

Wprowadzenie do Bioinformatyki

zajęcia laboratoryjne 6

Narzędzia bioinformatyki strukturalnej

1 Wprowadzenie

Celem zajęć jest zapoznanie się z narzędziami do modelowania struktur przestrzennych oraz zrozumienie, że jakość otrzymanej struktury III-rzędowej opiera się na jakości modelu struktury II-rzędowej. Do zaproponowania struktury II-rzędowej wykorzystane zostaną dwa narzędzia RNAstructure i CentroidFold, natomiast do predykcji struktury III-rzędowej wykorzystany zostanie RNAComposer.

2 Wykorzystywane narzędzia i bazy danych

1. **RCSB PDB** – Największe ogólnodostępne repozytorium przestrzennych struktur białek oraz kwasów nukleinowych, które są dostarczane za pomocą metod eksperymentalnych takich jak krystalografia rentgenowska, spektroskopia NMR czy mikroskopia elektronowa. Zasoby dostępne na stronie <https://www.rcsb.org>.
2. **RNAComposer** – zaawansowane narzędzie do modelowania struktur RNA. Jego celem jest budowanie modeli 3D w oparciu o ich struktury drugorzędowe, przy użyciu słownika RNA FRABASE dotyczącego struktury drugorzędowej RNA i elementów struktury trzeciorzędowej. Narzędzie dostępne jest pod adresem: <https://rnacomposer.cs.put.poznan.pl>. RNAComposer wykorzystuje różne narzędzia do przewidywania struktury drugorzędowej RNA. W ramach zajęć wyróżnione zostaną dwa z nich: **CentroidFold** (<http://rtools.cbrc.jp/centroidfold/>) i **RNAstructure** (<https://rna.urmc.rochester.edu/RNAstructureWeb/>).
3. **PyMol** – system wizualizacji molekularnej dostępny do pobrania na stronie: <https://pymol.org>.

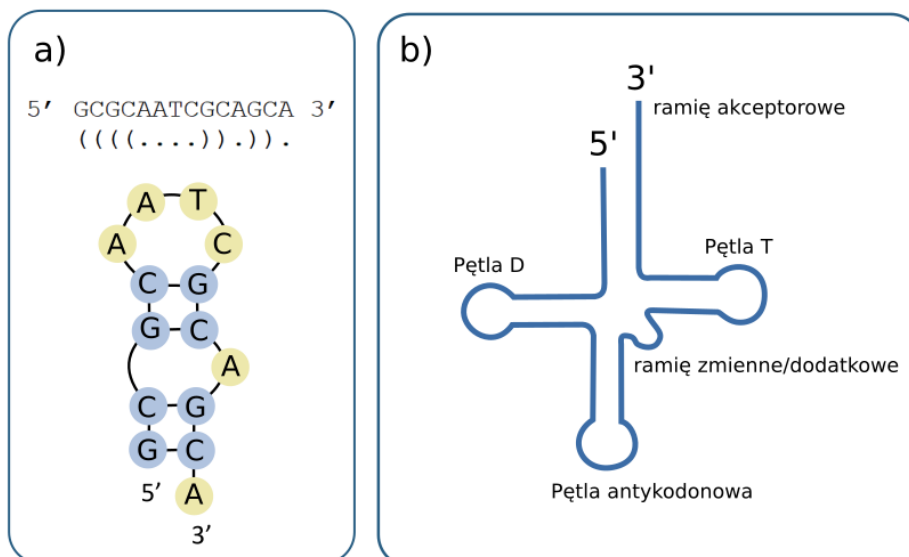
3 Zadania do wykonania

Odpowiedzi do poniższych zadań zapisz w pliku tekstowym, jeśli zadanie wymaga to zapisz grafiki lub zachowaj do okazania własnoręczny rysunek.

1. Przyjrzyj się poniższej notacji kropkowo-nawiasowej (Rysunek 1) oraz strukturze drugorzędowej tRNA przedstawionej na Rysunku 2b), a następnie na ich podstawie utwórz klasyczną reprezentację graficzną i zaznacz kluczowe regiony. Przykład klasycznej reprezentacji graficznej odpowiadającej pewnej sekwencji i notacji kropkowo-nawiasowej umieszczono na 2a).

```
5' GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA 3'  
(((((((..(((.....))))).(((.....)))).....((((.....)))))))))).....
```

Rysunek 1: Struktura II-rzędowa tRNA w reprezentacji kropkowo-nawiasowej ([link do pobrania](#)).



