

2006 7.012 Problem Set 3 KEY

Due before 5 PM on FRIDAY, October 13, 2006.

Turn answers in to the box outside of 68-120.

PLEASE WRITE YOUR ANSWERS ON THIS PRINTOUT.

1. Which reaction is catalyzed by each of the enzymes listed below? Answer by stating which specific type of bond is affected (e.g. "covalent" is not specific enough), and whether each enzyme catalyzes the formation or the breaking of that type of bond.
- (a) helicase

Helicase disrupts Hydrogen bonds between bases in opposing strands of a DNA double helix, thereby unwinding the double helix into two single strands of DNA.

- (b) DNA polymerase

DNA polymerase catalyzes the addition of deoxyribonucleotides to a growing strand of DNA. This condensation reaction results in the formation of a phosphodiester linkage between the existing strand of DNA and the newly added deoxyribonucleotide.

- (c) DNA ligase

DNA ligase catalyzes the formation of phosphodiester linkages between two pre-existing fragments of DNA, thereby ligating (i.e. joining) the 2 fragments into one continuous strand.

- (d) RNA polymerase

RNA polymerase catalyzes the joining of two ribonucleotides into a growing strand of RNA. This condensation reaction results in the formation of a phosphodiester linkage between the two ribonucleotides.

- (e) ribosome

The ribosome catalyzes formation of a peptide bond between two amino acids to form a polypeptide chain

2. A portion of the DNA sequence and the corresponding predicted amino acid of a protein-encoding gene is given below.

```
5 '   ggg act cgg tgc tgt gat tgt cgg gct gct cct   3 '  
3 '   ccc tga gcc acg aca cta aca gcc cga cga gga   5 '
```

ala thr ile thr pro

Name: _____ KEY _____

You decide to make a DNA primer that would hybridize to the mRNA produced from this gene. What would the sequence of the primer be? Limit your answer to the 15 nucleotides that would hybridize in the boxed region. Give the sequence of the primer that would hybridize to the mRNA and label the primer's 5' and 3' ends.

a DNA primer that would hybridize to this sequence would be

3' GGC TGT TAG TGT CGT 5'

because the mRNA produced from this gene would be

5' ...CCG ACA AUC ACA GCA... 3'

The lower strand must be the strand that looks like the mRNA, because it is the only one of the two strands that would encode for the protein sequence listed. If you want your primer to hybridize to the mRNA, then your primer must be complementary and antiparallel to your mRNA.

3. The picture shown below is a sequence alignment of an entire gene's DNA sequence and the entire sequence of the mRNA produced from that gene. The top line is the DNA and the bottom line is the mRNA. Each nucleotide in the gene is numbered. Vertical dashes indicate nucleotides that are identical in both sequences. Dots indicate nucleotides in the DNA sequence that are not found in the mRNA sequence.

@ represents a 5'-G-cap.

Name: _____ KEY _____

(a) At what nucleotide position in the gene does transcription begin?

206 – the first point that we see a mRNA product complementary to the genomic sequence

(b) At what nucleotide position in the gene does transcription end?

675 – the last point that we see a mRNA product complementary to the genomic sequence

(c) At what nucleotide position in the gene does translation begin?

284 – the A in the first AUG in the mRNA

(d) At what nucleotide position in the gene does translation end?

595 or 598 – the last nucleotide that codes for an amino acid is 595 and 598 is the position of the G in the UAG (which is the first in-frame stop sequence in the mRNA)

(e) How many introns does the gene have?

One. There is only one segment internal in the gene that has been removed from the pre-mRNA to make the final, processed mRNA. An intron is always found between 2 exons.

(f) For each intron, give the nucleotide positions of its beginning **and** its end.

Start- 353, end- 464

(g) How many exons does the gene have?

Two. These two exons flank the intron that has already been removed from the mRNA.

(h) How many amino acids long would the protein be that is encoded by this gene?

67 amino acids – corresponding to 67 triplet codons from the start codon at 284 to the stop codon at 596-598. The stop codon is not included because it doesn't code for an amino acid.

