

Introduction to Bio Science and Technology

Name : Chagam Krishnaja

Roll No: 22CS01024

Assignment Solutions :

A. AGU CUC UGU CUC CAU UUG AAG AAG GGG AAG GGG

a) By using the Genetic Code Table

		Second letter				
		U	C	A	G	
First letter	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA STOP UAG STOP	UGU } Cys UGC } UGA STOP UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } Ile AUC } AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } Val GUC } GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }	U C A G

B. The given mRNA is the decoded version of the given genetic code into an amino acid sequence .It is not a full length protein sequence, usually protein synthesis begins with Met and the Stop codons (or nonsense codons) terminate translation which cant be seen in the given mRNA.

➤ Amino Acid Sequence:

AGU CUC UGU CUC CAU UUG AAG AAG GGG AAG GGG

Ser Leu Cys Leu His Leu Lys Lys Gly Lys Gly

b)

B.

- 1) **BLAST** stands for **Basic Local Alignment Search Tool**. The primary purpose of BLAST is to compare protein and gene sequences against available public databases. It is a set of sequence comparison algorithms used to search databases for an optimal local alignment against the query. It breaks the obtained queries and database sequences into fragments and matches them.
- 2) Length of Query Sequence (A) = 2245
- 3) Sequence Name = H. Sapiens mRNA for Serum Albumin
Organism = Homo Sapiens
Accession ID = V00495.1
- 4) The Expect value (E) is a parameter that describes number of hits one can “expect” to see by chance when searching a database of a particular size.
It decreases exponentially when Score (S) of match increases.
When we get an E value zero for 100% identity, it means that the query sequence aligns perfectly with the database sequence.
- 5) “**Bos taurus mRNA for bovine serum albumin**” Sequence (B)

```
GTCTGACTTTGGCACAATGAAGTGGGTGACTTTTATTTCTCTTCTCCTTCTCTTCAGCTCTGCTTATTCCA
GGGGTGTGTTTTCTGTCGAGATACACACAAGAGTGAGATTGCTCATCGGTTTAAAGATTTGGGAGAAGAACA
TTTTAAAGGCCTGGTACTGATTGCCTTTTCTCAGTATCTCCAGCAGTGTCCATTTGATGAGCATGTAAAA
TTAGTGAACGAACTAACTGAGTTTGCAAAAACATGTGTTGCTGATGAGTCCCATGCCGGCTGTGAAAAGT
CACTTCACACTCTCTTTGGAGATGAATTGTGTAAAGTTGCATCCCTTCGTGAAACCTATGGTGACATGGC
TGACTGCTGTGAGAAACAAGAGCCTGAAAGAAATGAATGCTTCCTGAGCCACAAAGATGATAGCCCAGAC
CTCCCTAAATTGAAACCAGACCCCAATACTTTGTGTGATGAGTTTAAAGGCAGATGAAAAGAAGTTTTGGG
GAAAATACCTATACGAAATTGCTAGAAGACATCCCTACTTTTATGCACCAGAACTCCTTTACTATGCTAA
TAAATATAATGGAGTTTTTTCAAGAATGCTGCCAAGCTGAAGATAAAGGTGCCTGCCTGCTACCAAAGATT
GAACTATGAGAGAAAAAGTACTGACTTCATCTGCCAGACAGAGACTCAGGTGTGCCAGTATTCAAAAAT
TTGGAGAAAGAGCTTTTAAAGCATGGTCAGTAGCTCGCCTGAGCCAGAAAATTTCCCAAGGCTGAGTTTGT
AGAAGTTACCAAGCTAGTGACAGATCTCACAAAAGTCCACAAGGAATGCTGCCATGGTGACCTACTTGAA
TGCGCAGATGACAGGGCAGATCTTGCCAAGTACATATGTGATAATCAAGATACAATCTCCAGTAAACTGA
AGGAATGCTGTGATAAGCCTTTGTTGGAAAAATCCCACTGCATTGCTGAGGTAGAAAAAGATGCCATACC
TGAAAACCTGCCCCCATTAAGTCTGCTGACTTTGCTGAAGATAAGGATGTTTGCAAAAACCTATCAGGAAGCA
AAAGATGCCTTCCTGGGCTCGTTTTTTGTATGAATATTCAAGAAGGCATCCTGAATATGCTGTCTCAGTGC
TATTGAGACTTGCCAAGGAATATGAAGCCACACTGGAGGAATGCTGTGCCAAAGATGATCCACATGCATG
CTATTCCACAGTGTTTGACAACTTAAGCATCTTGTGGATGAGCCTCAGAATTTAATCAAACAAAACCTGT
GACCAATTCGAAAAACTTGGAGAGTATGGATTCCAAAATGCGCTCATAGTTCGTTACACCAGGAAAGTAC
CCCAAGTGTCAACTCCAACCTCTCGTGGAGGTTTCAAGAAGCCTAGGAAAAGTGGGTACTAGGTGTTGTAC
AAAGCCGGAATCAGAAAGAATGCCCTGTACTGAAGACTATCTGAGCTTGATCCTGAACCGGTTGTGCGTG
CTGCATGAGAAGACACCAGTGAGTGAAAAAGTCACCAAGTGCTGCACAGAGTCATTGGTGAACAGACGGC
CATGTTTCTCTGCTCTGACACCTGATGAAACATATGTACCCAAAGCCTTTGATGAGAAAATTGTTACCTT
CCATGCAGATATATGCACACTTCCCATACTGAGAAACAAATCAAGAAACAACTGCACTTGTTGAGCTG
TTGAAACACAAGCCCAAGGCAACAGAGGAACAACTGAAAACCGTCATGGAGAATTTTGTGGCTTTTGTAG
ACAAGTGCTGTGCAGCTGATGACAAAGAGGCCTGCTTTGCTGTGGAGGGTCCAAAACCTTGTTGTTTCAAC
TCAAACAGCCTTAGCCTAAACACGACACAACCACAGGCATCTCAGCCTACCCTGAGAGTCGAC
```

- 6) Length of Nucleotide Sequence (B) = 1883

7) 82.19% of identity was found between Sequence (A) and (B)

8) Length of Amino Acid sequence BSA = 607

9) Crystal Structure of Bovine Serum Albumin [Transport Protein]

PDB ID: 3V03

Journal Name: Structural and immunologic characterization of bovine, horse, and rabbit serum albumins

Mol Immunol (2012) **52** p.174-82

Authors: Majorek KA, Porebski PJ, Dayal A, Zimmerman MD, Jablonska K, Stewart AJ, Chruszcz M, Minor W

Crystal Structure of Bovine Serum Albumin [Transport Protein]

PDB ID: 4F5S

Journal Name: Structures of bovine, equine and leporine serum albumin

Acta Crystallogr D Biol Crystallogr (2012) **68** p.1278-89

Authors: Bujacz A

10) Loading 3V03 and 4F5S PDB Files into PyMOL and superimposing the two structures gives the following result:

