Fragalysis API

1.1

Contents

Chapter 1

Main Page

This API aims to allow any user to upload pdb files from the pdb or that they have created themselves, and analyse the ligand binding using the fragalysis webpage (https://fragalysis.diamond.ac.uk).

Installation

Not recommended: pip

To install with pip, you will need to install both pymol and rdkit separately, as these don't exist as pip packages.

To install fragalysis-api with pip:

```
pip install -e .
```

How to use API

1. Set up environment
2. Download PDB files and query the PDB for structures of the same protein bound to the same or different ligands
3. Submit PDB files - you will be given a query ID
4. Push your files into fragalysis and view them online :construction:
5. Analyse the binding of ligands to your target protein! :construction:

Other functionalities that are available:

- Import pdb files directly through the API
- Query the pdb for similar structures that also have ligands bound and have the option to import these structures

Enforced rules :scroll:

- The pdb file shall not be greater than 5mb.

- The pdb filename shall not contain non English language ascii characters and shall be between 4 and 20 characters in length.

- Each pdb file for alignment shall contain the same number of chains.

- All pdb files to be aligned must be in the same directory.

- If manually selecting a file for reference it must be in the same director as all pdb files for alignment.

- PDB file must abide by best practices set out at [https://www.cgl.ucsf.edu/chimera/docs/↩ UsersGuide/tutorials/pdbintro.html#note5](https://www.cgl.ucsf.edu/chimera/docs/UsersGuide/tutorials/pdbintro.html#note5)

- In PDB file, ligands must be referred to by same code in 'HET' lines in header and in 'HETATM' lines in main part of file

1. How to set install, update a fragalysis enviroment

Starting out by initialising an environment and activating it. Clone the repository and cd to the relevant directory. Install pymol and rdkit via conda, and the other dependcies via the setup.py file:

```
conda create -n fragalysis_env2 anaconda -y
conda activate fragalysis_env2
conda install -c schrodinger pymol -y
conda install -c rdkit rdkit -y
git clone "https://github.com/xchem/fragalysis-api.git"
cd fragalysis-api/
pip install -e .
cd ..
rm -rf fragalysis-api/
```

You can check if it has installed using: `conda list`

2.1 How to download PDB files

Move to the fragalysis-api/xcimporter directory.

You will need two bits of information:

1. Your 'user ID' - this is your name followed by your protein name. For example, Anna looking at protein ATAD2 wil have username 'Anna_ATAD2'.

2. The PDB code you would like to download. For example, '6epu'.

To download the PDB file

```
python pdbimporter.py -id [user_id] -pdb [pdb code]
```

In our example, this would be: `python pdbimporter.py -id Anna_ATAD2 -pdb 6epu`

Alternatively, you can upload your own PDB files.

2.2 How to query the PDB file

Note: you don't need to manually download a pdb file before querying the PDB for structures of the same protein

Similar to downloading PDB files, however initially all you need is the PDB code and the chain you would like to query. The command is

```
python pdbquery.py -pdb [pdb code] -chain [required chain]
```

The API will then query the PDB for structures of the same protein that also have ligands bound. You will be asked if you would like to see a list of these structures and ligands, and if you would like to download all of these pdb files in bulk. If you choose to download these pdb files, you will be asked for your user id. This is as before: your name followed by the protein name (e.g. Anna_ATAD2).

3.How to submit PDB files for conversion to a fragalysis friendly format (fff)

Once you have the files downloaded, they need to be processed before they can be visualised in fragalysis. This is done using

```
python xcimporter -id [user id]
```

Default directories are used. These can however be changed by using `-i [input directory]` or `-o [output directory]` if this is required.

The terminal will let you know when the conversion has been successful and if there are any files that have been found to be incompatible with the API. We are working to minimize any incompatibilities.

Who we are?

We are the 'Fragment 5', a group of students at the University of Oxford.

- Anna
- Maranga
- George (Joff Boff)
- Tobias
- Alister

We are looked after by Rachael (Queen).