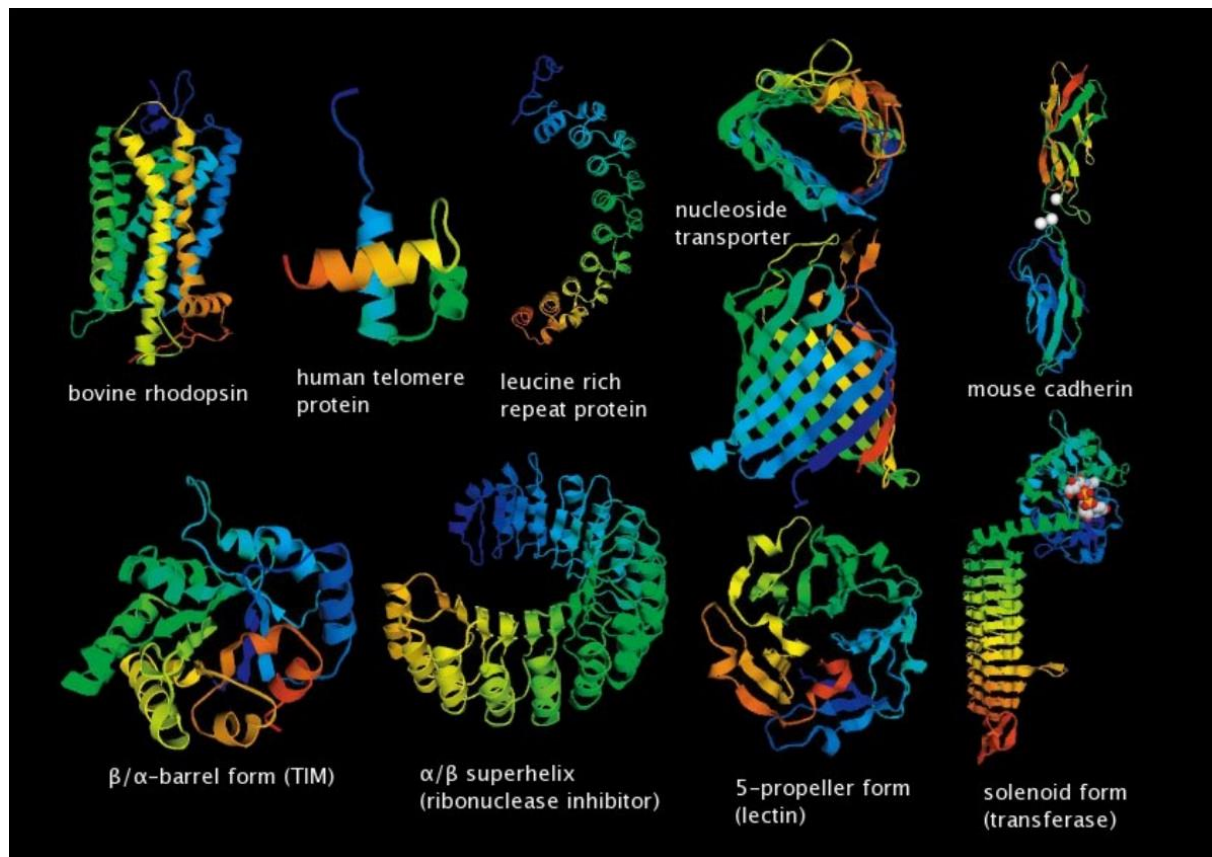


*Instytut Informatyki i Matematyki Komputerowej UJ,
opracowanie Ewa Matczyńska, Jacek Śmietański*

Przewidywanie struktur drugorzędowych białek

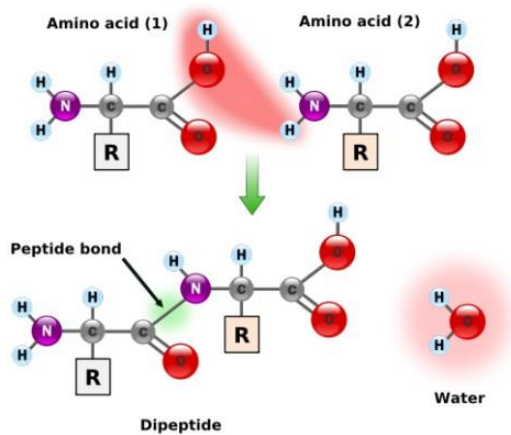
1. Białka



Rysunek 1: Różne typy białek.

