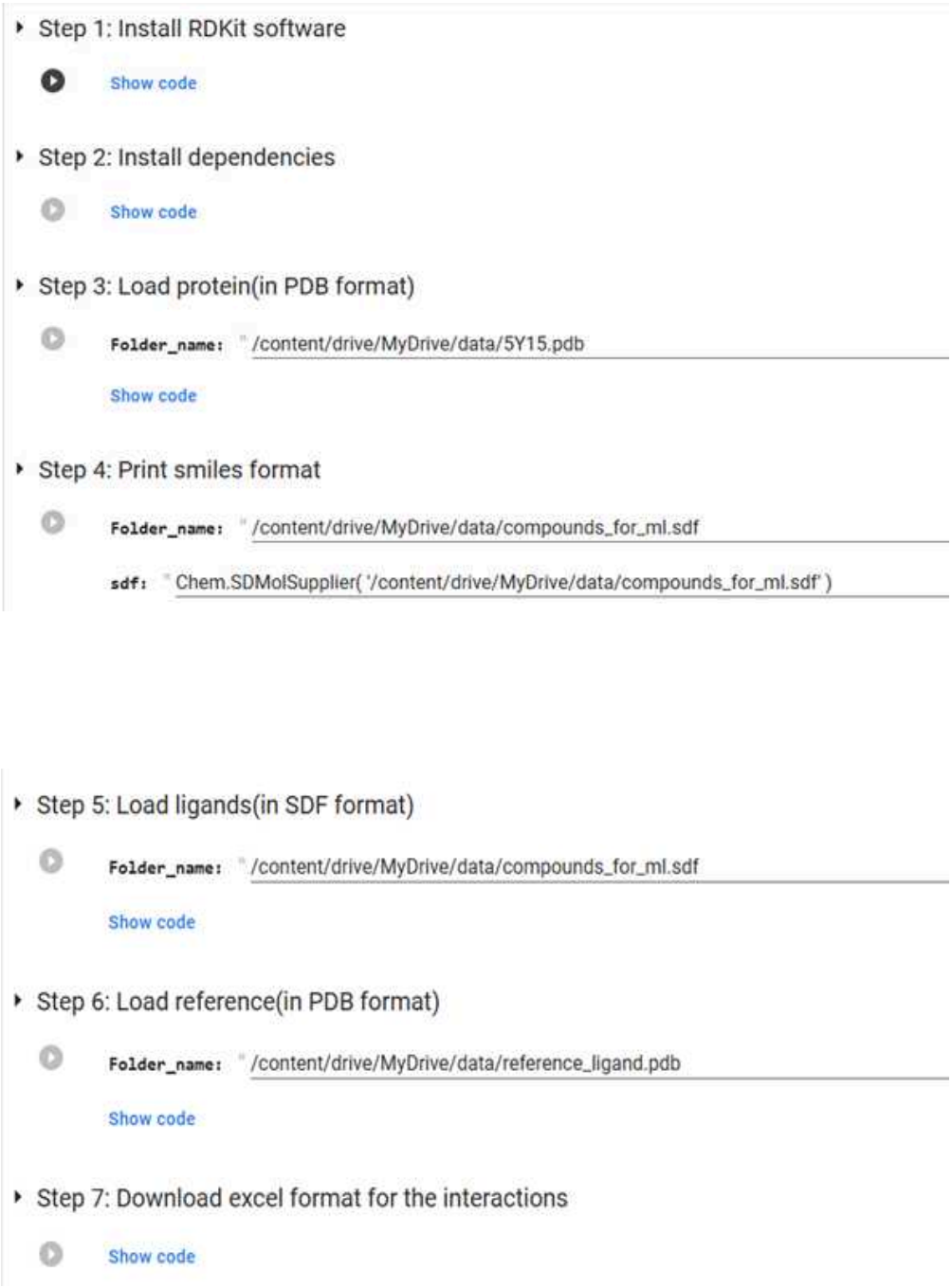
2 **Click on the URL provided below to get started.**

URL: Google Colab Notebook

Now, you will be directed towards the Google Colab Notebook. Towards the LHS side of the notebook, from the “files” section, click on the “mount drive” option. Then, a cell will be created, run the cell. Then click on “Run Anyway” and select “Connect to Google Drive” and give access. Now from the “drive” option, check whether the “data” folder created in the above step is seen in “MyDrive”.