(i) At what nucleotide positions does the 5' untranslated region begin **and** end?

206-283. This is the 5' part of the mRNA that is transcribed but not translated, so it is found at the beginning of the mRNA to the first AUG

(j) At what nucleotide positions does the 3' untranslated region begin **and** end?

It starts at 596 or 599 (right after the end of translation). This is the 3' part of the mRNA that is transcribed but not translated.

It ends at 675 or 694 (675 is the end of the part of the gene that is transcribed and 694 is the end of the poly-A tail)

Name: _____ KEY _____

4. The following sequence is a wild-type bacterial gene that encodes a short protein. The sequence given is from the point where transcription starts (called "+1") to the point where transcription ends.

5' -ACTTCGATATGTCTAAAATATCGATCGATCTGTGGGGCCTAGCTAGCTAACCAGAGACGCTACCG-3'
3' -TGAAGCTATACAGATTTTATAGCTAGCTAGACACCCCGGATCGATCGATTGGTCTCTGCGATGGC-5'

(a) Which strand (the upper or the lower) is used as the template in transcription?

Lower. The upper strand is the only strand that contains an AUG, so the upper strand must be the strand that looks like the mRNA. That means that the lower strand must be the strand that is used as a template.

(b) Write out the entire sequence of the RNA transcribed from this wild-type gene. Make sure to label the 5' and 3' ends of your molecule.

5' -ACUUCGAUAUGUCUAAAAUAUCGAUCGAUCUGUGGGGCCUAGCUAGCUAACCAGAGACGCUACCG-3'

Note that we gave you the sequence from the point where transcription starts to the point where transcription ends, so you need to include the entire sequence, which will look like the upper strand except that all of the Ts will instead be Us.

(c) Write out the amino acid sequence of any protein that is encoded by this wild-type gene. Make sure to label the N and C termini of your molecule.

N-met-ser-lys-ile-ser-ile-asp-leu-trp-gly-leu-ala-ser-C

You simply translate the mRNA from part (b) using the genetic code, from the first start codon to the first stop codon that is in frame after the start codon.

The following sequence is a mutant version of the above gene that is present in a mutant bacterial strain. The nature of the mutation is that **the base-pair bolded above in the wild-type sequence has been deleted**. The sequence given is from the point where transcription starts to the point where transcription ends.

5' -ACTTCGATATGTCTAAAATACGATCGATCTGTGGGGCCTAGCTAGCTAACCAGAGACGCTACCG-3'
3' -TGAAGCTATACAGATTTTATGCTAGCTAGACACCCCGGATCGATCGATTGGTCTCTGCGATGGC-5'

(d) Which strand (the upper or the lower) is used as the template in transcription?

lower – for every gene, which strand is used as a template is defined. For this gene, the lower strand will always be used as the template whether the allele of the gene is wild-type or mutant.

(e) Write out the amino acid sequence of any protein that is encoded by this mutated gene. Make sure to label the N and C termini of your molecule.

Name: _____ KEY _____

N-met-ser-lys-ile-arg-ser-ile-cys-gly-ala-C (bold= new amino acid sequence)

The deletion of a single basepair has caused a frameshift mutation that changes the amino acids from the point of the frameshift until a stop codon is reached in this new reading frame.

The following sequence is a wild-type gene that encodes a tRNA-ser molecule that recognizes the codon 5'-UCG-3' on all mRNAs in the bacterial cell. The sequence given is from the point where transcription starts (called "+1") to the point where transcription ends.