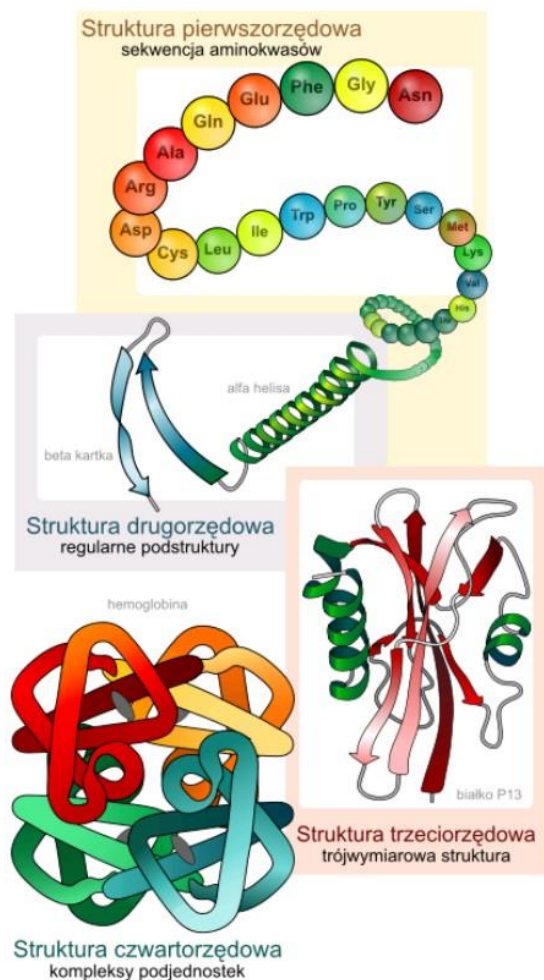
Jak pamiętamy, białka są zbudowane z aminokwasów, których mamy 20 rodzajów oznaczanych przez litery alfabetu: A R N D C E Q G H I L K M F P S T W Y V. Proces budowy białka polega na sekwencyjnym dołączaniu kolejnych aminokwasów, które tworzą coś w rodzaju łańcucha. Na skutek oddziaływań fizycznych łańcuch aminokwasów zwija się i formuje strukturę przestrzenną białka. Sekwencja aminokwasów determinuje strukturę 3D białka, a co za tym idzie, jego funkcję. Białka są kluczowe dla działania naszego organizmu:

- katalizują reakcje chemiczne (enzymy)
- odpowiadają za układ odpornościowy
- regulują pracę organizmu (hormony, np. insulina)
- są receptorami światła, neuroprzekaźników
- pełnią funkcje transportowe (np. hemoglobina)
- budują mięśnie, ścięgna, włosy
- itd...



Rysunek 2: Wiązanie peptydowe.

Sekwencja aminokwasów nazywana jest strukturą pierwszorzędową białka. Łańcuch aminokwasów powstaje poprzez dołączanie kolejnych, określonych przez sekwencję genu aminokwasów. Pomiedzy aminokwasami wytwarza się tzw. wiązanie peptydowe. Do tej pory zajmowaliśmy się właśnie sekwencjami, natomiast teraz będziemy próbować analizować strukturę przestrzenną białek.



Rysunek 3: Poziomy uporządkowania przestrzennego.

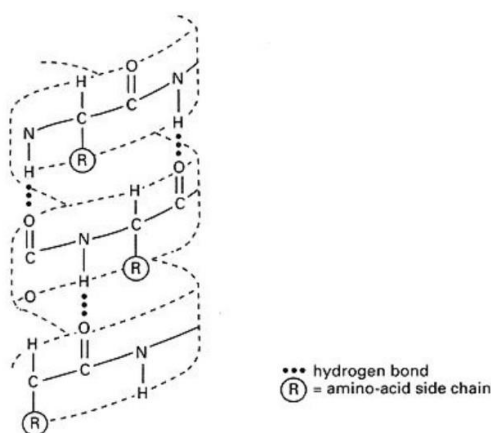
2. Struktura drugorzędowa białek

Struktura drugorzędowa białka to przypisanie określonym aminokwasom pewnych klas regularności przestrzennej, jest to poziom wyżej w opisie organizacji białka niż sekwencja, natomiast w dalszym ciągu nie wybiegamy za bardzo w przestrzeń gdyż możemy wprost przypisać strukturę drugorzędową do łańcucha aminokwasów. Upraszczając sytuację będziemy rozróżniać 3 podstawowe klasy struktur drugorzędowych:

