

## Ćwiczenie 08

### Modelowanie białek. Analiza właściwości fizykochemicznych. Analiza porównawcza struktury białka.

Kornelia Polok

#### 1. Modelowanie białek

##### 1.1. Sekwencjonowanie genomów a identyfikacja białek

➔ Wraz ze wzrostem liczby zsekwencjonowanych genomów, pojawia się coraz więcej białek, których istnienie przewiduje się na podstawie sekwencji nukleotydowej. Oznacza to, że nie znamy struktury danego białka ani funkcji. W tej sytuacji wykorzystujemy liczne algorytmy przewidujące strukturę przestrzenną, centra aktywne i potencjalne interakcje. Generalnie wyróżniamy trzy strategie pozwalające na przewidywanie struktury białka:

- modelowanie porównawcze (homology modeling) czyli tworzenie modelu przestrzennego na podstawie białek pokrewnych;
- analiza struktury przestrzennej (fold recognition, threading), która wykorzystuje podobieństwo struktur wyższego rzędu przy braku podobieństwa sekwencji;
- modelowanie *ab initio* w sytuacji braku homologii ze znanymi białkami.

1.1.1. Modelowanie na podstawie sekwencji zwykle wymaga przepisania sekwencji nukleotydowej na sekwencję aminokwasową. Na ogół nie wiemy, która nić jest sensowna, kodująca. Dlatego przepisując sekwencję nukleotydową na aminokwasową wykorzystujemy wszystkie możliwe ramki odczytu.

- **ORF (Open Reading Frame):** sekwencja nukleotydów z kodonem start i stop, która potencjalnie koduje białko (Rys. 1.1.1).
- **Nić sensowna:** nić kodująca, zapisana od końca 5' do 3'. Nić sensowna ma identyczną sekwencję jak mRNA. Jednakże jest to pojęcie umowne, gdyż dla sekwencji o nieznanym funkcjach nie wiadomo, która nić jest sensowna, a która nie aż do momentu zidentyfikowania produktu. W bazach danych zapisuje się sekwencje nukleotydowe od 5' do 3' i są one domyślnie nicią sensowną.
- **Nić antysensowna:** nić niekodująca, ale będąca matrycą do syntezy białka.

**Otwarta ramka odczytu (ang. ORF - open reading frame) to ciąg nukleotydów kodujących białko wraz z kodonem start i stop.**

