

# **CodeAlpha Bioinformatics Internship –**

## **Task 1 Report & Result**

### **Title**

**DNA / Protein Sequence Analysis using BLAST**

Student Name: **Gulam Sarwer**

Internship Platform: **CodeAlpha**

Domain: **Bioinformatics**

Task Number: **Task 1**

### **Introduction**

Bioinformatics is an interdisciplinary field that combines biology and computer science to analyze biological data. Sequence analysis is one of the fundamental tasks in bioinformatics, helping in identify similarities between DNA or protein sequences and understand their biological significance. This task focuses on performing **DNA/Protein sequence analysis using the BLAST tool**.

### **Objective**

To perform protein sequence analysis using BLAST

To identify homologous protein sequences

To analyze sequence similarity, identity, query coverage, and E-value

### **Tools and Databases Used**

NCBI BLAST (BLASTp)

UniProt Database

Web based bioinformatics tools

## Selected Sequence Details

Protein Name: Hemoglobin Subunit Beta

Organism: Homo sapiens

UniProt ID: P68871

Sequence Format: FASTA

## Methodology

Protein sequence was downloaded from UniProt in FASTA format.

The FASTA sequence was submitted to NCBI BLASTp.

Default parameters were used for analysis.

BLAST search was performed against the non-redundant protein database.

Results were analyzed based on identity, query coverage, and E-value.

## Results

High similarity with known hemoglobin proteins was observed.

Percent Identity: 100%

Query Coverage: 100%

E-value: 0.0

Alignment score indicated strong homology.

Screenshots of BLAST results and alignments were recorded for documentation.

The screenshot displays the UniProt entry for Hemoglobin subunit beta (P68871). The page is titled "P68871 · HBB\_HUMAN". The left sidebar contains navigation links: Function, Names & Taxonomy, Subcellular Location, Disease & Variants, PTM/Processing, Expression, Interaction, Structure, Family & Domains, Sequence, and Similar Proteins. The main content area shows protein details: Protein name (Hemoglobin subunit beta), Gene (HBB), Status (UniProtKB reviewed (Swiss-Prot)), and Organism (Homo sapiens (Human)). It also displays Amino acids (147), Protein existence (Evidence at protein level), and Annotation score (5/5). Below this, there are tabs for Entry, Variant viewer, Feature viewer, Genomic coordinates, Publications, and External links. The "Function" section is expanded, showing the protein's role in oxygen transport and its function as an endogenous inhibitor of enkephalin-degrading enzymes.

**Function**  
Involved in oxygen transport from the lung to the various peripheral tissues. [1 Publication](#)

LVV-hemorphin-7 potentiates the activity of bradykinin, causing a decrease in blood pressure.

**Spinorphin**  
Functions as an endogenous inhibitor of enkephalin-degrading enzymes such as DPP3, and as a selective antagonist of the DPP3 receptor which is involved in pain signaling. These properties indicate it as a regulator of pain and

AppsPlaces

Fri Jan 2 21:45

46%74%↓ 1.6 kB100%

WhatsApp(3190) Crime PaHBB - Hemoglorest.uniprot.org/unGitHubNCBI Blast:sp|P1A0U(MMDB) in

https://rest.uniprot.org/uniprotkb/P68871.fasta

☆🔍📄☰

For quick access, place your bookmarks here on the bookmarks toolbar. [Manage bookmarks...](#)

>sp|P68871|HBB\_HUMAN Hemoglobin subunit beta OS=Homo sapiens OX=9606 GN=HBB PE=1 SV=2  
MVHLTPEEKSAVTALWGKVNVDVGGGALGRLLVVPWTRFFESFGDLSTPDVAVMGNPK  
VKAHGKKVLGAFSDGLAHLNKGTFATLSELHCDKLHVDPENFRLLGNVLCVLAHHFG  
KEFTPPVQAAAYQKVAVGAVANALAHKYH

AppsPlaces

Fri Jan 2 21:46

42%79%↓ 110 kB100%

(39) WhatsApp(3190) Crime PaHBB - Hemoglorest.uniprot.org/unGitHubNCBI Blast:sp|P1A0U(MMDB) in

https://blast.ncbi.nlm.nih.gov/Blast.cgi

☆🔍📄☰

For quick access, place your bookmarks here on the bookmarks toolbar. [Manage bookmarks...](#)