5' - CCGTTGCTCAGATCTGGATA **TCG** TCCATCCTGCATGCACTTGCTCATGCTGATACGCGCAACGGT - 3'
3' - GGCAACGAGTCTAGACCTAT **AGC** AGGTAGGACGTACGTGAACGAGTACGACTATGCGCGTTGCCA - 5'

(f) Which strand (the upper or the lower) is used as the template in transcription? (Remember that tRNAs are DIRECTLY transcribed from tRNA-encoding genes. There is no mRNA intermediate in the production of a tRNA molecule from a tRNA gene!)

upper – If the tRNA recognizes 5'-UCG-3', then the tRNA itself must have the anticodon 3'-AGC-5'. Thus the strand that looks like the tRNA will contain this sequence 3'-AGC-5'. The lower strand contains this sequence, so the upper strand must be the one that is used as a template.

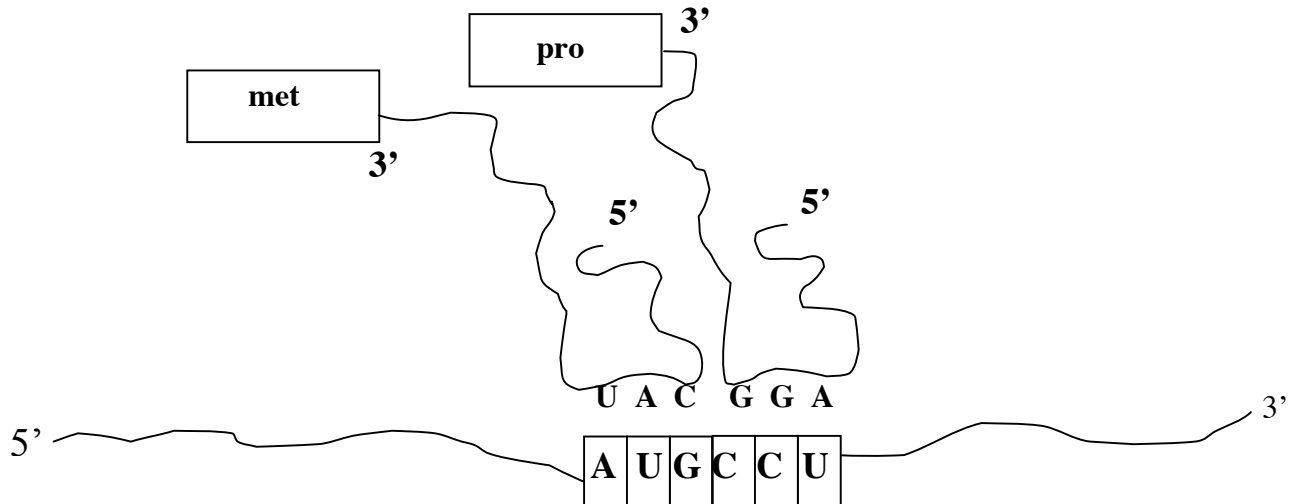
(g) Write out the amino acid sequence of any protein that is encoded by this wild-type gene. Make sure to label the N and C termini of your molecule.

None. Genes that encode tRNAs are transcribed but never translated. No protein is produced from a tRNA-encoding gene. The RNA transcript is the final product of this gene.

(h) Put a box around the double-stranded DNA portion of the wild-type tRNA gene that encodes the anticodon portion of the tRNA. (Do this in the drawing of the sequence that we provided for you at the top of the page.)

See diagram. If the tRNA recognizes 5'-UCG-3', then the tRNA itself must have the anticodon 3'-AGC-5'. Thus the strand that looks like the tRNA will contain this sequence 3'-AGC-5'. The lower strand contains the only copy of this sequence present in this gene.

5. Below is a diagram of two tRNAs and an mRNA in the active site of the ribosome during translation of the mRNA into protein. Three nucleotides from the sequence of each tRNA are shown for you.



(a) In the diagram above, label the 5' and 3' ends of each tRNA.

See diagram – there are four ends that must be labeled, two 5' ends and two 3' ends. The tRNA must be antiparallel to the mRNA that it is basepairing with, and the amino acid is always covalently attached to the 3' end of the tRNA molecule.

(b) In the diagram above, fill in the box attached to one end of each tRNA with the name of the amino acid that would be attached there.