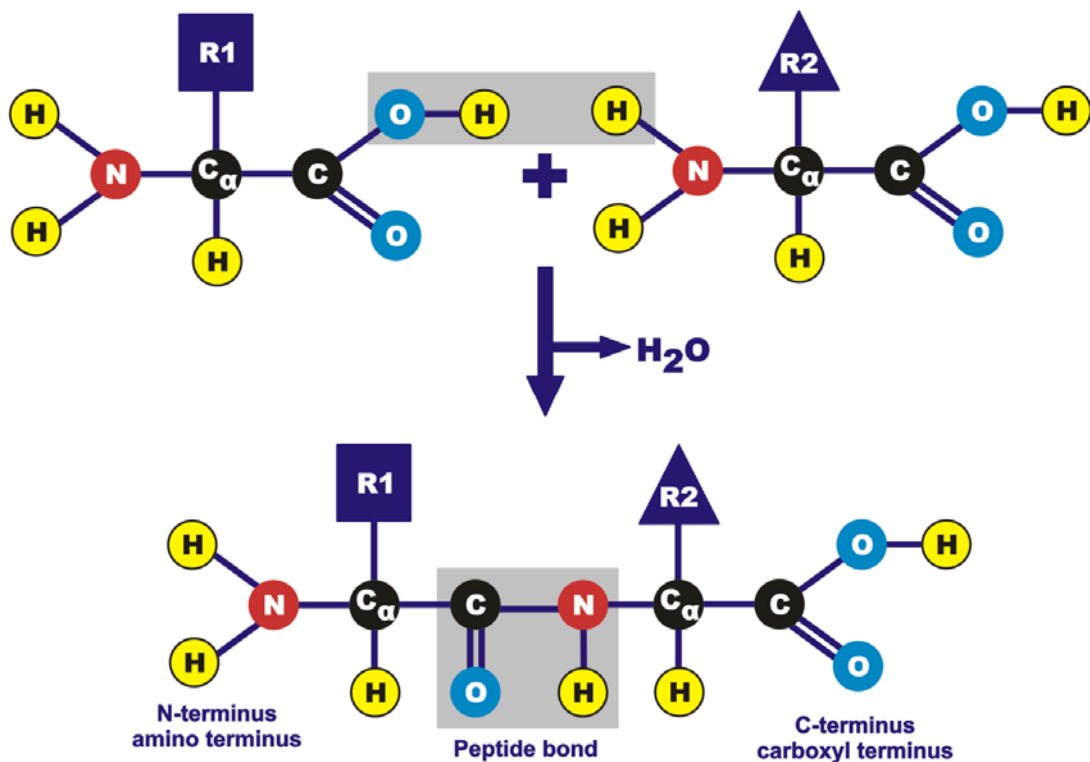
- ORF identyfikuje się w wyniku sekwencjonowania genomów, gdy nie są znane białka kodowane przez dany fragment DNA. Wówczas przewiduje się sekwencję białek na podstawie sekwencji DNA.
- Przyjmuje się, że ORF to najdłuższy odcinek DNA z trójką nukleotydów dla kodonu start i stop.
- „Zidentyfikowano 10 ORF” oznacza, że zidentyfikowano 10 obszarów będących potencjalnymi genami.

Każdy region DNA może być matrycą dla 6 typów mRNA różniących się początkiem odczytu.

**Ramki odczytu różnych genów mogą się czasami nakładać. Są to tzw. nakładające się geny.**

Rys. 1.1.1. Definicja i lokalizacja ORF.

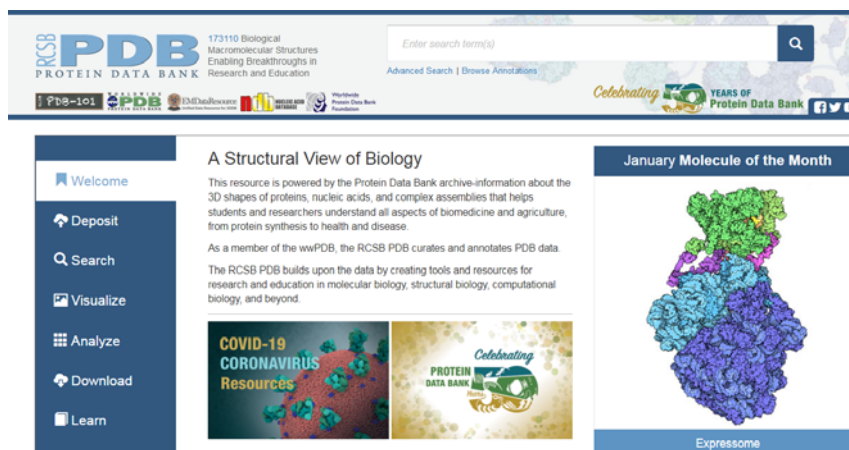
1.1.2. Strukturę pierwszorzędową białek tworzy sekwencja aminokwasów. Aminokwasy połączone są wiązaniem peptydowym, które powstaje między grupą karboksylową i aminową z wydzieleniem cząsteczki wody. W efekcie lewy koniec białka ma wolną grupę aminową,  $\text{NH}_2$ . Koniec ten określa się mianem N-końca (N-terminus). Natomiast prawy koniec białka ma wolną grupę karboksylową,  $\text{COOH}$ . Koniec ten określa się jako C-koniec (C-terminus). Białko zawsze zapisuje się od końca N do końca C.



Rys. 1.1.2. Tworzenie wiązania peptydowego.

