

1001001100001
0100001110100
0100001110100

Hands-On Three Transcription, Translation and Databases

Problem 1

a) Complete the following table. Assume that

- the reading is from left to right
- the columns represent transcriptional and translational alignments

C													DNA double helix
					T	C	A						mRNA transcribed
	C	A				A							Appropriate tRNA anticodon
								G	C	A			Amino acids incorporated into protein
	Trp (W)												

b) Label 5' and 3' ends of DNA and RNA, as well as the amino and carboxyl ends of proteins. [Introduction to Genetic Analysis by Griffiths et al., 2005]

Problem 2

	Transcription	Translation
Where does this process occur in a eukaryotic cell?		
What is the enzyme that carries out this process?		
What is the template that is read during this process?		
In what direction is the template read?		
What is the start signal/sequence for this process?		
What is the polymer that is formed?		
What monomer is used to form this polymer?		
What type of bond is formed between monomers?		
In what direction is the new polymer formed?		
What is the stop signal/sequence for this process?		

Problem 3

The following double-stranded DNA sequence is part of a hypothetical yeast genome which happens to contain a very small gene. Transcription starts at the Transcription Start Site (TSS),

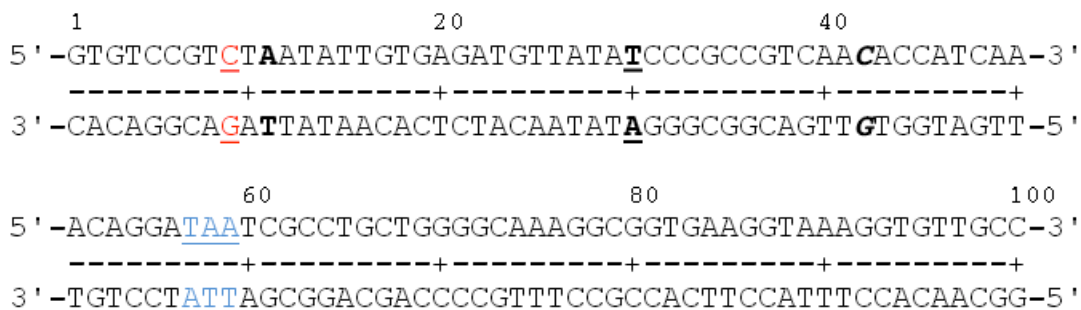
proceeds in the direction of the arrow and stops at the end of the Transcription Terminator (the boxed region).



- Which strand of DNA shown, the top or the bottom, is the template strand?
- What is the sequence of the mRNA produced from this gene? Label the 5' and 3' ends.
- What is the sequence of the protein produced from the mRNA in (b)? Label the N and C termini.
- If a mutation (an insertion) were found where a T/A (top/bottom) base pair were added immediately after the T/A base pair shown in bold, what would be the sequence of the mRNA? What would be the sequence of the protein?

Problem 4

The following double-stranded bacterial (*E. coli*) DNA sequence codes for a hypothetical protein. Both strands are shown; the top strand reads 5' to 3' left to right, while the bottom strand reads 5' to 3' right to left. The nucleotides are numbered 1 to 100. NOTE: Transcription begins with and includes the underlined C/G (top strand/bottom strand) base pair in position 9, and the RNA polymerase proceeds from left to right along the DNA.



- Which strand is used as a template for transcription, the top or the bottom?
- Where would the promoter be relative to the start of transcription?
- What are the first 15 nucleotides of the resulting mRNA? Indicate the 5' and 3' ends.
- What are the first 5 amino acids translated from the resulting mRNA? Indicate the amino and carboxyl termini of the protein.
- Do the underlined nucleotides TAA encode a stop codon for the protein? Explain your answer. Consider the situations in parts (f-h) independently.
- A mutation occurs which results in the insertion of an extra G/C (top strand/bottom strand) base-pair immediately after base pair 11 (A/T shown in bold). What effect will this insertion mutation have on the mRNA transcript and on the resulting protein?
- A different mutation results in the substitution of the T/A base pair at position 30 (shown in bold and underlined) with a G/C base pair. How would this mutation affect the sequence of the protein that is produced?

h) A different mutation occurs which results in the substitution of the C/G base pair at position 42 (shown in bold italics) to a T/A base pair. How would this mutation affect the sequence of the protein that is produced?

Problem 5

NCBI has a sample GenBank record at:

<http://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html>

Please go to that site, read the example and answer the following questions:

a) What does PLN stand for? _____.

How many divisions does GenBank have? _____.

b) What does CDS stand for? _____. Explain.

c) There are three occurrences of “CDS” under “Features”. Consider the first occurrence of “CDS”. One of its subfields is “/translation”.

Explain why “/translation” starts with the specific sequence of amino acids: **SSIYN**.

d) Consider the second occurrence of “CDS”. One of its subfields is “/translation”. Explain why “/translation” starts with the specific sequence of amino acids: **MTQLQ**.

e) [Optional] Consider the third occurrence of “CDS”. One of its subfields is “/translation”.

Explain why “/translation” starts with the specific sequence of amino acids: MNRWV.