

Lab 05: Central Dogma I: Integrated pipeline for Building Web-based DNA Translation Tool

Your Task

To receive full credit for this lab, you must **answer** all following questions.

Task 1: Get protein sequence for DNA sequence

(https://www.ncbi.nlm.nih.gov/nuccore/NM_207618.2), and compare it with sequence on NCBI website. Do you get same results?

```
GGTCAGAAAAAGCCCTCTCCATGTCTACTCACGATACATCCCTGAAAACCACTGAGGAAGTGGC
TTTTTCAGATCATCTTGCTTTGCCAGTTTGGGGTTGGGACTTTTGCCAATGTATTTCTCTTTGTC
TATAATTTCTCTCCAATCTCGACTGGTTCTAAACAGAGGCCAGACAAGTGATTTTAAGACACA
TGGCTGTGGCCAATGCCTTAACTCTCTTCCTCACTATATTTCCAAACAACATGATGACTTTTGC
TCCAATTATTCTCAAACTGACCTCAAATGTAAATTAGAATTCTTCACTCGCCTCGTGGCAAGA
AGCACAAACTTGTGTTCAACTTGTGTTCTGAGTATCCATCAGTTTGTCACTTGTTCCTGTTA
ATTGAGGTAAAGGAATACTCAGAGCAAGTGTCAAAACATGGCAAGTTATTCTTGTACAGTTG
TTGTTTCTTCAAGTGTCTTAAATAACATCTACATTCCAATTAAGGTCAGTGGTCCACAGTTAACA
GACAATAACAATACTCTAAAAGCAAGTTGTTCTGTTCCACTTCTGATTTCAAGTGTAGGCATTG
TCTTCTTGAGGTTTGCCCATGATGCCACATTCATGAGCATCATGGTCTGGACCAGTGTCTCCAT
GGTACTTCTCCTCCATAGACATTGTCAGAGAATGCAGTACATATTCACTCTCAATCAGGACCCC
AGGGGCCAAGCAGAGACCACAGCAACCCATACTATCCTGATGCTGGTAGTCACATTTGTTGGCT
TTTATCTTCTAAGTCTTATTTGTATCATCTTTTACACCTATTTTATATATTCTCATCATTCCTT
GAGGCATTGCAATGACATTTTGGTTTTCGGGTTTCCCTACAATTTCTCCTTTACTGTTGACCTTC
AGAGACCCTAAGGGTCCTTGTGTTCTGTGTTCTTCAACTGTTGAAAGCCAGAGTCACTAAAAATGC
CAAACACAGAAGACAGCTTTGCTAATACCATTAAATACTTTATTCCATAAATATGTTTTTAAAA
GCTTGTATGAACAAGGTATGGTGCTCACTGCTATACTTATAAAAGAGTAAGGTTATAATCACTT
GTTGATATGAAAAGATTTCTGGTTGGAATCTGATTGAAACAGTGAGTTATTCACCACCCTCCAT
TCTCT
```