## 1.2. Bazy białkowe

Struktury białek, które otrzymano eksperymentalnie lub na podstawie przewidywań przechowywane są w bazach danych. Najstarszą i najbardziej znaną jest PDB (Protein Data Bank), który



Rys. 1.2a. Strona PDB.

```

_cell.entry_id      4HHB
_cell.length_a     63.150
_cell.length_b     83.590
_cell.length_c     53.800
_cell.angle_alpha  90.00
_cell.angle_beta   99.34
_cell.angle_gamma  90.00
_cell.Z_PDB        4

```

Rys. 1.2b. Plik PDB w postaci wartości. Dane dotyczą typu komórki oraz wymiarów.

powstał w 1971 r.

PDB zawiera koordynaty atomowe oraz inne informacje opisujące białka. Pliki zawierają listę atomów oraz ich lokalizację w przestrzeni trójwymiarowej. Pliki występują w formacie PDB, mmCIF, XML. Typowy plik zawiera nagłówek, dane o strukturze, atomy z koordynatami oraz dane eksperymentalne. Dane mogą być przedstawione w postaci „kluczowych wartości” lub tabelarycznie. W każdym pliku podane są wszystkie atomy, ich nazwy, pozycja, co pozwala łatwo stworzyć model przestrzenny. Informacje zawarte w pliku PDB mogą być wizualizowane, przeglądane oraz zmieniane.

```

loop_
_atom_site.group_PDB
_atom_site.id
_atom_site.type_symbol
_atom_site.label_atom_id
_atom_site.label_alt_id
_atom_site.label_comp_id
_atom_site.label_asym_id
_atom_site.label_entity_id
_atom_site.label_seq_id
_atom_site.pdbx_PDB_ins_code
_atom_site.Cartn_x
_atom_site.Cartn_y
_atom_site.Cartn_z
_atom_site.occupancy
_atom_site.B_iso_or_equiv
_atom_site.pdbx_formal_charge
_atom_site.auth_seq_id
_atom_site.auth_comp_id
_atom_site.auth_asym_id
_atom_site.auth_atom_id
_atom_site.pdbx_PDB_model_num
ATOM 1 N N . VAL A 1 1 ? 6.204 16.869 4.854 1.00 49.05 ? 1 VAL A N 1
ATOM 2 C CA . VAL A 1 1 ? 6.913 17.759 4.607 1.00 43.14 ? 1 VAL A CA 1
ATOM 3 C C . VAL A 1 1 ? 8.504 17.378 4.797 1.00 24.80 ? 1 VAL A C 1
ATOM 4 O O . VAL A 1 1 ? 8.805 17.011 5.943 1.00 37.68 ? 1 VAL A O 1
ATOM 5 C CB . VAL A 1 1 ? 6.369 19.044 5.810 1.00 72.12 ? 1 VAL A CB 1
ATOM 6 C CG1 . VAL A 1 1 ? 7.009 20.127 5.418 1.00 61.79 ? 1 VAL A CG1 1
ATOM 7 C CG2 . VAL A 1 1 ? 5.246 18.533 5.681 1.00 80.12 ? 1 VAL A CG2 1

```

Rys. 1.2c. Plik PDB w postaci tabelarycznej. Między innymi przedstawione są koordynaty atomowe, x, y, z.

## 2. Analiza właściwości fizyko-chemicznych białka

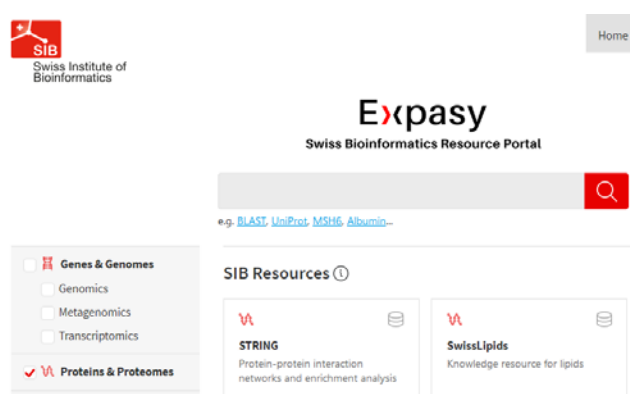
### 2.1. Właściwości fizyko-chemiczne wybranych białek