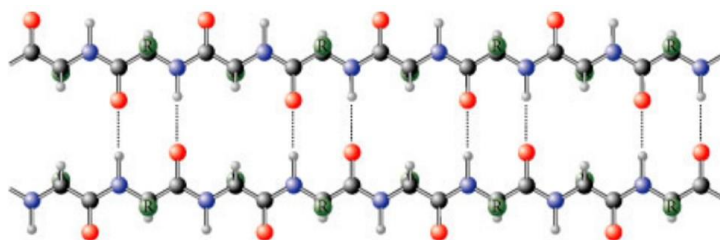
- α helisę (*helix*)
- β arkusz (*sheet*)
- zwrot (*turn*)

oraz elementy nieuporządkowane (*random coil*).

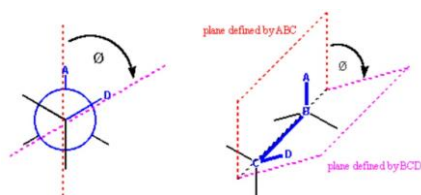
Helisy i arkusze to struktury, które są stabilizowane specjalnym rodzajem wiązania – wiązaniem wodorowym. Wspominaliśmy już o nim przy okazji budowy helisy DNA, która również jest stabilizowana wiązaniem wodorowym. Jest to wiązanie dość słabe, ale jeśli wiązań jest odpowiednio dużo, będą one dobrze stabilizować cząsteczkę. Wiązanie wodorowe polega na tym, że atom wodoru jest współdzielony przez dwa inne atomy, np. w podstawowym rodzaju helisy – helisie α , atom wodoru z grupy N-H jest częściowo dzielony z grupą C=O. W arkuszu β sytuacja jest podobna.



Rysunek 4: α helisa – schemat.



Rysunek 5: Arkusz β – schemat.



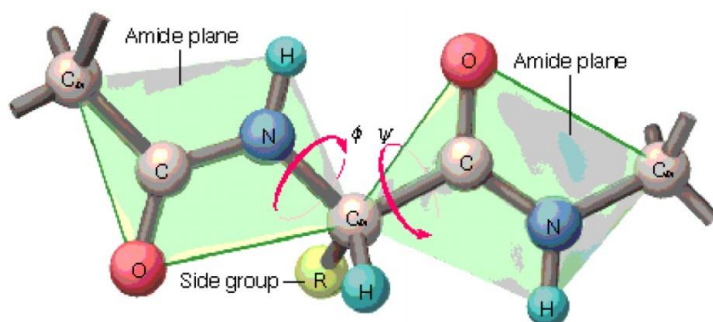
Rysunek 6: Kąty dwuścienne.

Do opisu struktury drugorzędowej przydają się kąty dwuścienne, które w strukturze białka będziemy określać jako

ϕ - opisany przez atomy C-N-C $_{\alpha}$ -C,

ψ - opisany przez atomy N-C $_{\alpha}$ -C-N,

gdzie C to węgiel związany z tlenem, C $_{\alpha}$ to węgiel do którego przyłączona jest reszta aminokwasowa. Specyfika wiązania peptydowego powoduje, że w jego płaszczyźnie nie może dojść do obrotów, możemy to sobie wyobrazić, jakbyśmy mieli w tym miejscu pewne sztywne płytki, które można obracać tylko na wiązaniach przy węglu C $_{\alpha}$.



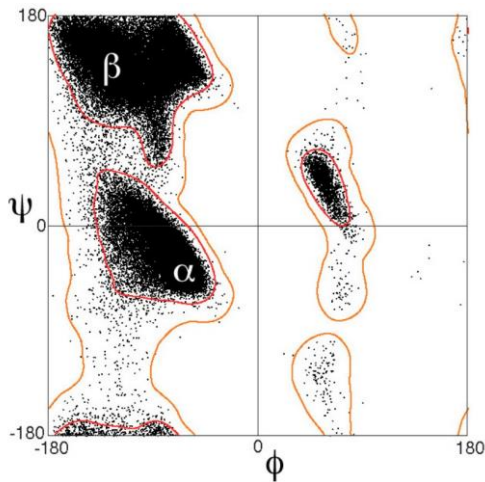
Rysunek 7: Kąty ϕ , ψ .

