

7.014 Solution Set 4

Question 1

Shown below is a fragment of the sequence of a hypothetical bacterial gene. This gene encodes production of CHWDWN, protein essential for metabolizing sugar yummoose. The transcription begins (and includes) the G/C base pair in bold and proceeds to the right.

5' ...TTCGAG**G**CTCTCGTCGTCGAGATACGCGATGATATTAGTGGTAATATGGGGATGCACT...3'
3' ...AAGCT**C**GAGAGCAGCAGCTCTATGCGCTACTATAATCACCATTATAACCCCTACGTGA...5'

- a) Give the sequence of the mRNA transcribed from this gene and indicate the 5' and 3' ends of the mRNA.

5'-GCUCUCGUCGUCGAGAUACGCGAUGAUUUAGUGGUAAUAUGGGGAUGCACU...3'


- b) Give the sequence of the peptide that will be translated from this mRNA. Label the amino and carboxy termini of the peptide.

N-Met-Ile-Leu-Val-Val-Ile-Trp-Gly-Cys-Thr...-C

- c) You study two different mutants, Mutant A and Mutant B.

- i) In the DNA sequence for Mutant A, you find the insertion of a G/C base pair between positions 22 and 23 (position of insertion is indicated by an arrow):

5' ...TTCGAG**G**CTCTCGTCGTCGAGATACGCGATGATATTAGTGGTAATATGGGGATGCACT...3'
3' ...AAGCT**C**GAGAGCAGCAGCTCTATGCGCTACTATAATCACCATTATAACCCCTACGTGA...5'



Give the sequence of the new peptide produced by mutant A. Label the amino and carboxy termini of the peptide.

N-Met-Thr-Arg-C

- ii) In the DNA sequence for Mutant B, two consecutive G/C base pairs are inserted between positions 22 and 23 (position of insertion is indicated by an arrow on the figure above).

Give the sequence of the new peptide produced by mutant B. Label the amino and carboxy termini of the peptide.

N-Met-Asp-Ala-Met-Ile-Leu-Val-Val-Ile-Trp-Gly-Cys-Thr...-C

Question 1, continued

d) One of these two mutants is fully functional, while the other is not. Which mutant peptide do you predict is functional and which one is not? Why?

Peptide A is truncated after three amino acids, so it is likely the one that is not functional. Peptide B has three extra amino acids on the amino terminus with respect to the wild-type peptide. These amino acids are not likely to affect the overall fold of the peptide. Thus, it is likely that peptide B is fully functional.

Question 2

You are fascinated by CHWDWN, and decide to continue your research over the summer.

A graduate student in your lab has developed a collection of strains of bacteria containing different mutant tRNAs.

a) In wild-type cells, what is the anticodon on the tRNA charged with trp? Indicate 5' and 3'.
5'-CCA-3' or 3'-ACC-5'

b) In strain X, the 5' nucleotide of the anticodon on the trp tRNA is changed to a G, and no wild-type trp tRNA is present.

i. Would you expect CHWDWN polypeptide production in X to be affected? If yes, explain how it would be affected. If no, explain why not.

The anticodon now reads 5'-GCA-3'. Because of the mutant anticodon, in strain X the tRNA charged with amino acid trp pairs with the codon 5'-UGC-3'. Moreover, no wild-type trp tRNA exists in strain X. This would cause translation of CHWDWN to terminate after the sixth codon, since no tRNA will be available to translate codon 7, UGG.

Based on the gene sequence in Question 1, CHWDWN protein contains at least one trp codon.

ii. What proteins other than CHWDWN would you expect to be affected? Why?

The codon 5'-UGC-3' normally encodes cys. Thus, whenever the codon UGC appears in frame in the coding sequence of a gene in strain X, part of the time amino acid trp will be inserted instead of cys. In addition, as illustrated in part i above, any time the codon for trp is present in frame in a coding sequence of a gene in strain X, translation will be terminated. So the proteins that would be affected are all those that have codons UGC or UGG in their coding sequences.

iii. Would you expect strain X to grow on media containing yummose as the only carbon source? If yes, how strong would you expect that growth to be with respect to the wild-type strain? If no, explain why you expect no growth.

In order to grow on yummose as the only carbon source, strain X would need to have a functional CHWDWN protein. Since translation of CHWDWN in strain X terminates after only six amino acids, it is very unlikely that strain X could grow on yummose as the only carbon source.

c) In strain Z, the tRNA with the anticodon for trp found in wild-type cells is actually charged with amino acid gln, and no wild-type trp tRNA is present.

i. Would you expect CHWDWN polypeptide production in Z to be affected? If yes, explain how it would be affected. If no, explain why not.

Since the amino acid that a trp tRNA is charged with changes in strain Z, the seventh amino acid of the CHWDWN protein will now be gln, instead of trp.

ii. What proteins other than CHWDWN would you expect to be affected? Why?

Any protein that in a wild-type strain contains amino acid trp will, in strain Z, contain amino acid gln in its place.