DownloadGenPeptGraphics

▼ Next ▲ Previous ◀ Descriptions

**beta globin [Homo sapiens]**  
Sequence ID: [AAZ39780.1](#) Length: 147 Number of Matches: 1  
[See 4 more title\(s\)](#) [See all Identical Proteins \(IPG\)](#)

Range 1: 1 to 147GenPeptGraphics

▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps
299 bits(766)	4e-102	Compositional matrix adjust.	146/147(99%)	147/147(100%)	0/147(0%)
Query 1	MVHLTPEEKSAVTALWGKVNVDVGGGALGRLLVVPWTRFFESFGDLSTPDVAVMGNPK	60			
Sbjct 1	MVHLTP+EKSAVTALWGKVNVDVGGGALGRLLVVPWTRFFESFGDLSTPDVAVMGNPK	60			
Query 61	VKAHGKKVLGAFSDGLAHLNKGTFATLSELHCDKLHVDPENFRLLGNVLCVLAHHFG	120			
Sbjct 61	VKAHGKKVLGAFSDGLAHLNKGTFATLSELHCDKLHVDPENFRLLGNVLCVLAHHFG	120			
Query 121	KEFTPPVQAAAYQKVAVGAVANALAHKYH	147			
Sbjct 121	KEFTPPVQAAAYQKVAVGAVANALAHKYH	147			

**Related Information**  
[Gene](#) - associated gene details  
[Identical Proteins](#) - Identical proteins to AAZ39780.1

DownloadGenPeptGraphics

▼ Next ▲ Previous ◀ Descriptions

**hemoglobin beta chain [Homo sapiens]**  
Sequence ID: [AAD19696.1](#) Length: 147 Number of Matches: 1

Range 1: 1 to 147GenPeptGraphics

▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps
299 bits(766)	6e-102	Compositional matrix adjust.	146/147(99%)	146/147(99%)	0/147(0%)

**Related Information**  
[Gene](#) - associated gene details

Feedback

Apps Places Fri Jan 2 21:46 55% 81% 14 kB 100%

WhatsApp (3190) Crime Po HBB - Hemogio rest.uniprot.org/un GitHub NCBI Blast:sp 1A0U(MMDB)

https://blast.ncbi.nlm.nih.gov/Blast.cgi

For quick access, place your bookmarks here on the bookmarks toolbar. [Manage bookmarks...](#)

Descriptions Graphic Summary **Alignments** Taxonomy

Alignment view Pairwise Restore defaults Download

67 sequences selected

Download GenPept Graphics Next Previous Descriptions

**hemoglobin subunit beta [Homo sapiens]**  
Sequence ID: [NP\\_000509.1](#) Length: 147 Number of Matches: 1  
[See 51 more title\(s\)](#) [See all Identical Proteins \(IPG\)](#)

Range 1: 1 to 147 GenPept Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps
301 bits(770)	1e-102	Compositional matrix adjust.	147/147(100%)	147/147(100%)	0/147(0%)

Query 1 MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTORFFESFGDLSTPDAMGNPK 60  
Sbjct 1 MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTORFFESFGDLSTPDAMGNPK 60

Query 61 VKAHGKKVLGAFSDGLAHLNLIKGTFTATLSELHCDKLHVDPENFRLLGNVLCVLAHHFG 120  
Sbjct 61 VKAHGKKVLGAFSDGLAHLNLIKGTFTATLSELHCDKLHVDPENFRLLGNVLCVLAHHFG 120

Query 121 KEFTPPVQAAAYQKVAGVANALAHKYH 147  
Sbjct 121 KEFTPPVQAAAYQKVAGVANALAHKYH 147

**Related Information**  
[Gene](#) - associated gene details  
[AlphaFold Structure](#) - 3D structure displays  
[Genome Data Viewer](#) - aligned genomic context  
[Identical Proteins](#) - Identical proteins to NP\_000509.1

Feedback

Apps Places Fri Jan 2 21:46 55% 81% 14 kB 100%

WhatsApp (3190) Crime Po HBB - Hemogio rest.uniprot.org/un GitHub NCBI Blast:sp 1A0U(MMDB)

https://blast.ncbi.nlm.nih.gov/Blast.cgi

For quick access, place your bookmarks here on the bookmarks toolbar. [Manage bookmarks...](#)

