**Internship Task 1:**

**Protein Sequence Analysis using NCBI BLAST**

**Name:** [Navya Thalathoti]  
**Internship:** CodeAlpha  
**Course / Department:** [Bioinformatics] B.TECH

**Gmail**:navyathalathoti@gmail.com

**Title:**

**Identification of Homologous Proteins for Spike glycoprotein SARS-CoV-2 using NCBI BLASTp**

**Aim:**

To analyze the **P0DTC2 (SPIKE\_SARS2)**. protein sequence from *Homo sapiens* using the BLASTp tool and identify homologous proteins across different organisms based on sequence similarity and alignment.

**Introduction:**

The **Basic Local Alignment Search Tool (BLAST)** is a widely used algorithm in bioinformatics that compares an input DNA or protein sequence against databases to find regions of local similarity.  
In molecular biology, BLAST helps researchers identify homologous genes or proteins, predict functional domains, and study evolutionary relationships. Protein BLAST (**BLASTp**) aligns an amino-acid query sequence against a protein database to reveal conserved regions and structural or functional similarities.

**Materials and Methods:**

**Tools and Databases Used:**

* Database: **UniProt** (for protein sequence retrieval)
* Tool: **NCBI BLASTp** (https://blast.ncbi.nlm.nih.gov/Blast.cgi)
* Database used for BLAST: **nr (non-redundant protein sequences)**

**Procedure:**

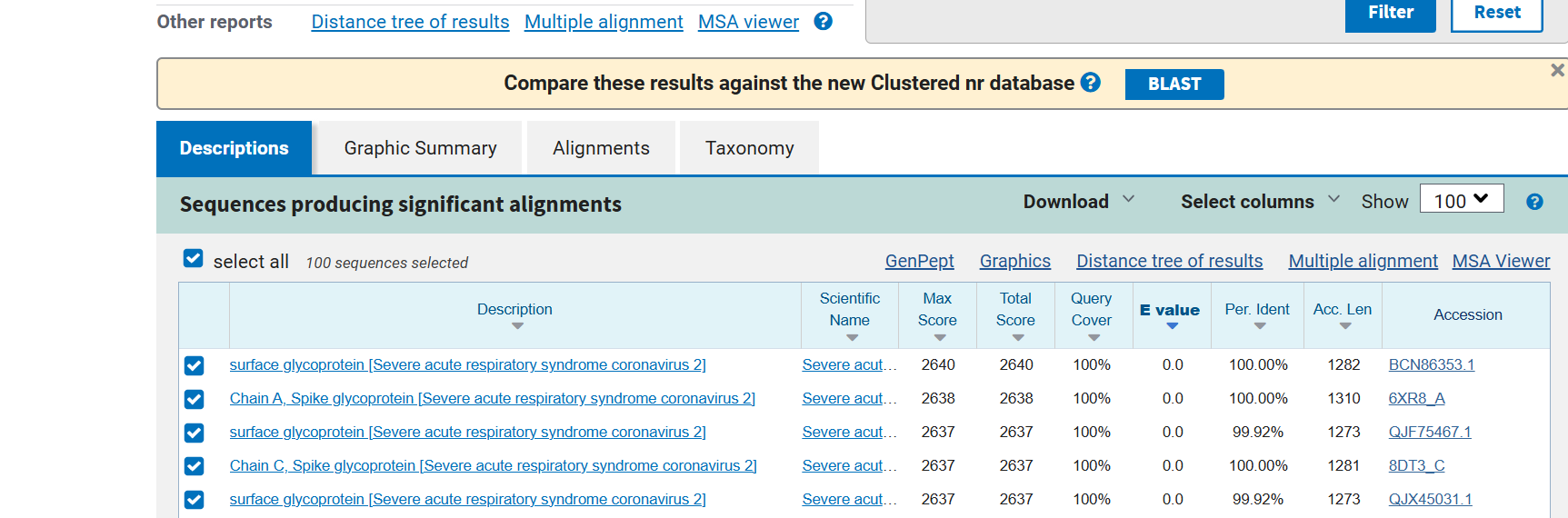
1. Opened UniProt.
2. Searched for Spike glycoprotein SARS-CoV-2 (UniProt ID: P0DTC2).
3. Clicked the **FASTA** link and copied the amino-acid sequence.
4. Navigated to the **NCBI BLAST** page → selected **Protein BLAST (blastp)**.
5. Pasted the sequence into the search box.
6. Kept default parameters and selected database **nr**.
7. Clicked **BLAST** and waited for the results.
8. Downloaded and recorded the summary table and alignment section.