## Run the cells

3 Below are the seven steps that need to be followed in order to get the Protein-Ligands Interactions.



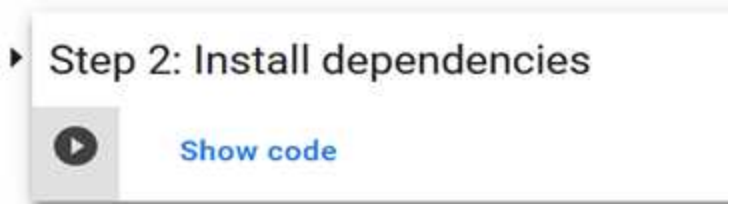

***Step 1: Install RDKit software***

RDKit is an open source Chemoinformatics & Machine Learning Software. It provides tools for different kinds of similarity search. To install RDKit, click on the run cell option beside the show code(in blue).


In a few seconds RDKit will be installed.


***Step 2: Install dependencies***

Click on the run cell option beside the show code(in blue) to install the dependencies(libraries). The required libraries will be installed in a few seconds.




### *Step 3: Load protein(in PDB format)*

The protein used here is 5Y15. To upload your protein of interest, change “5Y15” to that protein name in the “folder_name” section. Now run the cell.


### *Step 4: Print smiles format*

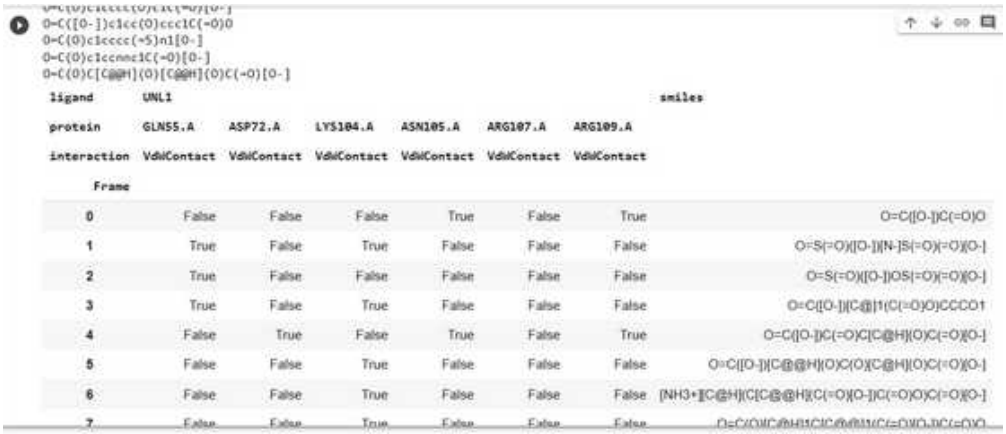
Change “compounds_for_ml” to your ligand name in the “folder_name” and "sdf" section.
Now run the cell.
Here ligands are converted from SDF to SMILES format.


The smiles format will be printed.

### *Step 5: Load ligands(in SDF format)*

Change “compounds_for_ml” to your ligand name in the “folder_name” section(same as step 4). Run the cell.


Following output will be obtained.

