

**GEBZE TECHNICAL UNIVERSITY, FACULTY OF ENGINEERING,  
DEPARTMENT OF BIOENGINEERING  
2024-2025 FALL SEMESTER  
BENG451: Introduction to Bioinformatics**

**HOMEWORK I**

Please submit a PDF file including your answers/solutions.

For each of the assignments listed below, provide screenshots detailing the steps you have followed.

**Pairwise protein sequence alignment using web-based tools**

The European Bioinformatics Institute provides a simple interface (EMBOSS Needle) to perform pairwise protein sequence alignments using the Needleman-Wunsch algorithm at [https://www.ebi.ac.uk/Tools/psa/emboss\\_needle/](https://www.ebi.ac.uk/Tools/psa/emboss_needle/). The user only needs to paste two protein sequences into their respective boxes on the screen, and click on “Submit” button. Advanced options such as selecting the Scoring/Substitution Matrix, and setting custom Gap Opening/Closing/Extension penalties can be set by using the “... More Options” button.

Take some time to get familiar with the EMBOSS Needle interface, and then perform the following analysis:

- Get familiar with dihydrofolate reductase by reading its PDB molecule of the month description at <https://pdb101.rcsb.org/motm/34>
- Retrieve the amino acid sequences of the enzyme **dihydrofolate reductase** from **human and three additional organisms** in FASTA format. You can use the UniProt web interface to do this. The selection of the organisms is up to you.
- Perform pairwise alignments for each possible pair of these four sequences. Experiment with different scoring matrices (e.g., BLOSUM and PAM families) and tweak gap insertion/extension penalties.
- Report alignment results to highlight matches, mismatches, and gaps, as clearly as possible. You are welcome to use external alignment viewers such as Jalview.
- Compare and discuss the alignment results, including identity and similarity scores. Which of the alignments was the most successful one? Why? (Hint: you may also include a short discussion of the structure of the protein based on what you know about protein sequence-structure-function relationship of the protein).

**Please try to submit a comprehensive discussion. Your responses will be evaluated with great emphasis on your discussion.**