**Results:**

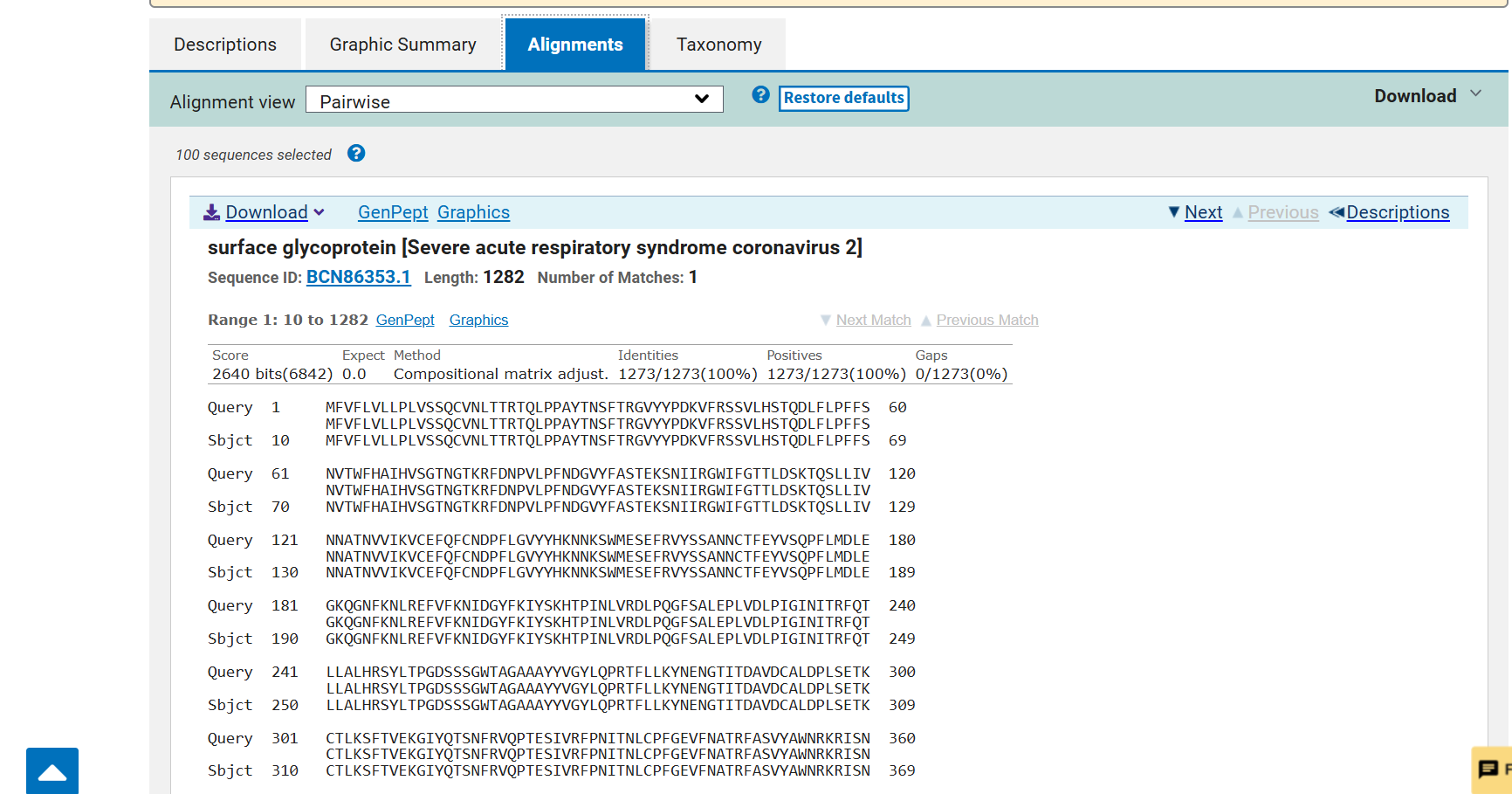
**1. FASTA Sequence :**



**2.Summary of Top BLAST Hits:**

****

**3.** **Alignment Section:**

****

**Interpretation:**

The BLAST analysis of the SARS-CoV-2 spike glycoprotein showed very high similarity with spike proteins of other coronaviruses. The top hits were from different SARS-CoV-2 strains from many countries showing 99–100% identity. Other close matches were from bat and pangolin coronaviruses with 92–96% similarity. This shows that SARS-CoV-2 may have evolved from these animal viruses. The receptor-binding part of the protein was mostly similar but had a few small changes that might help the virus attach to human cells more easily.

**Discussion:**The spike protein of SARS-CoV-2 is highly conserved, meaning it is almost the same in all virus samples around the world. This makes it an important protein for infection and for vaccine development. The few changes found in some parts, especially in the receptor-binding domain, may help the virus infect humans more efficiently. Studying these similarities helps scientistsunderstand how the virus evolved, how it spreads, and how to make better vaccines and medicines.