- Your Protein Sequence: ['M', 'S', 'T', 'H', 'D', 'T', 'S', 'L', 'K', 'T', 'T', 'E', 'E', 'V', 'A', 'F', 'Q', 'I', 'T', 'L', 'L', 'C', 'Q', 'F', 'G', 'V', 'G', 'T', 'F', 'A', 'N', 'V', 'F', 'L', 'F', 'V', 'Y', 'N', 'F', 'S', 'P', 'I', 'S', 'T', 'G', 'S', 'K', 'Q', 'R', 'P', 'R', 'Q', 'V', 'I', 'L', 'R', 'H', 'M', 'A', 'V', 'A', 'N', 'A', 'L', 'T', 'L', 'F', 'L', 'T', 'I', 'F', 'P', 'N', 'N', 'M', 'M', 'T', 'F', 'A', 'P', 'I', 'I', 'P', 'Q', 'T', 'D', 'L', 'K', 'C', 'K', 'L', 'E', 'F', 'F', 'T', 'R', 'L', 'V', 'A', 'R', 'S', 'T', 'N', 'L', 'C', 'S', 'T', 'C', 'V', 'L', 'S', 'I', 'H', 'Q', 'F', 'V', 'T', 'L', 'V', 'P', 'V', 'N', 'S', 'G', 'K', 'G', 'I', 'L', 'R', 'A', 'S', 'V', 'T', 'N', 'M', 'A', 'S', 'Y', 'S', 'C', 'Y', 'S', 'C', 'W', 'F', 'F', 'S', 'V', 'L', 'N', 'N', 'T', 'Y', 'I', 'P', 'I', 'K', 'V', 'T', 'G', 'P', 'Q', 'L', 'T', 'D', 'N', 'N', 'N', 'N', 'S', 'K', 'S', 'K', 'L', 'F', 'C', 'S', 'T', 'S', 'D', 'F', 'S', 'V', 'G', 'I', 'V', 'F', 'L', 'R', 'F', 'A', 'H', 'D', 'A', 'T', 'F', 'M', 'S', 'I', 'M', 'V', 'W', 'T', 'S', 'V', 'S', 'M', 'V', 'L', 'L', 'L', 'H', 'R', 'H', 'C', 'Q', 'R', 'M', 'Q', 'Y', 'I', 'F', 'T', 'L', 'N', 'Q', 'D', 'P', 'R', 'G', 'Q', 'A', 'E', 'T', 'T', 'A', 'T', 'H', 'T', 'I', 'L', 'M', 'L', 'V', 'V', 'T', 'F', 'V', 'G', 'F', 'Y', 'L', 'L', 'S', 'L', 'I', 'C', 'I', 'I', 'F', 'Y', 'T', 'Y', 'F', 'I', 'Y', 'S', 'H', 'H', 'S', 'L', 'R', 'H', 'C', 'N', 'D', 'I', 'L', 'V', 'S', 'G', 'F', 'P', 'T', 'I', 'S', 'P', 'L', 'L', 'L', 'T', 'F', 'R', 'D', 'P', 'K', 'G', 'P', 'C', 'S', 'V', 'F', 'F', 'N', 'C']

NCBI's Protein Sequence: MSTHDTSLKTTEEVAFQIILLCQFGVGTFFANVFLFVYNFSPIST
 GSKQRPRQVILRHMAVANALTLFLTIFPNNMMTFAPIIPQTDLKCKLEFFTRLVARST
 NLCSTCVLSIHQFVTLVPVNSGKGILRASVTNMASYSCYSCWFFSVLNNIYIPIKVTG
 PQLTDNNNNNSKSKLFCSTSDFSVGIVFLRFAHDATFMSIMVWTSVSMVLLLHRHCQRM
 QYIFTLNQDPRGQAETTATHTILMLVVTFFVGFYLLSLICIFITYFYIYSHHSLRHCND
 ILVSGFPTISPLLLTFRDPKGPSCSVFFNC

Task 2: Get protein sequence for DNA sequence

(https://www.ncbi.nlm.nih.gov/nuccore/NM_008220.5), and compare it with sequence on NCBI website. Do you get same results?

TTACGTTTGCTTCTGATTCTGTTGTGTTGACTTGCAACCTCAGAAACAGACATCATGGTGCACC
 TGACTGATGCTGAGAAGGCTGCTGTCTCTGGCCTGTGGGGAAAGGTGAACGCCGATGAAGTTGG
 TGGTGAGGCCCTGGGCAGGCTGCTGGTTGTCTACCCTTGGACCCAGCGGTACTTTGATAGCTTT
 GGAGACCTATCCTCTGCCTCTGCTATCATGGGTAATGCCAAAGTGAAGGCCCATGGCAAGAAAG
 TGATAACTGCCTTTAACGATGGCCTGAATCACTTGGACAGCCTCAAGGGCACCTTTGCCAGCCT
 CAGTGAGCTCCACTGTGACAAGCTGCATGTGGATCCTGAGAACTTCAGGCTCCTGGGCAATATG
 ATCGTGATTGTGCTGGGCCACCACCTGGGCAAGGATTTACACCCCGCTGCACAGGCTGCCTTCC
 AGAAGGTGGTGGCTGGAGTGGCTGCTGCCCTGGCTCACAAGTACCACTAAGCCCCCTTTCTGCT
 ATTGTCTATGCACAAAGGTTATATGTCCCCTAGAGAAAAACTGTCAATTGTGGGGAAATGATGA
 AGACCTTTGGGCATCTAGCTTTTATCTAATAAATGATATTTACTGTCATCTCAAAAAAAAAAAAA
 AAAAAA