➔ Pomimo, iż właściwości fizyko-chemiczne białek są podobne (wszystkie białka są bezbarwne, bezsmakowe, mają duży ciężar właściwy oraz tworzą koloidy), skład aminokwasowy wpływa na ładunek białka, rozpuszczalność, zdolność do tworzenia soli. Przykładowo rozkład aminokwasów hydrofobowych i hydrofilnych na powierzchni białka oraz ładunek decydują o jego rozpuszczalności. Białka klasyfikowane są do poszczególnych grup, rodzin między innymi na podstawie właściwości fizyko-chemicznych. I tak wysoka zawartość lizyny i argininy oraz niski ciężar właściwy wskazują na histony. Na podstawie właściwości fizyko-chemicznych białka dzieli się na globularne i fibrylarne.

- Masa cząsteczkowa (MW): określana jest w Daltonach (Da). Masa 24 500 DA oznacza, że jeden mol substancji waży 24 500 g. Masę cząsteczkową można obliczyć na podstawie wzoru chemicznego i mas atomowych.
- Ładunek: sumaryczny ładunek, np. 16 dodatnich ładunków i 15 ujemnych daje ładunek +1.
- Punkt izoelektryczny: jest to pH, w którym cząsteczka nie ma ładunku. Ładunki pozytywne i negatywne danej cząstki równoważą się.
- Indeks alifatyczny: opisuje zawartość aminokwasów, które nie mają pierścieni aromatycznych w łańcuchach bocznych czyli zawartość alaniny, waliny, leucyny i izoleucyny. Wartość tę podaje się jako frakcję molową czyli liczbę moli danej substancji podzieloną przez całkowitą liczbę moli. Wartość często podaje się w procentach.
- Hydrofobowość: powinowactwo do wody. Cząsteczki hydrofilne są polarne, a hydrofobowe są niepolarne. Większość aminokwasów jest hydrofilna, jedynie siedem aminokwasów jest hydrofobowych: alanina, cysteina, leucyna, isoleucyna, metionina, fenyloalanina i walina. Aminokwasy hydrofobowe najczęściej występują w regionach transmembranowych lub rdzeniu białka. Znając rozkład aminokwasów hydrofobowych można przewidzieć regiony transmembranowe i aktywne.
- Czas półtrwania: okres, w którym połowa ilości białka w komórce ulega degradacji.

### 2.2. Baza Expasy

➔ Expasy (<https://www.expasy.org/>) zawiera narzędzia bioinformatyczne obejmujące około 160 różnych baz wraz z odpowiednim oprogramowaniem. Bazy i programy niezbędne do analizy białek są zgrupowane na podstronie "Protein I Proteomics".



Rys. 2.2. Serwer Expasy, strona główna.

### 2.3. Analiza właściwości fizyko-chemicznych białka AAM68164.1



2.3.1. Numer akcesyjny AAM68164.1 oznacza pewne białko. Proszę podać nazwę tego białka uwzględniając wszystkie synonimy oraz gatunek z, którego białko pochodzi (**1 punkt**). Proszę zapisać sekwencję aminokwasów tego białka w formacie Fasta, np. w pliku tekstowym.

