

*Instytut Informatyki i Matematyki Komputerowej UJ,
opracowanie dr Jacek Śmietański*

Modelowanie homologiczne struktur białkowych

Opracowano na podstawie:
<http://godziklab.org/SSBC/modeling.html>

1. Przygotowanie środowiska

Będziemy korzystać z narzędzia SwissModel: <http://swissmodel.expasy.org/>

Jeśli pracujesz na własnym komputerze, możesz ściąnąć i zainstalować **Swiss-PdbViewer** (<http://spdbv.vital-it.ch/>) w celu lepszej wizualizacji modelowanych struktur.
(Nie jest to konieczne, wymodelowana struktura jest również wyświetlana w przeglądarce)

2. Konstruujemy pierwszy model: „łatwy cel”

Będziemy modelować strukturę białka na podstawie jego sekwencji aminokwasowej, wykorzystując jako szablon jeden z jego homologów. Utworzony model porównamy z rzeczywistą strukturą modelowanego białka.

Pierwsze modelowane białko nazwaliśmy „łatwym celem”, ponieważ znalezienie dla niego szablonu nie powinno nastręczyć problemów.

Białko: Ligand Binding Domain of the Vitamin D Nuclear Receptor
Plik fasta z sekwencją tego białka znajduje się w pliku: **lab14-seq1.fasta**

2a. Modelowanie w pełni automatyczne

Modelowanie w pełni automatyczne: Modelling -> myWorkspace -> BuildModel.
Modelowanie może trochę potrwać. Po zakończeniu zapisz pierwszy wygenerowany model.
A w międzyczasie w nowym oknie zacznij realizować kolejne zadanie.

2b. Ręczna identyfikacja szablonu

Znajdź szablon za pomocą BLASTa.
Szukamy w bazie Protein Data Bank, za pomocą algorytmu iteracyjny PSI-BLAST.

Jakie są kryteria wyboru szablonu?

- która struktura na najwyższy / najniższy procent identyczności z sekwencją zapytania?
- jak długie jest dopasowanie pomiędzy pierwszym / ostatnim potencjalnym szablonem?

- które dopasowanie ma największą / najmniejszą E-wartość? Która E-wartość jest najlepsza?

Dopasuj wybrany szablon do sekwencji zapytania.

- zapisz sekwencję aminokasową wybranego szablonu. Dopasuj programem Clustal Omega: <http://www.ebi.ac.uk/Tools/msa/clustalo/>
- uzyskany rezultat może być użyty bezpośrednio w programie SwissModel.

3. Porównanie modelu z rzeczywistą strukturą.

Możesz w tym celu wykorzystać np. narzędzie dostępne na stronie.

<http://fatcat.burnham.org/fatcat-cgi/cgi/fatcat.pl?-func=pairwise>

A jeśli wolisz rozwiązać problem programistycznie, to w biopythonie jest moduł *Superimposer* (<http://biopython.org/DIST/docs/tutorial/Tutorial.html#htoc177>)

Struktura z przykładu powyżej to 2HCD.A

4. Modelowanie trudnej struktury

Przeprowadź podobny schemat modelowania dla białek, dla których trudno jest znaleźć dobre homologi. Jako wejście podajemy oczywiście sekwencję, a szablonu szukamy w bazie PDB. Pełne struktury udostępnione są aby możliwe było oszacowanie jakości modelu.

(pliki *lab14_protein2.pdb*, *lab14_protein3.pdb* – sekwencje trzeba wyłuskać z plików)

Rozwiązanie zadanie prześlij mailem do niedzieli, **28.01.2019** włącznie, na adres:

jacek.smietanski@ii.uj.edu.pl

Temat wiadomości proszę opatrzyć przedrostkiem **[Bio] Lab 14**. Rozwiązaniem ma być **tylko jeden plik** – dokument PDF (oraz skrypt pythona, jeżeli część rozwiązań zaimplementowałeś). Proszę o nazwanie pliku wg schematu: **Imie.Nazwisko.14.pdf** (analogicznie Imie.Nazwisko.14.py)