

PDBmutator

PDBmutator is a GUI enabled script to batch run swapna, swapaa and delete residue commands in chimera on a PDB file. It is tested on UCSF Chimera 1.14 (<https://www.cgl.ucsf.edu/chimera/>) installed Linux Mint 19.3. It also requires gedit, python kivy, perl-tk and pdb4amber (from AmberTools 18 <https://ambermd.org/>) or directly from GitHub (<https://github.com/Amber-MD/pdb4amber>) pdb4amber is only needed if an alignment file is used as the source of the PDB variants.

To run:

```
python PDBMUTATOR.py
```

To install dependencies:

```
sudo apt-get install gedit
sudo apt-get install perl-tk
sudo apt-get install python-tk
sudo apt-get install python-kivy
```

To install pdb4amber from AmberTools:

```
conda install ambertools -c conda-forge
```

To install pdb4amber from GitHub:

```
pip install git+https://github.com/amber-md/pdb4amber
```

To get UCSF Chimera:

<https://www.cgl.ucsf.edu/chimera/download.html>

Within a selected range of a given chain, mutator will create all possible base substitutions (for DNA chain) or all possible amino acid replacements (protein chain) and deposit all possible mutant PDB files in a directory. PDBmutator can also model all amino acid replacements present in 1 or more genetic variants aligned to a reference sequence of a CLUSTAL (.aln) file. This can be used to create structures for ortholog or paralog sequences provided the PDB file sequence

You must install perl, python and UCSF Chimera. Type 'python PDBMUTATOR.py' at terminal or command line to launch the GUI. Select the type of procedure to run (variants from protein region, DNA region, or protein alignment file and follow the directions on each GUI and Linux terminal. Example files for each procedure are given on each GUI. This includes creating all 1521 possible amino acid replacements in ubiquitin (PDB: 1ubq), creating all 49 DNA substitutions possible in the DNA bound to TATA binding protein (1cdw_bound.pdb), OR creating the 8 lysozyme paralogs in the CLUSTAL alignment (1rex_align.aln) from the PDB file 1rex.pdb.

NOTE: make sure path to chimera executable is correct for your machine on the GUI

NOTE: To model all AA replacements between position 5 and 15 in 1ubq.pdb, simply open GUI, enter these as Start position and Stop position and then select chain type = amino acid and data input to create all possible mutants and then execute mutations.

NOTE; To model all the AA replacements in the CLUSTAL alignment given (1rex_align.aln), enter the file name (without the .aln extension), and the pdb ID = 1rex then click 'process variants in alignment' and enter 8 when prompted to enter the number of variants at the terminal, then click 'execute replacements and deletions'. If deletions are present, they will create gaps in the structure that are repaired by manually removing the TER lines created by each deletion from the PDB file, and then running the 'seal deletions and finalize' button. This button

uses pdb4amber to repair and renumber chains in the structure.

NOTE: For DNA mode, the script assumes numbering of nucleotides on forward chain increases (e.g. 1cdw chain b) while numbering on the reverse chain decreases (e.g. 1cdw chain c). To model all DNA substitutions at all 16 sites on the DNA segment in this file, select start position = 1, stop position = 16 and start position on reverse chain = 116

ALSO NOTE: script creates the 3 possible variants for a DNA substitution at a any given site of mutation and all 20 possibilities an amino acid variant.