- Your Protein Sequence: ['M', 'V', 'H', 'L', 'T', 'D', 'A', 'E', 'K', 'A', 'A', 'V', 'S', 'G', 'L', 'W', 'G', 'K', 'V', 'N', 'A', 'D', 'E', 'V', 'G', 'G', 'E', 'A', 'L', 'G', 'R', 'L', 'L', 'V', 'V', 'Y', 'P', 'W', 'T', 'Q', 'R', 'Y', 'F', 'D', 'S', 'F', 'G', 'D', 'L', 'S', 'S', 'A', 'S', 'A', 'T', 'M', 'G', 'N', 'A', 'K', 'V', 'K', 'A', 'H', 'G', 'K', 'K', 'V', 'I', 'T', 'A', 'F', 'N', 'D', 'G', 'L', 'N', 'H', 'L', 'D', 'S', 'L', 'K', 'G', 'T', 'F', 'A', 'S', 'L', 'S', 'E', 'L', 'H', 'C', 'D', 'K', 'L', 'H', 'V', 'D', 'P', 'E', 'N', 'F', 'R', 'L', 'L', 'G', 'N', 'M', 'I', 'V', 'I', 'V', 'L', 'G', 'H', 'L', 'G', 'K', 'D', 'F', 'T', 'P', 'A', 'A', 'Q', 'A', 'A', 'F', 'Q', 'K', 'V', 'V', 'A', 'G', 'V', 'A', 'A', 'A', 'L', 'A', 'H', 'K', 'Y', 'H']

NCBI's Protein Sequence:

MVHLTDAEKAAVSGLWGKVNADDEVGGEALGRLLVVYPWTQRYFD

SFGDLSSASAIMGNAKVKAHGKKVITAFNDGLNHLDSLKGTFASLSELHCDKLHVDPE
 NFRLGNMIVIVLGHHLGKDFTPAQAQAFQKVAGVAAALAHKYH

Task 3: Get protein sequence for DNA sequence

(https://www.ncbi.nlm.nih.gov/nuccore/NM_025580.2), and compare it with sequence on NCBI website. Do you get same results?

```
CGACGGGGAGCGCAGTGAAGGGGGGTGGGACCCGAACATGGCGGCGGTGGTAGCTGCTACGGCG
CTGAAGGGCCGGGGGGCGAGAAATGCCCGCGTCTCTCCGGGGGATCCTCTCTGGAGCCACAGCTA
ACAAGGCTTCCCAGAACAGGACCAGAGCACTGCAGAGCCACAGCTCACCAGAATGCAAGGAGGA
GCCCCAGCCCCCTCTCCCCTGAGCTAGAATACATTCCCAGAAAGAGGGGCAAGAACCCCATGAAA
GCCGTGGGGCTAGCCTGGGCCATCGGCTTCCCCTGTGGTATCCTCTTCTTCGTCCTCACCAAGC
AGGAAGTGGACAAGGACCGATTGAAGCAGATGAAAGCTCGTCAGAACATGCGGGTGTCCAACAC
GGGCGAGTATGAGAGCCAGAGGTTTCAGGGCTTCACCCCAACAAGCCCAGTTCCCTGAGGTTGGG
TCTGGGGCCCAGACCTGAGTGAGGATCACTGCAACCTTGTGCTAGACTATTGACTGCCGTCTTG
TGTGGCCATGGCTCCTGAGGCCCACCAGAGGGCGCAGGATGCCCAAGCTGCTGCCAGCACCCAA
CTCTTCCACACCGTGGAAGGAGCCAAATGCAAATAAAGTTATTAAGTACTGACACGGAAAAAAA
AA
```