Zadanie 1.

Znajdź w bazie PDB strukturę **1ERA**, jest to struktura erabutoksyny, bardzo silnej toksyny zawartej w jadzie węża koralowego. Uruchom wizualizację za pomocą appletu JSMol.

- poobracaj i przybliż cząsteczkę
- poeksperymentuj z menu po prawej stronie, które kontroluje wyświetlanie cząsteczki
- wyświetl tylko łańcuch główny
- zmierz kąty torsyjne (*dihedrals*) dla wybranego aminokwasu
- zapisz obraz do pliku graficznego
- wyświetl strukturę w trybie cartoon. Jakie struktury drugorzędowe występują w tym białku?

Kąty ϕ i ψ są bardzo ważne przy opisie struktury białka, dobrze widać to na wykresie nazywanym *Mapą Ramachandrana*, która pokazuje zależności tych kątów. Okazuje się, że tylko niektóre pary kątów są możliwe. Ponadto położenie kątów na mapie świadczy o przybranej strukturze drugorzędowej.



Rysunek 8: Mapa Ramachandrana sporządzona na białkach o znanych strukturach. Zaznaczone obszary struktur drugorzędowych

3. Przewidywanie struktur drugorzędowych

Zadanie 2.

Przeanalizuj następujące pary fragmentów białek o identycznych sekwencjach:

- 1) 1IAL (292-300), 1PKY (413-421): KGVVPQLVK
- 2) 1CGU (121-127), 1BGL (835-841): LITTAHA
- 3) 1EFV (119-124), 1P04 (114-119): LLPRVA

1. Pobierz struktury przestrzenne wskazanych fragmentów z bazy PDB (implementacja w pythonie).
2. Na stronie <http://npsa-pbil.ibcp.fr/> udostępniono szereg narzędzi do analizy białek. Interesuje nas przewidywanie struktury drugorzędowej. Przetestuj metodę statystyczną GOR IV oraz bazującą na hierarchicznych sieciach neuronowych metodę HNN. Porównaj uzyskane wyniki między sobą oraz strukturą drugorzędową otrzymaną na podstawie struktury (informacje dostępne na stronie PDB – sekcja *sequence*: tu porównaj struktury otrzymane metodami DSSP, STRIDE oraz strukturę zadeklarowaną przez autorów).

Zadanie 3. (4pkt)

Zweryfikuj następującą hipotezę:

„Liczba elementów struktury drugorzędowej w losowej sekwencji aminokwasowej jest istotnie mniejsza lub ich długość jest istotnie krótsza niż dla struktury drugorzędowej rzeczywistego białka zbudowanego z tych samych aminokwasów”.

Napisz program generujący losowe sekwencje białkowe o zadanym składzie (w praktyce, zwracający n losowych permutacji dla zadanej na wejściu liczby n oraz danej sekwencji białkowej) i zaimplementuj algorytm weryfikujący powyższą hipotezę.

Do przewidywania struktury drugorzędowej wykorzystaj uważany za jeden z lepszych, algorytm *Jpred* (<http://www.compbio.dundee.ac.uk/jpred/>). Za pomocą biblioteki *jpredapi* (<https://github.com/MoseleyBioinformaticsLab/jpredapi>) możesz wykonywać zapytania bezpośrednio z poziomu pythona.

Algorytm przetestuj na kilku sekwencjach prawdziwych białek o znanej strukturze drugorzędowej. Czy hipoteza się potwierdziła?

Rozwiązanie zadania 3 prześlij mailem do piątku, **28.01.2020** włącznie, na adres:

jacek.smietanski@ii.uj.edu.pl

Temat wiadomości proszę opatrzyć przedrostkiem **[Bio] Lab 13**. Rozwiązaniem ma być **tylko jeden plik** – skrypt zgodny z Pythonem w wersji 3.x, zawierający wszystkie niezbędne funkcje oraz procedurę wykonawczą. Proszę o nazwanie pliku wg schematu: **Imie.Nazwisko.13.py**