

# Bioinformatics Tools Assignment

**Name: Laiba Suleman**

**Program: BS 7<sup>th</sup> (M)**

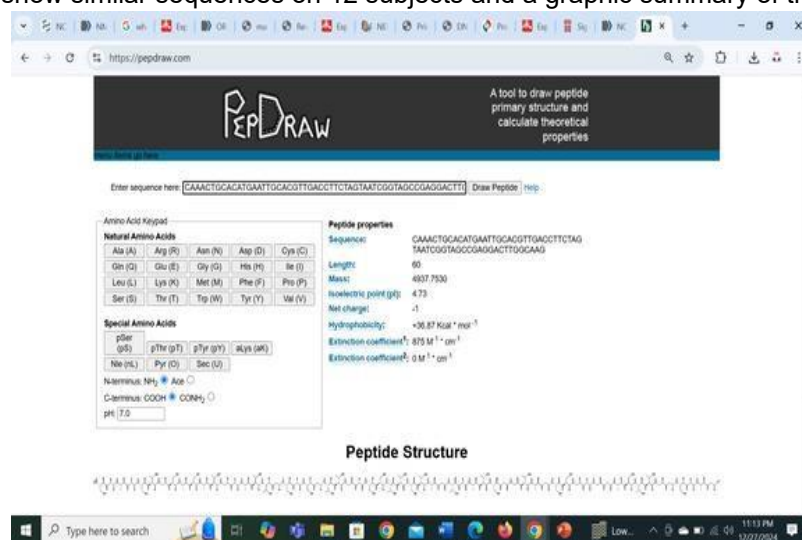
**Reg: 04032213027**

## DNA Sequence Used

```
CAAAC TGCACATGAATTGCACGTTGACCTTCTAGTAATCGGTAGCCGAGGACTTGGC
AAGATTAAGAGGGCTTTCTTGGGTAGTGTGAGCGACTATTGTGCTCATCATGTGCAA
TGTCCTATTCTCATTGTGAAGCCACCCAAGGAAAACCCCAAAGCAGCTAATTAAC
AAGGATATATCATGTGCCAGTCTATGAACAATGTCATTTACTTGATGATGAATAAAC
ACAAGTTCATTGAGGATTATAAACTTATGTATATATTAATTATATTATCAATAAAT
GGAGGTTGTATCTGTTGTCTCAATAGCTAGCCAGTCCACAGGCAATAGGTTTTTGG
GTAAAAGCCAGTCGTCACAAC TAGTGCAATCTGAATATTTGACCCTCATTCTCTGG
TAACGCAAGTTGCTGCAGCTTATGTGGCTCTGTATTTGGTTTTAATAAATATAACAA
TTCAAATTAGAGATATCTAGCACTTCAACCTCGATGGAACAAA
```

## BLAST

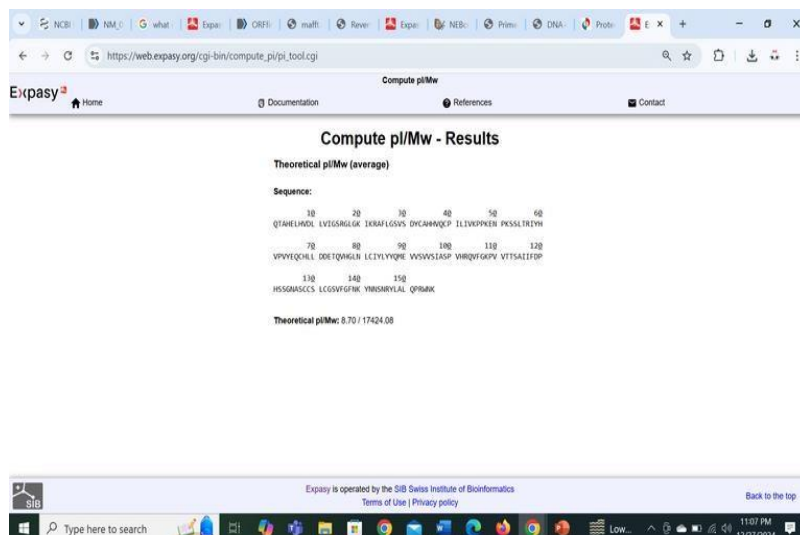
BLAST results show similar sequences on 12 subjects and a graphic summary of the top hits.