- Your Protein Sequence: ['M', 'A', 'A', 'V', 'V', 'A', 'A', 'T', 'A', 'L', 'K', 'G', 'R', 'G', 'A', 'R', 'N', 'A', 'R', 'V', 'L', 'R', 'G', 'T', 'L', 'S', 'G', 'A', 'T', 'A', 'N', 'K', 'A', 'S', 'Q', 'N', 'R', 'T', 'R', 'A', 'L', 'Q', 'S', 'H', 'S', 'S', 'P', 'E', 'C', 'K', 'E', 'E', 'P', 'E', 'P', 'L', 'S', 'P', 'E', 'L', 'E', 'Y', 'I', 'P', 'R', 'K', 'R', 'G', 'K', 'N', 'P', 'M', 'K', 'A', 'V', 'G', 'L', 'A', 'W', 'A', 'I', 'G', 'F', 'P', 'C', 'G', 'I', 'L', 'F', 'F', 'V', 'L', 'T', 'K', 'Q', 'E', 'V', 'D', 'K', 'R', 'L', 'K', 'Q', 'M', 'K', 'A', 'R', 'Q', 'N', 'M', 'R', 'V', 'S', 'N', 'T', 'G', 'E', 'Y', 'E', 'S', 'Q', 'R', 'F', 'R', 'A', 'S', 'P', 'Q', 'Q', 'A', 'Q', 'F', 'P', 'E', 'V', 'G', 'S', 'G', 'A', 'Q', 'T']

NCBI's Protein Sequence:

MAAVVAATALKGRGARNARVLRGILSGATANKASQNRTRALQSH

SSPECKEEPEPLSPELEYIPRKRKGKNPMKAVGLAWAIGFPCGILFFVLTKQEVDKDRL
KQMKARQNM RVSN TGEYESQRFRASPQQAQFPEVGSGAQT

Task 4: Get protein sequence for DNA sequence

(https://www.ncbi.nlm.nih.gov/nucore/NM_009735.3), and compare it with sequence on NCBI website. Do you get same results?

```
TTTCAGTGGCTGCTACTCGGCGCTTCAGTCGCGGTTCGCTTCAGTCGTCAGCATGGCTCGCTCGG
TGACCCTGGTCTTTCTGGTGCTTGTCTCACTGACCGGCCTGTATGCTATCCAGAAAACCCCTCA
AATTCAAGTATACTCACGCCACCCACCGGAGAATGGGAAGCCGAACATACTGAACTGCTACGTA
ACACAGTTCCACCCGCCTCACATTGAAATCCAAATGCTGAAGAACGGGAAAAAAATTCCTAAAG
TAGAGATGTCAGATATGTCCTTCAGCAAGGACTGGTCTTTCTATATCCTGGCTCACACTGAATT
CACCCCCACTGAGACTGATACATACGCCTGCAGAGTTAAGCATGCCAGTATGGCCGAGCCCAAG
ACCGTCTACTGGGATCGAGACATGTGATCAAGCATCATGATGCTCTGAAGATTCATTTGAACCT
GCTTAATTACAAATCCAGTTTCTAATATGCTATACAATTTATGCACGCAGAAAGAAATAGCAAT
GTACACATCACCTTCTTTATATCTTACTTTAAATATTTTATGCATGTTTTCAAAAAATTGAAA
TATCCTAGATAGCTGAGCAATAAATCTTCAATAAGTATTTTGATCAGAATAATAAATATAATTT
TAAGAACAATAGTTGATCATATGCCAAACCCTCTGTACTTCTCATTACTTGGATGCAGTTACTC
ATCTTTGGTCTATCACAACATAAGTGACATACTTTCCTTTTGGTAAAGCAAAGAGGCCTAATTG
AAGTCTGTCACTGTGCCCAATGCTTAGCAATTCTCACCCCCAACCCCTGTGGCTACTTCTGCTTT
TGTTACTTTTACTAAAAATAAAAAAC
```

- Your Protein Sequence: ['M', 'A', 'R', 'S', 'V', 'T', 'L', 'V', 'F', 'L', 'V', 'L', 'V', 'S', 'L', 'T', 'G', 'L', 'Y', 'A', 'I', 'Q', 'K', 'T', 'P', 'Q', 'I', 'Q', 'V', 'Y', 'S', 'R', 'H', 'P', 'P', 'E', 'N', 'G', 'K', 'P', 'N', 'I', 'L', 'N', 'C', 'Y', 'V', 'T', 'Q', 'F', 'H', 'P', 'P', 'H', 'I', 'E', 'I', 'Q', 'M', 'L', 'K', 'N', 'G', 'K', 'K', 'I', 'P', 'K', 'V', 'E', 'M', 'S', 'D', 'M', 'S', 'F', 'S', 'K', 'D', 'W', 'S', 'F', 'Y', 'I', 'L', 'A', 'H', 'T', 'E', 'F', 'T', 'P', 'T', 'E', 'D', 'T', 'Y', 'A', 'C', 'R', 'V', 'K', 'H', 'A', 'S', 'M', 'A', 'E', 'P', 'K', 'T', 'V', 'Y', 'W', 'D', 'R', 'D', 'M']