2.3.2. Na serwerze Expasy proszę wybrać program ProtParam i po wejściu na serwer proszę wklepiować sekwencję aminokwasową. Na podstawie uzyskanych danych proszę podać następujące wartości:

- A. Liczbę aminokwasów oraz wzór chemiczny (**1 p**)
- B. Masę cząsteczkową oraz punkt izoelektryczny oraz proszę określić zasadowość/kwasowość tej cząsteczki w pH 7.0 (**2 p**)
- C. Który aminokwas występuje najczęściej, a który najrzadziej? Proszę podać tę częstość. (**1 punkt**)
- D. Proszę podać ładunek białka (net charge) i wyjaśnić sposób jego obliczenia. (**2 p**)

*Czas wykonania: 15 minut*

## 3. Analiza porównawcza struktury białka

### 3.1. Porównanie wielu sekwencji

➔ Analiza zbieżności wielu sekwencji (uliniowanie wielokrotne, multiple alignment) polega na porównaniu sekwencji w taki sposób aby znaleźć regiony podobne. Regiony te są elementem podobieństwa funkcjonalnego, i strukturalnego, a także zależności ewolucyjnych. Najczęściej wykorzystuje się techniki hereustyczne, w których zbieżność analizuje się:

- **progresywnie** (hierarchicznie, metodą drzewa): polega na porównaniu parami zaczynając od dwóch sekwencji najbardziej podobnych i dodawaniu sekwencji coraz bardziej odległych; metodę tę wykorzystuje CLUSTAL Omega i T-Coffee;
- **iteratywnie** (powtarzające): polega na wykorzystaniu metod progresywnych oraz kilkukrotnym ponownym uliniowaniu porównywanych sekwencji, co pozwala uniknąć błędów generowanych przez metody progresywne, metodę tę wykorzystuje PRRN/PRRP;
- **na bazie bloków**: metoda polega na podziale sekwencji na bloki i uliniowaniu bez luk poszczególnych bloków, metodę tę wykorzystuje DIALIGN2.

Rys. 3.1.1. Serwer CLUSTAL Omega



### 3.1.1. CLUSTAL Omega

Program pozwala na analizę sekwencji RNA, DNA i białka. Po zaakceptowaniu wyboru (submit) program zwróci uliniowane sekwencje.

### 3.1.2. T-Coffee

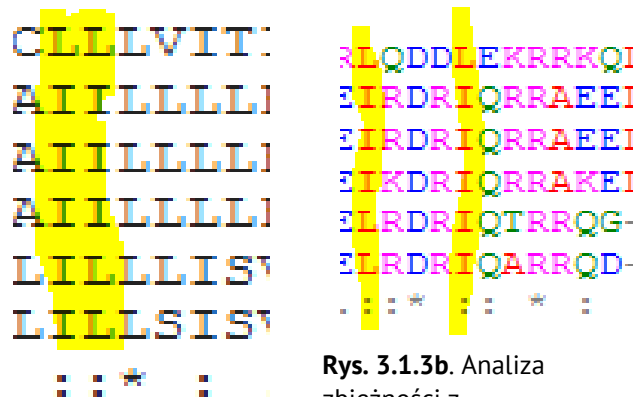
Program pozwala na analizę sekwencji RNA, DNA i białka. W odniesieniu do białek można wykorzystać kilka opcji. Espresso umożliwia uliniowanie na podstawie podobieństw strukturalnych. PSI Coffee opiera się na homologii i pozwala na wiarygodne uliniowanie sekwencji odległych ewolucyjnie. M-Coffee jest to kombinacja kilku, najbardziej popularnych metod analizy zbieżności.



Rys. 3.1.2. Serwer T-Coffee.

### 3.1.3. Odczyt wyników

- Gwiazdka „\*” lub nazwa aminokwasu oznacza aminokwasy, które są identyczne we wszystkich sekwencjach.
- Dwukropek „:” oznacza prawie idealne dopasowanie, np. tylko jeden aminokwas jest inny, przy czym jest on podobny. Jest to konserwatywne dopasowanie.
- Kropka „.” Oznacza dopasowanie większości, ale nie wszystkich aminokwasów, przy czym aminokwasy są podobne, a więc dopasowanie jest semi-konserwatywne.
- Przykładowo na rys. 3.1.3b na żółto zaznaczono zamianę leucyny w izoleucynę, które są aminokwasami podobnymi o podobnej funkcji biologicznej. Pozycje te oznaczone są dwukropkiem.