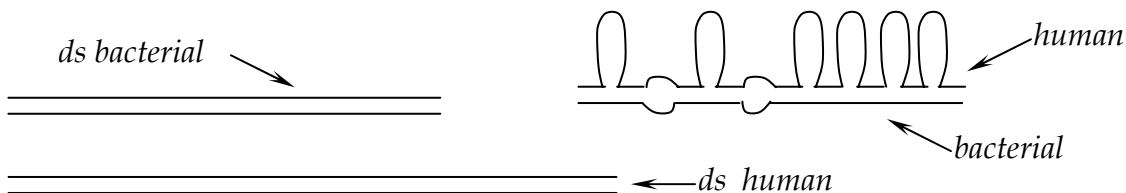
Question 2, continued

- iii. Would you expect strain Z to grow on media containing yummose as the only carbon source? If yes, how strong would you expect that growth to be with respect to the wild-type strain? If no, explain why you expect no growth.

In order to grow on yummose as the only carbon source, strain Z would need to have a functional CHWDWN protein. At least in position 7 of the protein, trp has been substituted by gln. If the codon for trp is contained anywhere else in the gene, the amino acid substitution will occur in that position as well. Gln is a smaller amino acid than trp is, and it is also polar, while trp is hydrophobic. Thus, is it likely that substitution will have some impact on the protein, but it is unclear whether it will render the protein unable to perform its function.

- d) You find that the protein sequence of CHWDWN is highly conserved (~80%) in humans. Excited, you acquire DNA fragments encoding bacteria and human CHWDWN proteins. You
1. combine both samples into one test tube
 2. briefly treat the sample in the test tube with heat
 3. let the sample cool
 4. examine the contents of the test tube in electron microscope.

You find that you have three types of complexes in your sample:



You reason that two of the types are the original double stranded bacteria and human DNA, and that the third was made when a strand of the human DNA base paired with a strand of bacteria DNA.

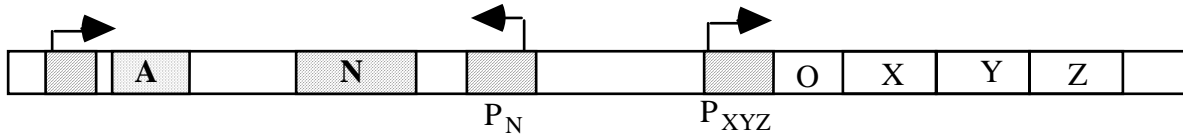
On the figure above, identify each complex. For the bacteria-human hybrid, indicate which strand is bacterial, and which is human.

Briefly justify your choices.

Human DNA for the same protein is likely to be significantly longer than a bacterial gene encoding homologous bacterial protein because the human gene is likely to have introns. Those are represented by the large loops in the human-bacterial hybrid in the figure above. Furthermore, because the protein sequence is not 100% identical between the human and bacterial versions, some portions of the human and bacterial genes are likely to not be similar enough to base pair (even with mismatches). Those regions are represented by the small bulges that occur on both the human and bacterial strands in the hybrid figure above.

Question 3

As a UROP student, you are working in bacteria on the metabolism of the monosaccharide, *theose*. Metabolism of *theose* requires the enzymes X, Y, and Z. The genes encoding these enzymes are part of one operon, and the product of gene N (N protein) regulates the transcription of these three genes. In a normal cell, protein A is always produced. A diagram of this operon is shown below.



a) Give a brief definition of an operon.

An operon is two or more genes that are regulated by the same control region(s).

b) In a cell where there is a high level of N protein, you detect no transcription of genes X, Y, and Z. You conclude that...

The N protein is (circle one)

a repressor.

an activator.

c) In further study, you discover that

- transcription of the N gene is controlled by protein A (protein A is the product of gene A shown in the diagram), and
- theose* binds to the N protein.

You examine the transcription of genes N, X, Y, and Z in cells where gene A is normal (A^+) and where gene A is not functional (A^-) and in the presence (+) and absence (-) of *theose*. The data is shown below.

gene A	<i>theose</i>	Transcription of gene N	Transcription of genes X, Y, and Z
A^-	+	-	+
A^+	-	+	-
A^-	-	-	+
A^+	+	+	+

i) Given the results above, what does protein A do?

Protein A promotes the transcription of gene N.

ii) Given the results above, what does *theose* do?

Theose binds to protein N and prevents it from binding to the O region.

This prevents repression by N.

