**Project 4**

**Prediction of Protein Function from Sequence Data Using BLAST and Pfam**

**Objective:** Predict function of a protein sequence using homology and conserved domain search.

**Data Source:**

* Use a hypothetical protein sequence from NCBI (choose one annotated as “uncharacterized protein”).

**Steps:**

1. Retrieve protein sequence (FASTA) from NCBI.
2. Run **BLASTp (NCBI)** to find homologous proteins and record % identity, organism, function.
3. Use **Pfam (online)** to search for conserved domains.
4. Combine BLAST + Pfam evidence to suggest a likely function.
5. Compare with literature if available.

**Deliverables:**

* BLAST results table
* Pfam domain screenshot
* 2–3 page report: predicted function, confidence, and possible biological role