NCBI's Protein Sequence:

MARSVTLVFLVLVSLTGLYAIQKTPQIQVYSRHPPENGKPNILN

CYVTQFHPPHIEIQMLKNGKKIPKVEMSDMSFSKDWSFYILAHTFTPTETDTYACRV
KHAASMAEPKTVYWDRDM

Task 5: Get protein sequence for DNA sequence

(https://www.ncbi.nlm.nih.gov/nucore/NM_011662), and compare it with sequence on NCBI website. Do you get same results?

```
GTCACACCAGGTCCCACCAGCCCCTGGACTGTGGTGTCCAGTGCATATCTGGCCACCATGGGGG
CTCTGGAGCCCTCCTGGTGCCTTCTGTTCCCTTCCTGTCCTCCTGACTGTGGGAGGATTAAGTCC
CGTACAGGCCCAGAGTGACACTTTCCTCAAGATGCGACTGTTCTTCCGTGAGCCCTGGTGTACTG
GCTGGGATTGTTCTGGGTGACTTGGTGTGACTCTGCTGATTGCCCTGGCTGTGTACTCTCTGG
GCCGCCTGGTCTCCCGAGGTCAAGGGACAGCGGAAGGGACCCGAAACAACACATTGCTGAGAC
TGAGTCGCCTTATCAGGAGCTTCAGGGTCAGAGACCAGAAGTATACAGTGACCTCAACACACAG
AGGCAATATTACAGATGAGCCCACTCTATGCCCATCAGCGGCCTGATGCCCCGGATCCGGTCATT
CCAGATGCCTACTCAACAAGCCCTCTCTGGGATCAGGACTCCCGTTGGAATACAGATCCACAGG
GTACCTCCCTGAGATATCTGACATTGTACCATTCTGTCCCCAAATAAAAGACAGAGCAAAAAA
AAAAAAAAAAAAAAAAAAAA
```

- Your Protein Sequence: ['M', 'G', 'A', 'L', 'E', 'P', 'S', 'W', 'C', 'L', 'L', 'F', 'L', 'P', 'V', 'L', 'L', 'T', 'V', 'G', 'G', 'L', 'S', 'P', 'V', 'Q', 'A', 'Q', 'S', 'D', 'T', 'F', 'P', 'R', 'C', 'D', 'C', 'S', 'S', 'V', 'S', 'P', 'G', 'V', 'L', 'A', 'V', 'L', 'G', 'D', 'L', 'V', 'L', 'T', 'L', 'L', 'T', 'A', 'L', 'A', 'V', 'Y', 'S', 'L', 'G', 'R', 'L', 'V', 'S', 'R', 'G', 'Q', 'G', 'T', 'A', 'E', 'G', 'T', 'R', 'K', 'Q', 'H', 'I', 'A', 'E', 'T', 'S', 'P', 'Y', 'Q', 'E', 'L', 'Q', 'G', 'Q', 'R', 'P', 'E', 'V', 'Y', 'S', 'D', 'L', 'N', 'T', 'Q', 'R', 'Q', 'Y', 'Y', 'R']

NCBI's Protein Sequence:

MGALEPSWCLLFLPVLLTVGGLSPVQAQSDTFPRCDCSSVSPGV

LAGIVLGDVLTLIALAVYSLGRLVSRGQGTAEGTRKQHIAETESPYQELQGQRPEV
YSDLNTQRQYYR

Task 6: Give a brief overview of your work, discussing any thoughts, successes or challenges you faced.

Overall, the work for the main parts was not hard, just reusing the code from lab04 and implementing the flask and html to run together. Then it was just a matter of copy pasting everything. The hardest part was the optional part which I got stuck on because I thought I was messing up the first frame, but I was actually getting it right. I was just comparing it to the frame from lab04 but the frame from lab04 is actually the third reading frame, but once I realized that the rest wasn't bad, just a few pauses to think and then just coding it all.

Additional note: for the optional I did not make it print them in separate boxes like ExPasy does, they're all in the one box with which reading frame it is at the start and break written between when a end codon is hit