See diagram – the one on the left would be attached to methionine and the one on the right would be attached to proline. This way this is determined is that you know that the mRNA must read 5'-A,U,G,C,C,U-3' (since it must be antiparallel and complementary to the tRNAs it is basepairing with), and 5'-AUG-3' encodes methionine and 5'-CCU-3' encodes proline, according to the genetic code.

(c) In the diagram above, fill in the boxes in the mRNA with the 6 nucleotides that would be present there.

See diagram – it should read 5'-A,U,G,C,C,U-3'. The sequence of the codons on the mRNA must be complementary and antiparallel to the sequence of the anticodons on the tRNA, since the two molecules are basepairing with each other.

(d) Which tRNA is about to transfer its attached amino acid over to the other tRNA and exit the active site: the tRNA on the left or the tRNA on the right?

mRNA is translated 5' to 3', and thus the tRNA on the left would be the first to leave. The ribosome would be reading the template 5'→3' and making the protein N to C.

Name: _____ **KEY** _____

(e) After the tRNA you mentioned in part (d) leaves, what will the sequence (so far) of the newly made protein be? Be sure to label the N and C termini of the growing polypeptide chain.

So far, the polypeptide chain would be N- met-pro -C. The ribosome would catalyze the reaction of forming a peptide bond between the met and the pro. Proteins are always made in the direction N to C, so met would be on the N terminal end of the growing polypeptide chain.

6. Tryptophan is one of the 20 amino acids used by all organisms to make their proteins. Bacteria can synthesize tryptophan from scratch, or they can directly use any tryptophan that is available to them from the environment. The genes that encode enzymes that synthesize tryptophan are transcriptionally regulated in a logical way that minimizes energy expenditure.

(a) Should these tryptophan synthesis genes be expressed when tryptophan:
... is present in the environment? ... is absent from the environment?

NO

YES

Bacteria only need to synthesize tryptophan when there is none available in the environment. If there is some available already, it makes a lot more sense in terms of energy consumption to just take it up from the environment, instead of synthesize it from scratch.

(b) A mutant bacterium has no activity for one of these tryptophan synthesis enzymes. Does this result prove that there is a mutation in the gene encoding this enzyme?

No. The lack of activity could result from a number of possibilities. A mutation in any gene that affects the activity of the tryptophan synthesis gene (i.e. regulates the tryptophan synthesis gene) could affect the activity of this gene. For instance, the mutation could have been in an activator protein that promotes transcription of the tryptophan synthesis gene such that the activator protein no longer could work, and thus the tryptophan synthesis gene could no longer be transcribed at all. This would eliminate activity of the tryptophan synthesis enzyme, even though the tryptophan synthesis gene itself would have been wild-type.

(c) One method of gene regulation is to control whether or not a gene is transcribed and translated. Another method is to control activity of a protein by whether or not it is phosphorylated. Which of these two methods of regulation allows for a faster response to changing cellular conditions?

It is much quicker to phosphorylate a protein than synthesize it from scratch. Synthesis from scratch involves catalyzing the formation of a bond between every nucleotide in the

Name: _____ KEY _____

mRNA that is being made (mRNAs are often 1000s of nucleotides long, so this would be 1000s of bonds) and then involves catalyzing the formation of a bond between every amino acid in the protein that is being made (proteins are often 100s of amino acids long, so this would be 100s of bonds). Putting a phosphate group on a protein only involves the catalysis of the formation of one bond, the bond between the R group of the amino acid and the third phosphate group of ATP.

(d) Which of these two methods of regulation requires less energy to express an active gene product?

It takes less energy to phosphorylate a protein than to synthesize it from scratch. Every time you synthesize an mRNA, every phosphodiester bond you form uses up one nucleotide tri-phosphate. Every time you synthesize a protein, every peptide bond you form costs multiple ATPs. Thus transcription and translation are very energy costly events. Putting a phosphate group onto a protein only requires a single ATP, i.e. the one whose phosphate group is torn off and then attached to the R group of the amino acid on the protein that is being phosphorylated.