Question 3, continued

d) You construct the following diploids containing two copies of the *theose* operon. Some of the genes or control regions are mutated where (-) = complete loss of function and (+) = normal. Predict whether you would see transcription of genes X, Y, and Z for each diploid below.

diploid	Transcription of X, Y, and Z?	
	- theose	+ theose
$\frac{P_A^+ A^+ N^+ P_N^+ P_{XYZ}^+ O^+}{P_A^+ A^+ N^+ P_N^+ P_{XYZ}^+ O^+}$	no	yes
$\frac{P_A^+ A^+ N^- P_N^+ P_{XYZ}^+ O^+}{P_A^+ A^- N^+ P_N^+ P_{XYZ}^+ O^+}$	no	yes
$\frac{P_A^+ A^+ N^+ P_N^- P_{XYZ}^+ O^+}{P_A^+ A^+ N^- P_N^+ P_{XYZ}^+ O^+}$	yes	yes
$\frac{P_A^+ A^+ N^+ P_N^+ P_{XYZ}^- O^+}{P_A^+ A^+ N^+ P_N^+ P_{XYZ}^+ O^-}$	yes	yes
$\frac{P_A^- A^+ N^+ P_N^+ P_{XYZ}^- O^-}{P_A^+ A^+ N^- P_N^+ P_{XYZ}^+ O^+}$	no	yes

The Genetic Code

	U	C	A	G	
U	UUU phe UUC phe UUA leu UUG leu	UCU ser UCC ser UCA ser UCG ser	UAU tyr UAC tyr UAA STOP UAG STOP	UGU cys UGC cys UGA STOP UGG trp	U C A G
C	CUU leu CUC leu CUA leu CUG leu	CCU pro CCC pro CCA pro CCG pro	CAU his CAC his CAA gln CAG gln	CGU arg CGC arg CGA arg CGG arg	U C A G
A	AUU ile AUC ile AUA ile AUG met	ACU thr ACC thr ACA thr ACG thr	AAU asn AAC asn AAA lys AAG lys	AGU ser AGC ser AGA arg AGG arg	U C A G
G	GUU val GUC val GUA val GUG val	GCU ala GCC ala GCA ala GCG ala	GAU asp GAC asp GAA glu GAG glu	GGU gly GGC gly GGA gly GGG gly	U C A G

Question 4

Recall (from problem set 1) that sickle-cell anemia is a disease that results from the presence of abnormal hemoglobin (HbS) in the red blood cells. In order to have the disease, a person needs to have only HbS hemoglobin. Recall also that a carrier of the disease is a person who has both HbS and HbA (wild-type hemoglobin) in their red blood cells.

Suppose a colleague at your lab created some human stem cell lines from adult human carriers of sickle cell disease. As it happened, she left a dish with some of these cells in a hood where you were performing your UV mutagenesis experiments on yeast.

Later she told you that, strangely, when she coaxed those cells to differentiate into red blood cells (RBCs) and placed these new RBCs into low O₂ environment, some cells assumed rigid sickle-like shapes.

- a) Do you think that appearance of the sickle-like cells is related to your colleague leaving the dish in the hood? If yes, how? If no, what caused this phenomenon?

UV light is a mutagen. Leaving the dish of stem cells in the hood exposed to UV light likely caused mutations in at least some of the individual cells. Because some of these cells developed into RBCs that behave as if they only have HbS hemoglobin, it is logical to conclude that the stem cells from which these particular RBCs came suffered a mutation in the copy of the gene for β -subunit that coded for the wild-type subunit. As a result, no HbA hemoglobin was produced in those cells, and all the hemoglobin produced was HbS.

- b) Is the DNA in the sickle-like cells different from that in the normal-shape cells? If so, how is it different? If not, explain why not.

Actually, RBCs do not have a nucleus, and therefore, have no DNA. But the stem cells that differentiated into the sickle-like cells had to only have genes encoding for the formation of HbS while the stem cells that differentiated into normal-shaped cells must have had genes encoding for both the HbS and HbA variants.

- c) Is the protein content of the sickle-like cells different from that in the normal-shape cells? If so, how is it different? If not, explain why not.

Yes, sickle-like cells must contain only HbS, while normal-shaped cells in general must contain HbA, and may contain some HbS. Because the stem cells in this experiment came from carriers of sickle-cell disease, the normal-shape RBCs must also contain some HbS.

- d) Are your answers to parts b and c related? If yes, how are they related? If no, explain why they are not related.

The stem cells in the dish all had the same versions of the genes. Because they came from carriers of sickle-cell anemia, they had one copy of a gene encoding wild-type β -subunit, leading to formation of HbA, and another that encoded a mutant β -subunit, leading to formation of HbS. If they were not exposed to a mutagen, they would likely all differentiate into RBCs that retained their normal shape even at low levels of O₂.

After being exposed to UV, some of the cells underwent mutations, and in some number of these cells the mutations occurred in the wild-type copy of the β -subunit. Those mutations altered the gene sequence, such that the mutated sequence now coded for a β -subunit that leads to the formation of HbS, instead of HbA. As a result, the RBCs differentiated from such a stem cell would contain only HbS, and would become deformed into the sickle-like shape if exposed to low levels of O₂.

The stem cells that did not undergo a mutation in the β -subunit gene would differentiate into RBCs that retained normal shape even at low levels of O_2 because they contain both HbS and HbA.