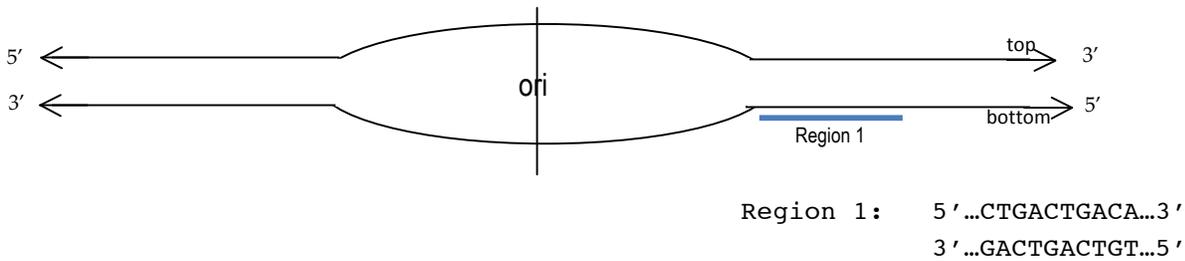


Solutions to Molecular Biology Unit Exam

Question 1

Consider the following origin of replication that is found on a chromosome. The sequence of region 1 is shown below.



a) Within Region 1, which strand will be the template for leading strand synthesis, the top or the bottom?
The bottom strand will be the template for leading strand synthesis.

b) If we assume that a lagging strand fragment is made from region 1, what will be its sequence?

5' T G T C A G T C A G 3'

c) You examine DNA replication in an *E. coli* mutant, which has a partially defective DNA polymerase. In vitro experiments using the mutant DNA polymerase gives an error rate of 10^{-3} , as compared to the expected error rate of 10^{-6} . Which of the following activities is the mutant polymerase likely to be missing, as compared to the normal polymerase? Circle all that apply.

5'→3' polymerase

3'→5' exonuclease

5'→3' exonuclease

3'→5' polymerase

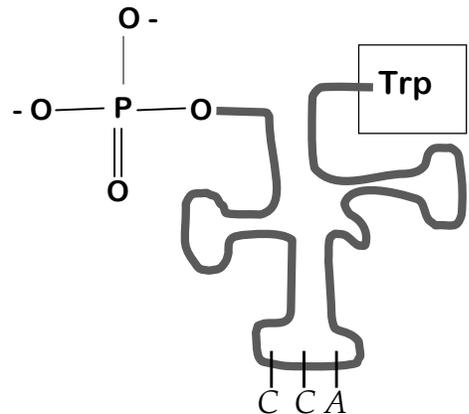
5'→3' recombinase

3'→5' recombinase

d) Below is a schematic of the molecule that inserts the fourth amino acid (a tryptophan) into the mutant polymerase. A codon chart is found on the final page of the exam.

i) This schematic represents a tRNA

ii) On the schematic, give the nucleotides of the anticodon.



Question 2

Below are 210 consecutive base pairs of DNA that includes only the beginning of the sequence of gene X. The underlined sequence (from position 20-54) represents the promoter for gene X and the underlined and italicized sequence (from position 71-90) encodes the gene X ribosome binding (RBS) site. Transcription begins at and includes the T/A base pair at position 60 (bold)

```

      1          10          20          30          40          50          60          70
      I-----I-----I-----I-----I-----I-----I-----I-----I
5' ATCGGTCTCGGCTACTACATAAAACGCGCGCATATATCGATATCTAGCTAGCTATCGGTCTAGGCTACTAC
3' TAGCCAGAGCCGATGATGTATTTTGCGCGGTATATAGCTATAGATCGATCGATAGCCAGATCCGATGATG
                                Promoter

      80          90          100         110         120         130         140
      -----I-----I-----I-----I-----I-----I-----I-----I-----I
5' CAGGTATCGGTCTGATCTAGCTAGCTTCTCTTCTCTCTCTCCCCGCGGGGGCTGTACTATCATGCGTCG
3' GTCCATAGCCAGACTAGATCGATCGAAGAGAAGAGAGAGAGGGGGCGCCCCGACATGATAGTACGCAGC
                                RBS

      150         160         170         180         190         200         210
      -----I-----I-----I-----I-----I-----I-----I-----I-----I
5' TCTCGGCTACTACGTAAACGCGCGCATATATCGATATCTAGCTAGCTATCGGTCTCGGCTACTACGTAAA
3' AGAGCCGATGATGCATTTGCGCGCGTATATAGCTATAGATCGATCGATAGCCAGAGCCGATGATGCATTT

```

a) What are the first 6 nucleotides of the mRNA from gene X?