## Hydropathy Plot

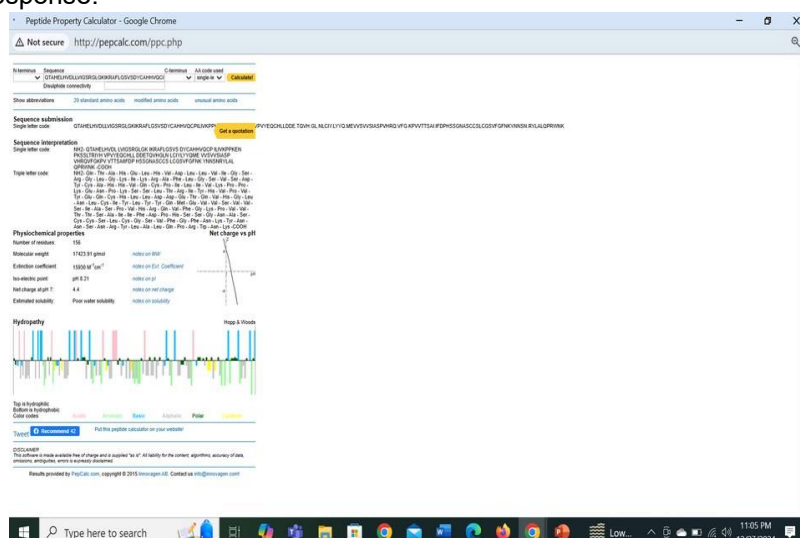
The hydropathy plot, which shows the hydrophobicity of amino acids on protein sequence.

The analysis shows four potential open reading frames, with the largest spanning 182-391 nucleotides, and predicts a signal peptide with a cleavage site between positions 16 and 17.



## Conserved Domain (CDD)

The analysis identified a conserved domain, HUP A, in the input sequence, suggesting a potential role in stress response.



## DNA Stats

The analysis shows that the 498 bp linear sequence has 38% GC content and 62% AT content and can be cut by 19 different restriction enzymes, including both frequent and rare cutters.

**Transcription and Translation Tool**  
Converts sequences from DNA to RNA and from this to protein.

**DNA (deoxyribonucleic acid)**

- permanent copy of the genetic information
- uses "T" rather than "U"
- without 2' OH group
- more stable than RNA
- lower error rate during replication than RNA

**RNA (ribonucleic acid)**

- messenger RNA (mRNA) is a temporary copy of the sequence of the gene that codes for the protein

**Protein**

- consists on amino acids linked by amide bonds ("peptide bonds")
- most enzymes and many structural components in cells are proteins

**DNA sequence:**

**transcription**

**RNA sequence:**

**translation**

☐ detect start codon  
☐ detect stop codon

**protein sequence:**

[Help with amino acid](#) [abbreviations](#)

## Expsy Translate Tool

The Expsy Translate tool translated the nucleotide sequence into six possible protein sequences based on different reading frames.

**Graphical View**

**Enzyme List**

**Sequence**

**ORF Summary**

**Flanking Sites**

**Custom Digest**

**Results For:**

**Enzymes:**

- Supplier: NEB
- Type: Type I, II, III, Homing, Nicking
- ORFs:
  - Min: 100 aa
  - Searched

**Sequence Information:**

498 bp, Linear  
GC = 38%, AT = 62%

**Display:**

- ☐ Circular
- ☒ Linear
- ☐ Alternative
- ☐ Show flanks

**Cleavage:**

- ☒ Blunt-End Cut
- ☐ Cuts 1 strand
- ☐ 1' 5' Extension
- ☐ 2' 3' Extension
- ☐ List 0 cutters

