---

author: Batuhan Seyhan S20343
title: lab4_Biology
date:
keywords: 11C
geometry: margin=2cm
output: pdf_document
fontsize: 12pt
lang: en

TASK 1

https://www.rcsb.org/structure/6ID1

TASK 2

https://ibb.co/NrN7ggQ

TASK 3

https://ibb.co/SvPZFTM

TASK 4

It is a three-dimensional structure. The protein is represented by a series of connected sticks, where each stick represents a covalent bond between two atoms. The sticks are mostly represented
in green or red color.

TASK 5

In this method, the protein is represented by a simplified cartoon-like diagram that shows the backbone of the protein as a series of connected spiral shaped lines.

TASK 6

https://ibb.co/LPqvDPF

TASK 7

In this method, the protein is represented by a series of flat, ribbon-like structures that follow the
path of the protein’s backbone. The widths of the ribbons differ according to the structures of protein.

TASK 8

We are dealing with hot spot.
Both of the curves which represents alpha-helices and beta-sheets accordingly goes under the X
axis, therefore we are dealing with so-called “hotspots”.

TASK 9

The "St" column in the GOR algorithm output refers to the predicted secondary structure of the amino acid residues in the protein sequence, where "C" denotes coil, "H" denotes helix, and "E" denotes strand.

TASK 10

The number of predicted transmembrane helices
68.6888999999999

TASK 11

Exp number of AAs in TMHs refers to expected number of amino acids that are typically found within the transmembrane helices of a given membrane protein.

TASK 12

The plot of probabilities in TMHMM shows the probability that each amino acid in a protein sequence is located within a transmembrane helix, providing insights into the potential topology and structural features of the protein.

TASK 13

The GOR algorithm uses the amino acid sequence to forecast the secondary structure (alpha helix,
beta strand, coil) of a protein, whereas the graphs generated by TMHMM predict transmembrane
helices by taking into account both the hydrophobicity and amino acid sequence.

TASK 14

https://ibb.co/Bc5QNh1

TASK 15

protein structure AlphaFold [1] and OmegaFold [2] and transmembrane protein topology "DeepTMHMM [3]. The Mol* toolkit [4] is used for visualisation

TASK 16

https://doi.org/10.2210/pdb4UUM/pdb

TASK 17

https://ibb.co/VjY3qnc

TASK 18

Serine and Glutomic Acid

TASK 19

Orthogonal coordinates for X,y,z in Angstroms

TASK 20

Fold assignment methods are used to determine the structure of a protein from its amino acid sequence. They typically use computational algorithms to search for similar protein structures in a database and predict the structure of the protein of interest based on the detected similarities.

When selecting "Fast" as the fold assignment method, it means that the computational algorithm used for predicting the protein structure will prioritize speed over accuracy. This can be useful for quickly predicting the structure of a large number of proteins, but may not produce the most accurate predictions compared to slower, more computationally intensive methods.

TASK 21

Run-Name: ModWeb20230513_10570
No. of sequences submitted: 1 No. of hits detected : 79 No. of models calculated : 56 Model selection criteria : MPQS TSVMOD LONGEST_DOPE DOPE No. of models selected : 1

TASK 22

Model is considered credible

TASK 23

https://ibb.co/xS9VKG2

TASK 24

```
protein_sequence = input("Enter protein sequence: ")
new_sequence = ""

for i in range(len(protein_sequence)):
if protein_sequence[i] == "P":
new_sequence += "*P"
else:
new_sequence += protein_sequence[i]

print("\nMarked sequence: ", new_sequence)
```