5' U A G G C U 3'

b) What are the first 4 amino acids encoded by gene X? (A codon chart is found on the final page)

N met arg arg leu C

You have found two different mutations of gene X, mutation 1 and mutation 2.

c) In mutation 1, there is an insertion of the following three base pairs immediately after the C/G base pair at position 100 (shown in bold).

```

5' TGT 3'
3' ACA 5'

```

i) Would the mRNA expressed from this version of gene X be longer, shorter, or the same as that produced from the normal gene X? Explain and if longer or shorter, indicate by how many in bases.
The insertion is after the promoter so the mRNA would be longer by 3 nucleotides.

ii) If the mRNA can be translated, ...

- ...would you expect the protein to be longer, shorter, or the same as that produced from the normal gene X? If longer or shorter, indicate by how many in amino acids.
The protein would be the same as the normal protein because the insertion is before the start codon.
- ...do you expect that the protein produced will have the same function as the normal protein X? Explain your thinking.
The protein produced is identical to the normal protein, so it will have the same structure and function.

Question 2, continued

This is the same sequence as shown on the previous page. It is repeated for your convenience.

Below are 210 consecutive base pairs of DNA that includes only the beginning of the sequence of gene X. The underlined sequence (from position 20-54) represents the promoter for gene X and the underlined and italicized sequence (from position 71-90) encodes the gene X ribosome binding (RBS) site. Transcription begins at and includes the T/A base pair at position 60 (bold)

```

      1          10          20          30          40          50          60          70
      I-----I-----I-----I-----I-----I-----I-----I-----I
5' ATCGGTCTCGGCTACTACATAAAACGCGCGCATATATCGATATCTAGCTAGCTATCGGTCTAGGCTACTAC
3' TAGCCAGAGCCGATGATGTATTTTGCGCGGTATATAGCTATAGATCGATCGATAGCCAGATCCGATGATG
                                Promoter

      80          90          100         110         120         130         140
      -----I-----I-----I-----I-----I-----I-----I-----I
5' CAGGTATCGGTCTGATCTAGCTAGCTTCTCTTCTCTCTCTCCCCGCGGGGGCTGTACTATCATGCGTCTG
3' GTCCATAGCCAGACTAGATCGATCGAAGAGAAGAGAGAGAGGGGGCGCCCCGACATGATAGTACGCAGC
      RBS

      150         160         170         180         190         200         210
      -----I-----I-----I-----I-----I-----I-----I-----I
5' TCTCGGCTACTACGTAAACGCGCGCATATATCGATATCTAGCTAGCTATCGGTCTCGGCTACTACGTAAA
3' AGAGCCGATGATGCATTTGCGCGGTATATAGCTATAGATCGATCGATAGCCAGAGCCGATGATGCATTT
  
```

d) In mutation 2, there is an insertion of the following four base pairs immediately after the A/T base pair at position 130 (shown in bold).

```

5' ATGT 3'
3' TACA 5'
  
```

i) Would the mRNA expressed from this version of gene X be longer, shorter, or the same as that produced from the normal gene X? Explain and if longer or shorter, indicate by how many in bases. *The insertion is after the promoter so the mRNA would be longer by 4 nucleotides.*

ii) If the mRNA can be translated,...

- ...what are the first four amino acids produced?
N met phe met arg C

- ...would you expect the protein to be longer, shorter, or the same as that produced from the normal gene X? If longer or shorter, indicate by how many in amino acids.
The protein encoded by mutant 2 will be longer by 2 amino acids.

- ...do you expect that the protein produced will have the same function as the normal protein X? Explain your thinking.
This mutant protein is nearly identical to the normal protein. It is different in that it has two additional amino acids at the very N terminus, but these two terminal amino acids are unlikely to change the overall tertiary structure of the protein, so it is unlikely that this insertion will change the function of the protein.