**Supplier:**

- ☒ NEB
- ☐ Other supplier
- ☐ Not commercially available

**Site:**

- ☒ Affected by CpG methylation
- ☐ Affected by other methylation
- ☐ Ambiguous site

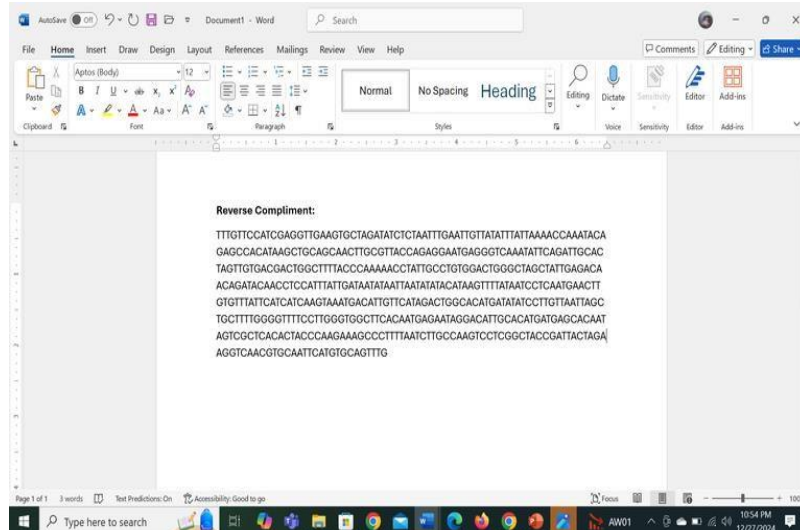
**Graphical View:**

1 498

Restriction enzymes and their cleavage sites are shown on the DNA sequence.

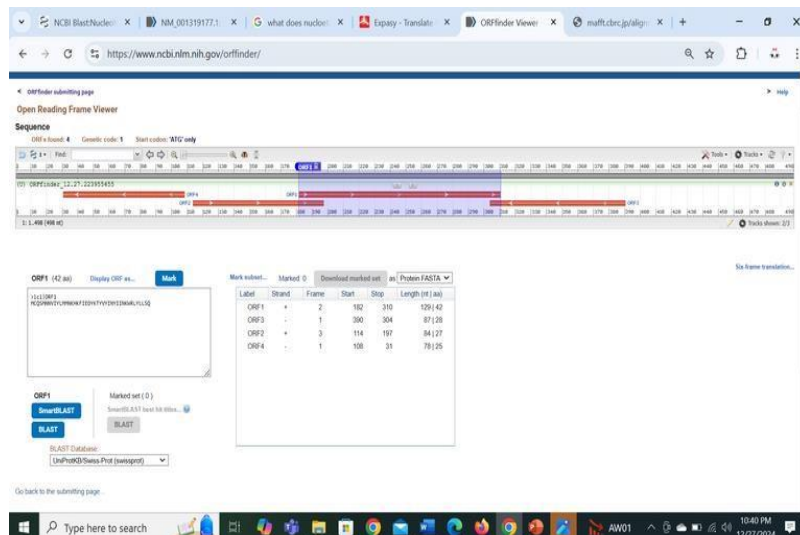
## Primer 3 & Transcription/Translation

It shows primer pairs, pair 1 is good pair and below I use Transcription and translation tool to analyze Amino Acid sequence from DNA Sequence.



## Expsy Compute pI/Mw (ProtParam)

The Expsy Compute pI/Mw tool calculated the (pI) of the protein : 8.70 and (Mw) to be 17424.08 Da.



## Additional Bioinformatics Tools

### STRING Tool

The STRING tool is used to predict protein–protein interaction networks, helping to understand functional associations between proteins derived from the analyzed sequence.

## **CELLO Tool**

The CELLO tool predicts the subcellular localization of proteins based on their amino acid composition, providing insight into where the translated protein may function inside the cell.

## **RNAfold Tool**

The RNAfold tool predicts the secondary structure of RNA sequences using minimum free energy models, helping to understand RNA stability and folding patterns derived from the DNA sequence.