Rys. 3.1.3a. Oznaczenie pozycji w przykładowej analizie zbieżności.

Rys. 3.1.3b. Analiza zbieżności z uwzględnieniem oznaczeń kolorami.

#### Konserwatywne aminokwasy obejmują sześć grup:

- ▶ alifatyczne aminokwasy czyli glicyna (G), alanina (A), walina (V), leucyna (L), izoleucyna (I);
- ▶ z grupami hydroksylowymi lub siarkowymi czyli seryna (S), cysteina (C), treonina (T), metionina (M);
- ▶ aromatyczne czyli fenylalanina (F), tyrozyna (Y), tryptofan (W);
- ▶ zasadowe czyli histydyna (H), lizyna (K), arginina (R);
- ▶ amidy i kwasowe czyli asparagina (D), glutamina (E), kwas asparaginowy (N), kwas glutaminowy (Q).
- Grupy konserwatywne dzieli się na podgrupy, dlatego oprócz konserwatywnych podstawień wyróżnia się podstawienia semi-konserwatywne.

### 3.2. Analiza zbieżności białka AAM68164.1

3.2.1. W notatniku proszę przygotować sekwencje białek o podanych poniżej numerach akcesyjnych. Proszę wszystkie sekwencje wkopiować w formacie FASTA. Plik w notatniku proszę zapisać na swoim komputerze, tak aby można było jego zawartość wkopiować do programu analizującego zbieżność.

Lista sekwencji:

- AAM68164.1
- ABB73024.1
- XP\_028701900.1
- XP\_025319403.1
- KAF0880361.1
- BBG56793.1

3.2.2. Proszę podać nazwę każdej sekwencji oraz nazwę łacińską i polską gatunku, z którego ona pochodzi. **(1 punkt/3 sekwencje)**.

3.2.3. Proszę wejść na stronę: <https://www.ebi.ac.uk/Tools/msa/clustalo/>, Clustal Omega. W ramce proszę wkopiować sekwencję aminokwasów w formacie FASTA dla wszystkich podanych numerów akcesyjnych. W tym celu najprościej przekopiować zawartość notatnika z punktu 3.2.1. Proszę pozostawić domyślne parametry i zatwierdzić analizę za pomocą „submit”. Po otrzymaniu wydruku proszę wybrać opcję „show color”.

3.2.4. W każdym rzędzie znajduje się 60 pozycji. Przyjmując za punkt odniesienia pierwszą sekwencję na liście, proszę podać, który fragment jest najbardziej konserwatywny we wszystkich sześciu sekwencjach. Proszę podać procent aminokwasów identycznych oraz konserwatywnych. **(2 punkty)**

3.2.5. Korzystając z opcji „Result Summary” oraz macierzy podobieństwa procentowego proszę podać gatunek, którego sekwencja jest najbardziej podobna do AAM68164.1 oraz gatunek, którego sekwencja jest najmniej podobna do AAM68164.1. Proszę podać wartość tego podobieństwa. **(2 punkty)**

3.2.6. Do sekwencji, którego gatunku jest najbardziej podobna sekwencja hieny cętkowanej. Proszę podać nazwę gatunku oraz wartość podobieństwa. **(1 punkt)**.

**Czas wykonania: 15 minut**

## Odpowiedzi

### 2. Analiza właściwości fizyko-chemicznych

#### 2.3. Analiza właściwości fizyko-chemicznych białka AAM68164.1

2.3.1. Numer akcesyjny AAM68164.1 oznacza pewne białko. Proszę podać nazwę tego białka uwzględniając wszystkie synonimy oraz gatunek z, którego białko pochodzi (**1 punkt**). Proszę zapisać sekwencję aminokwasów tego białka w formacie Fasta, np. w pliku tekstowym.