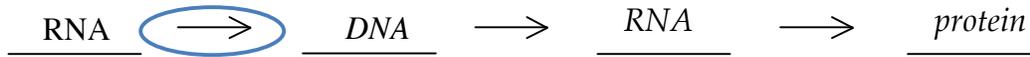
Question 3

You have discovered a new virus that contains only RNA as its genetic material. Curious as to how this virus works, you infect host cells with this virus and discover that after infection, you find that the host cell makes lots of viral RNA molecules and various viral proteins. Upon further examination, you discover that the host genome now encodes viral proteins.

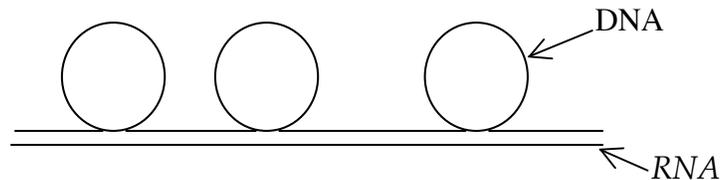
a) What type of virus you have discovered.

A Retrovirus.

b) Outline the steps of the information flow for this virus by filling in the blanks below. Each arrow represents a process, for example, copying of DNA from a DNA template. Circle the arrow that represents a process not found naturally in the host cell.



c) There is a human gene that encodes a protein identical to one of the proteins produced by the virus. You isolate a fragment of DNA that includes the shared gene, heat the fragment to separate the two DNA strands and allow the human DNA to base pair with the viral RNA. You find the following hybrid molecule using electron microscopy.



i) Label which strand is human DNA and which is viral RNA in the picture.

ii) Why can some regions form complementary base pairs, while other regions cannot?

The DNA contains introns that are not present in the viral RNA.

iii) If you were to isolate mature mRNA from the human cell and allow it to base pair with the viral RNA, would you see the same type of hybrid molecule? Explain your thinking.

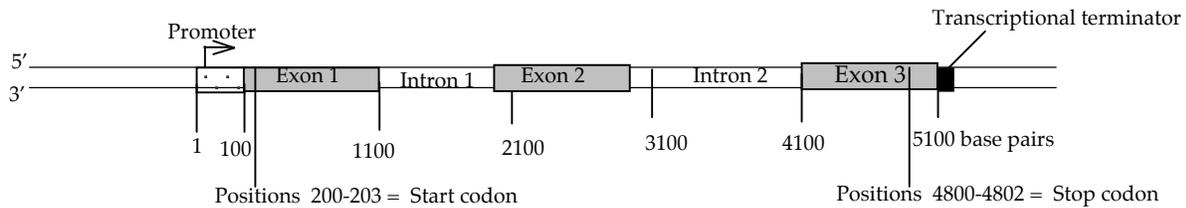
The mature human mRNA results from the processing of the original transcript. During processing, the introns are spliced out. Because the two proteins are identical, you would expect that the two molecules would be the same, so you would not see the type of hybrid molecule shown above.

d) In an actively dividing, normal cell, indicate whether the following statements are true or false for the virus discussed above.

- | | | |
|-------------------------------|--------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| True | <input type="checkbox"/> False | RNA-dependent RNA polymerase is used at some point during the viral life cycle to copy the viral genome. |
| <input type="checkbox"/> True | False | RNA-dependent DNA polymerase is used at some point during the viral life cycle to copy the viral genome. |
| <input type="checkbox"/> True | False | DNA-dependent DNA polymerase is used at some point during the viral life cycle to copy the viral genome. |
| <input type="checkbox"/> True | False | The polymerase used by the virus to copy its genome forms a covalent bond between a 5' phosphate and a 3' hydroxyl. |
| True | <input type="checkbox"/> False | The polymerase used by the virus to copy its genome forms a covalent bond between an N-terminal amine and a C-terminal carboxyl. |
| <input type="checkbox"/> True | False | Polymerization of the viral genome would be 5' → 3'. |
| True | <input type="checkbox"/> False | Polymerization of the viral genome would be 3' → 5'. |
| True | <input type="checkbox"/> False | Deoxyribonucleotides would be incorporated into the packaged viral genome. |
| True | <input type="checkbox"/> False | Di-deoxyribonucleotides would be incorporated into the packaged viral genome. |

Question 4

a) Below is a schematic of gene Y, which encodes protein Y. The promoter region is indicated by the dotted box. Transcription begins immediately following the promoter.