Descriptions Graphic Summary **Alignments** Taxonomy

Alignment view Pairwise Restore defaults Download

67 sequences selected

Download GenPept Graphics Next Previous Descriptions

**hemoglobin subunit beta [Homo sapiens]**  
Sequence ID: [NP\\_000509.1](#) Length: 147 Number of Matches: 1  
[See 51 more title\(s\)](#) [See all Identical Proteins \(IPG\)](#)

Range 1: 1 to 147 GenPept Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps
301 bits(770)	1e-102	Compositional matrix adjust.	147/147(100%)	147/147(100%)	0/147(0%)

Query 1 MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTORFFESFGDLSTPDAMGNPK 60  
Sbjct 1 MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTORFFESFGDLSTPDAMGNPK 60

Query 61 VKAHGKKVLGAFSDGLAHLNLIKGTFTATLSELHCDKLHVDPENFRLLGNVLCVLAHHFG 120  
Sbjct 61 VKAHGKKVLGAFSDGLAHLNLIKGTFTATLSELHCDKLHVDPENFRLLGNVLCVLAHHFG 120

Query 121 KEFTPPVQAAAYQKVAGVANALAHKYH 147  
Sbjct 121 KEFTPPVQAAAYQKVAGVANALAHKYH 147

**Related Information**  
[Gene](#) - associated gene details  
[AlphaFold Structure](#) - 3D structure displays  
[Genome Data Viewer](#) - aligned genomic context  
[Identical Proteins](#) - Identical proteins to NP\_000509.1

Feedback

Apps Places Fri Jan 2 21:46 55% 80% 0.6 kB 100%

(39) WhatsApp (3190) Crime Po HBB - Hemogio rest.uniprot.org/un GitHub NCBI Blast:sp 1A0U(MMDB)

https://blast.ncbi.nlm.nih.gov/Blast.cgi

For quick access, place your bookmarks here on the bookmarks toolbar. [Manage bookmarks...](#)

Download GenPept Graphics Next Previous Descriptions

**beta globin chain variant [Homo sapiens]**  
Sequence ID: [AAN84548.1](#) Length: 147 Number of Matches: 1

Range 1: 1 to 147 GenPept Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps
299 bits(766)	4e-102	Compositional matrix adjust.	146/147(99%)	147/147(100%)	0/147(0%)

Query 1 MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTORFFESFGDLSTPDAMGNPK 60  
Sbjct 1 MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTORFFESFGDLSTPDAMGNPK 60

Query 61 VKAHGKKVLGAFSDGLAHLNLIKGTFTATLSELHCDKLHVDPENFRLLGNVLCVLAHHFG 120  
Sbjct 61 VKAHGKKVLGAFSDGLAHLNLIKGTFTATLSELHCDKLHVDPENFRLLGNVLCVLAHHFG 120

Query 121 KEFTPPVQAAAYQKVAGVANALAHKYH 147  
Sbjct 121 KEFTPPVQAAAYQKVAGVANALAHKYH 147

Download GenPept Graphics Next Previous Descriptions

**beta-globin [Homo sapiens]**  
Sequence ID: [ACU56984.1](#) Length: 147 Number of Matches: 1

Range 1: 1 to 147 GenPept Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps
299 bits(766)	4e-102	Compositional matrix adjust.	146/147(99%)	147/147(100%)	0/147(0%)

Query 1 MVHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTORFFESFGDLSTPDAMGNPK 60

**Related Information**  
[Gene](#) - associated gene details  
[AlphaFold Structure](#) - 3D structure displays

Feedback

## Discussion

The BLAST analysis confirmed that the selected sequence belongs to Hemoglobin Subunit Beta protein. The high identity and low E-value indicate strong evolutionary conservation and biological relevance of the sequence.

## Conclusion

This task successfully demonstrated the use of BLAST for protein sequence analysis. It enhanced understanding of sequence alignment and interpretation of BLAST outputs in bioinformatics research.

## Learning Outcome

- Understanding FASTA format
- Hands-on experience with BLAST
- Ability to interpret alignment results

## Acknowledgment

I would like to thank **CodeAlpha** for providing this opportunity to learn and practice bioinformatics concepts through hands-on tasks.