

GEBZE TECHNICAL UNIVERSITY, FACULTY OF ENGINEERING,  
DEPARTMENT OF BIOENGINEERING  
2024-2025 FALL SEMESTER  
BENG451: Introduction to Bioinformatics  
BSB511: Bioinformatics Fundamentals

## HOMEWORK I

Please submit a single Jupyter notebook including your answers/solutions.

### Pairwise protein sequence alignment

- 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 using Biopython.
- 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. Again, you should use Biopython's align2 module to perform this.
- Report alignment results to highlight matches, mismatches, and gaps, as clearly as possible.
- Your code should generate a FASTA file that includes your alignment results.
- 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 as well.**