Rysunek 2: Schemat struktury II-rzędowej tRNA.

2. Wykorzystaj bazę RCSB PDB, aby znaleźć strukturę tRNA o ID 1EHZ. Dla znalezionej struktury podaj następujące informacje: klasyfikacja, organizm, metoda otrzymania oraz rozdzielczość. Ponadto zapisz jak wygląda struktura pierwszorzędowa, jej długość i nazwę molekuly (sekcja macromolecules). Wykorzystaj możliwość wizualizacji “1D-3D View”, czy umiesz określić gdzie znajdują się kluczowe regiony, zgodne z Rysunkiem 2?

3. Wykorzystaj PyMOL do wizualizacji struktur przestrzennych (dostępny na linuxie).

- Wczytaj strukturę przestrzenną (1EHZ) używając funkcji fetch. Przykład użycia tej funkcji sprawdź na PyMOLWiki. Alternatywne podejście, pobierz strukturę w formacie .PDB z bazy RCSB PDB i wczytaj do PyMOLa (file/open...).
- Dostosuj wizualizację poprzez ukrycie niektórych elementów. Użyj symbolu “H” po prawej stronie, aby ukryć wodę (water) oraz kule (spheres).
- Dostosuj kolor struktury, tak aby był jednolity, użyj symbolu “C”.
- Wyświetl sekwencję analizowanej struktury, wykorzystując symbol “S”, dostępny w prawym dolnym rogu.
- Za pomocą notacji kropkowo nawiasowej z pierwszego punktu oraz wyświetlonej sekwencji, wyznacz pętlę D, pętlę T oraz pętlę antykodonu, każdą z nich oznacz innym kolorem.

Przykładowo:

GCGGAUUUAGCUCAGUUGGGAGAGCG
 (((((((((..(((.....)))))))).

Na podstawie powyższego fragmentu notacji kropkowo-nawiasowej, można zauważyć iż pętla obejmuje zakres od 14 do 21 nukleotydu (fragment AGUUGGGA). Zaznacz ten fragment w PyMOLu (wybierając odpowiednie nukleotydy w sekwencji) i prawym przyciskiem myszy wybierz dowolny kolor, tak aby oznaczony fragment się wyróżniał. Analizując Rysunek 2b) można zauważyć, że omawiany przykład dotyczy pętli D (pierwsza pętla od końca 5’).

- Na podstawie odpowiednio wyznakowanej struktury 3D oraz schematu 2D (patrz Rysunek 2), odpowiedz które pętle są skrócone do siebie. Ułożenie to wpływa na strukturę przestrzenną i tworzy charakterystyczny kształt litery “L”. Zauważ, że struktura drugorzędowa przypomina czterolistną koniczynę.

4. Wykorzystaj sekwencję (strukturę pierwszorzędową) uzyskaną w punkcie 2 i zamodeluj strukturę przestrzenną w oparciu o narzędzie RNAComposer. Użyj w tym celu dwóch różnych metod predykcji struktur drugorzędowych: metody CentroidFold i metody RNAStructure. Jako wyniki zapisz dwa uzyskane pliki PDB oraz struktury drugorzędowe w notacji kropkowo-nawiasowej. Analizując notację kropkowo-nawiasową odpowiedz ile pętli zawiera struktura 2D wygenerowana przez RNAStructure, a ile przez CentroidFold? Czy te struktury są do siebie podobne?
5. Wykorzystaj uzyskane w punkcie 4 struktury przestrzenne i otwórz je w PyMolu, obok struktury 1EHZ. Każdą z nowo dodanych struktur oznacz innym kolorem. Następnie nałóż na siebie strukturę uzyskaną na podstawie metody CentroidFold i metody RNAStructure - użyj funkcji align. Przykład użycia tej funkcji sprawdź na PyMOLWiki. Co możesz powiedzieć o tych strukturach? Czy zastosowane metody pozwalają uzyskać taki sam model 3D?
6. Wykorzystaj strukturę 1EHZ jako referencyjną i nałóż na nią strukturę uzyskaną na podstawie metody metody CentroidFold, a w następnej kolejności wykonaj nałożenie 1EHZ ze strukturą uzyskaną z metody RNAStructure (funkcja align). W danym momencie wizualizacji, możesz ukryć niepotrzebną strukturę dla czytelności wizualizacji ("H/everything" dla wybranej struktury). Która z badanych metod pozwoliła na uzyskanie modelu zbliżonego do struktury ustalonej eksperymentalnie?