

Zawartości katalogu „Projekt”:

- podkatalog „*PDB_structures*” – zawiera listę PDB ID („*PDB_IDs_batch.txt*”) użytych struktur białek oraz skrypt pobierający udostępniany przez bazę danych RCSB PDB („*batch_download.sh*”). Podkatalog służy również jako repozytorium pobranych struktur PDB,
- podkatalog „*Domain_architectures*” – repozytorium na pliki XML zawierające informacje o architekturach domenowych analizowanych białek,
- skrypt „*find_domains.py*” – autorski skrypt wyszukujący domeny białkowe wiążące wskazane ligandy w zadanym zbiorze struktur białkowych,
- moduł „*bioinf_tools.py*” – zestaw autorskich narzędzi bioinformatycznych użytych z poziomu skryptu „*find_domains.py*”,
- dokumentacja modułu „*bioinf_tools.py*” – plik „*bioinf_tools.html*”,
- plik konfiguracyjny „*config*” – przykładowy plik konfiguracyjny skryptu „*find_domains.py*”,
- plik tekstowy „*log.txt*” – (plik log) szczegółowe wyniki analizy przeprowadzone dla zbioru struktur uzyskanego na podstawie listy PDB ID z „*PDB_IDs_batch.txt*”,
- plik tekstowy „*gen_stat.txt*” – statystyka dotycząca wszystkich znalezionych domen we wspomnianym zbiorze,
- plik tekstowy „*matched_stat.txt*” – statystyka dotycząca domen wiążących wskazane ligandy we wspomnianym zbiorze,
- plik tekstowy „*where.txt*” – skład aminokwasowy znalezionych miejsc wiążących w kontekście domen białkowych znalezionych we wspomnianym zbiorze.

Przykłady użycia skryptu „*find_domains.py*”:

- aby uzyskać wyniki podobne do tych zaprezentowanych w plikach tekstowych: „*log.txt*”, „*gen_stat.txt*”, „*matched_stat.txt*”, „*where.txt*” należy użyć komendy:
`./find_domains.py --config config --verbose`
- możliwe jest wybranie ilości struktur do testowego uruchomienia skryptu:
`./find_domains.py --config config --verbose --batch_s 73`
- aby uzyskać pliki log bez wyszczególnionych miejsc kontaktu ligand – białko, należy użyć komendy: `./find_domains.py --config config`

Plik konfiguracyjny:

<code>pdb_directory: PDB_structures</code>	- ścieżka wymagana
<code>xml_directory: Domain_architectures</code>	- ścieżka wymagana
<code>log_file: log.txt</code>	- ścieżka opcjonalna
<code>general_statistics_file: gen_stat.txt</code>	- ścieżka opcjonalna
<code>matched_statistics_file: matched_stat.txt</code>	- ścieżka opcjonalna
<code>where_ligand_file: where.txt</code>	- ścieżka opcjonalna
<code>ligands: {CLR MHQ 94R ERG LNP LAN VD3 DVE HC9 HC3 C01 C3S B81 Y01 20B CLL 5JK HCR HC2 HCD 0GV YK8 2DC PLO AND XCA K2B}</code>	- wymagane podanie skrótu PDB co najmniej jednego liganda