- Gatunek: Homo sapiens
- Białko: envelope glycoprotein, *ERVWE1* – locus B, *HERV-W*, syncytyna
- Seqwencja aminokwasowa w formacie FASTA

```
>AAM68164.1 envelope glycoprotein (endogenous virus) [Homo sapiens]
MALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFLWRMQRPGNIDAPSYRSLSKGTPTFTAHTHMPR
NCYHSATLCMHANTHYWTGKMINPSCPGGLGVTWCWYFTQTGMSDGGGVQDQAREKHVKEVISQLTRVH
GTSSPYKGLDLSKLHETLRTHRLVSLFNTTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPVPEQWNNFS
TEINTTSLVGLVSNLEITHSNTLTCVKFSNTTYTNSQCIRWVTPPTQIVCLPSGIFVCGTSAYRCL
NGSSESMCFLSFLVPPMTIYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALGTGIGGITTSTQFYKYL
SQELNGDMERVADSLVTLQDQLNSLAAVLQNRALDLLTAERGGTCLFLGEECCYYVNSQSGIVTEKVKKE
IRDRIQRRAEELRNTGPWGLLSQWMPWILPFLGPLAAIIIIIIIFGPCIFNLLVNFVSSRIEAVKLQMEPK
MQSKTKIYRRPLDRPASPRSDVNDIKGTPPEEISAAQPLLRPNSAGSS
```

2.3.2. Na serwerze ExPASy proszę wybrać program ProtParam i po wejściu na serwer proszę wkopiować sekwencję aminokwasową. Na podstawie uzyskanych danych proszę podać następujące wartości:

**A.** Liczbę aminokwasów oraz wzór chemiczny (**1 p**)

**538 aa, wzór  $C_{2676}H_{4203}N_{723}O_{773}S_{31}$**

**B.** Masę cząsteczkową oraz punkt izoelektryczny oraz proszę określić zasadowość/kwasowość tej cząsteczki w pH 7.0 (**2 p**)

**MW = 59 866,05 Da, pI = 8,8**

**pH = 7.0 < pI = 8.8, zatem białko będzie pobierało jony  $H^+$ , będzie miało ono ładunek dodatni, zatem będzie zasadą.**

**Każda cząsteczka w  $pH < pI$  jest zasadą (pobiera jony  $H^+$ ) oraz kwasem w  $pH > pI$  czyli oddaje jony  $H^+$ .**

**C.** Który aminokwas występuje najczęściej, a który najrzadziej? Proszę podać tę częstość. (**1 punkt**)

**Najczęściej występuje leucyna, 11.2%**

**Najrzadziej występuje tryptofan, 1.7%**

**D.** Proszę podać ładunek białka (net charge) i wyjaśnić sposób jego obliczenia. (**2 p**)

Białko posiada 36 ujemnie naładowanych aminokwasów oraz 46 dodatnio naładowanych, co daje ładunek +10.



### 3. Analiza porównawcza sekwencji

#### 3.2. Analiza zbieżności białka AAM68164.1

3.2.2. Proszę podać nazwę każdej sekwencji oraz nazwę łacińską i polską gatunku, z którego ona pochodzi. (1 punkt/3 sekwencje).

- **AAM68164.1:** białko kopertowe, syncytyna, *Homo sapiens*, człowiek
- **ABB73024.1:** syncytyna 1, *Gorilla gorilla*, goryl zachodni
- **XP\_028701900.1:** syncytyna 1, *Macaca mulatta*, makak królewski
- **XP\_025319403.1:** podobna do syncytyny 1, *Canis lupus dingo*, Wilk szary, podgatunek dingo
- **KAF0880361.1:** syncytyna, fragmenty; *Crocuta crocuta*, krokuta (hiena) cętkowana
- **BBG56793.1:** białko płaszcz, *Simian rethrowirus, SRV*; wirus małpi.