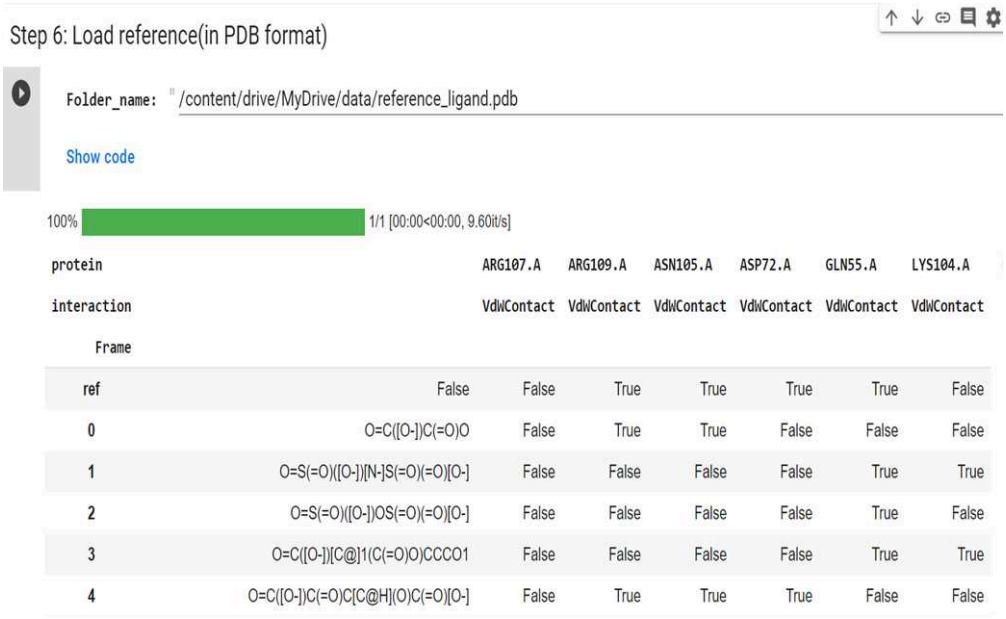


| ligand | UNL1 | | | | | | smiles |
|---|---|---|---|---|---|---|---|
| protein | GLN55.A | ASP72.A | LYS104.A | ASN105.A | ARG107.A | ARG109.A | |
| interaction | VdWContact | VdWContact | VdWContact | VdWContact | VdWContact | VdWContact | |
| Frame | | | | | | | |
| 0 | False | False | False | True | False | True | O=C([O-])C(=O)O |
| 1 | True | False | True | False | False | False | O=S(=O)([O-])[N-]S(=O)(=O)[O-] |
| 2 | True | False | False | False | False | False | O=S(=O)([O-])OS(=O)(=O)[O-] |
| 3 | True | False | True | False | False | False | O=C([O-])[C@]1(C(=O)O)CCCO1 |
| 4 | False | True | False | True | False | True | O=C([O-])C(=O)C[C@H](O)C(=O)[O-] |
| 5 | False | False | True | False | False | False | O=C([O-])[C@@H](O)C(O)[C@H](O)C(=O)[O-] |
| 6 | False | False | True | False | False | False | [NH3+][C@H](C[C@@H](C(=O)[O-])C(=O)O)C(=O)[O-] |
| 7 | False | False | True | False | False | False | O=C(O)[C@H]1CC[C@@H](C(=O)[O-])C(=O)O |

### *Step 6: Load reference(in PDB format)*

Change “reference_ligand” to your ligand name in the “folder_name” section. Now run the cell.

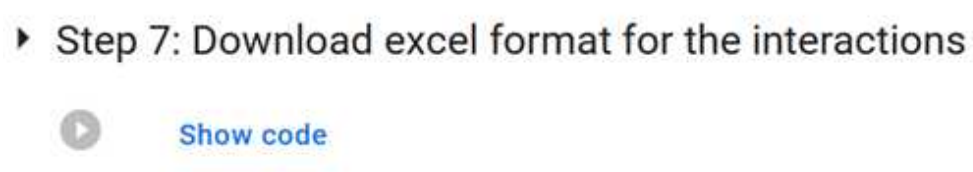

Following output will be obtained.




| protein | | ARG107.A | ARG109.A | ASN105.A | ASP72.A | GLN55.A | LYS104.A |
|---|---|---|---|---|---|---|---|
| interaction | | VdWContact | VdWContact | VdWContact | VdWContact | VdWContact | VdWContact |
| Frame | | | | | | | |
| ref | False | False | True | True | True | True | False |
| 0 | O=C([O-])C(=O)O | False | True | True | False | False | False |
| 1 | O=S(=O)([O-])[N-]S(=O)(=O)[O-] | False | False | False | False | True | True |
| 2 | O=S(=O)([O-])OS(=O)(=O)[O-] | False | False | False | False | True | False |
| 3 | O=C([O-])[C@]1(C(=O)O)CCCO1 | False | False | False | False | True | True |
| 4 | O=C([O-])C(=O)C[C@H](O)C(=O)[O-] | False | True | True | True | False | False |

### *Step 7: Download excel format for the interactions*

Download the excel file by running the cell. The downloaded excel file will be saved in the file section(LHS of colab notebook).




Citation

Bouysset C, Fiorucci S (2021)
. ProLIF: a library to encode molecular interactions as fingerprints..
Journal of cheminformatics.

https://doi.org/10.1186/s13321-021-00548-6 LINK

. J Cheminform**13**, 72 (2021). https://doi.org/10.1186/s13321-021-00548-6