3.2.3. Proszę wejść na stronę: <https://www.ebi.ac.uk/Tools/msa/clustalo/>, Clustal Omega. W ramce proszę wkopiować sekwencję aminokwasów w formacie FASTA dla wszystkich podanych numerów akcesyjnych. W tym celu najprościej przekopiować zawartość notatnika z punktu 3.2.1. Proszę pozostawić domyślne parametry i zatwierdzić analizę za pomocą „submit”.

3.2.4. W każdym rzędzie znajduje się 60 pozycji. Przyjmując za punkt odniesienia pierwszą sekwencję na liście, proszę podać, który fragment jest najbardziej konserwatywny we wszystkich sześciu sekwencjach. Proszę podać procent aminokwasów identycznych oraz konserwatywnych. (2 punkty)

```

BBG56793.1      LAEVLQNRRLDLLTAEQGGICLALQEKCCFYANKSGIVRDKIKRLQDDLEKRRKQLID      512
AAM68164.1      LAAVVLQNRRALDLLTAERGGTCLFLGEECCYYVNSGIVTEKVKIIRDRIQRRAEELRN      434
ABB73024.1      LAAVVLQNRRALDLLTAERGGTCLFLGEECCYYVNSGIVTEKVKIIRDRIQRRAEELRN      434
XP_028701900.1 LAAVVLQNRALDLLTAERGGTCLFLGEECCYYVNSGIISEKVKIIRDRIQRRAKELQN      434
XP_025319403.1 LAAVTLQNRRALDLLTAERGGTCLYLNEECCYFVNQSGIVTSKVKELRDRIQTRRQG---      486
KAF0880361.1    LAAIALQNRRALDLLTAERGGTCLYLKEECCYYVNSGIVTSKVELLRDRIQARRQD---      320
** :.***:*.*****:** * * * :***:.*:***: .*:::* : : * :

```

- Pierwszą uliniowaną sekwencją jest BBG56793.1. Podajemy względem tej sekwencji. Jest to fragment od 453 do 512 aminokwasu. Na tym obszarze prawie wszystkie aminokwasy są identyczne lub podobne.
- 28 aminokwasów czyli 46,7% jest identycznych we wszystkich analizowanych sekwencjach.
- 15 aminokwasów czyli 25% to podstawienia konserwatywne.

3.2.5. Korzystając z opcji „Result Summary” oraz macierzy podobieństwa procentowego proszę podać gatunek, którego sekwencja jest najbardziej podobna do AAM68164.1 oraz gatunek, którego sekwencja jest najmniej podobna do AAM68164.1. Proszę podać wartość tego podobieństwa. (2 punkty)

1: BBG56793.1	100.00	29.64	29.23	28.60	29.27	30.79
2: AAM68164.1	29.64	100.00	97.21	90.71	39.58	41.99
3: ABB73024.1	29.23	97.21	100.00	90.71	39.58	41.47
4: XP_028701900.1	28.60	90.71	90.71	100.00	37.64	38.89
5: XP_025319403.1	29.27	39.58	39.58	37.64	100.00	72.51
6: KAF0880361.1	30.79	41.99	41.47	38.89	72.51	100.00

- Sekwencja AAM68164.1. kodująca białko kopertowe (syncytynę 1) u człowieka jest najbardziej podobna do sekwencji ABB73024.1 pochodzącej od goryla zachodniego (97,21%), natomiast najmniej podobna do sekwencji BBG56793.1 kodującej białko płaszczka u wirusa małpiego SRV (29,64%).

3.2.6. Do sekwencji, którego gatunku jest najbardziej podobna sekwencja hieny cętkowanej. Proszę podać nazwę gatunku oraz wartość podobieństwa. (1 punkt).

- Sekwencja hieny cętkowanej ma numer KAF0880361.1. Najwyższe podobieństwo wynoszące 72,51% wykazuje ona względem sekwencji XP\_025319403.1, która pochodzi od wilka szarego, podgatunku dingo.

3.2.7. Drzewo filogenetyczne sekwencji otrzymane przy pomocy Neighbour Joining odzwierciedla powiązania filogenetyczne między syncytyną pochodzącą od różnych gatunków.