- The transcript first produced by this gene would be approximately how many nucleotides long?
The transcript first produced by this gene would be approximately 5000 nucleotides long.
- Two different transcripts are produced from this gene, one is approximately 2000 nucleotides long, the other is approximately 3000 nucleotides long. Explain how two different transcripts can be produced from this gene.
Alternative splicing could produce two different transcripts. One transcript could contain exons 1, 2 and 3, the other could contain only exons 1 and 3.
- Assume each transcript produces a protein. Given your answer above and the diagram, what is the approximate size of the two proteins produced from this gene?
The transcript contain all three exons would encode a protein beginning at the ATG at position 200 and end at the stop codon. This is about 2600 nucleotides, which represents about 866 codons. SO the two proteins expected would be about 866 amino acids long and about 533 amino acids long.

b) Assume gene Y was mutated such that the base pair found at position 200 was changed from an A/T to a G/C. Transcription and translation of the altered gene Y still occurs.

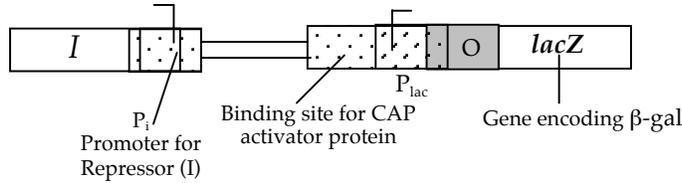
- Would the first transcript be the same length, shorter or longer than the first transcript produced from the wild type gene Y? Give all possible answers and explain your thinking.
This mutation changes a single base pair well after the promoter, so it is unlikely to affect transcription. The first transcript is expected to be the same length.
- Would the protein produced be the same length, shorter or longer than the protein produced from the wild type gene Y? Give all possible answers and explain your thinking.
This mutation changes a single base pair, eliminating the start codon. It is impossible to tell what affect this will have on the protein. If there is another start codon that maintains the original reading frame, then the protein will be shorter. If there is another start codon that changes the reading frame, then the protein will certainly be different. Because the location of the stop codon is also altered in this alternative reading frame, the resulting protein might be longer or shorter or even the same.

c) Now assume gene Y was mutated such that only the base pair found at position 2100 was changed from a C/G to an A/T. Transcription and translation of gene Y still occur.

- Would the first transcript be the same length, shorter or longer than the first transcript produced from the wild type gene Y? Give all possible answers and explain your thinking.
This mutation changes a single base pair well after the promoter, so it is unlikely to affect transcription. The first transcript is expected to be the same length.
- Would the protein produced be the same length, shorter or longer than the protein produced from the wild type gene Y? Give all possible answers and explain your thinking.
This mutation changes a single base pair of a single codon. There are two possibilities. If the new codon still encodes an amino acid, then the resulting protein will be the same length. If the new codon creates a stop codon, the resulting protein will be the shorter.

Question 5

You design a summer class where you recreate experiments studying the lac operon in *E. coli* (see schematic below). In your experiments, the activity of the enzyme β -galactosidase (β -gal) is measured by including X-gal and IPTG in the growth media. X-gal is a lactose analog that turns blue when metabolized by β -gal, but it does not induce the lac operon. IPTG is another analog of lactose that is an inducer of the lac operon but is not metabolized by β -gal.



a) Which of the following would you expect to bind to β -galactosidase? Circle all that apply.

Lactose (or allolactose)

X-gal

IPTG

b) Which of the following would you expect to bind to the lac repressor? Circle all that apply.

Lactose (or allolactose)

X-gal

IPTG

After mutagenesis you find 7 mutants that never turn blue as shown in the table below. Each mutant has a single loss-of-function mutation.

| Cell Type | Media | | | |
|-------------|---------------------------------|--------------------------------|----------------------------------|---------------------------------|
| | + glucose No lactose + X-gal | + glucose + lactose + X-gal | No glucose No lactose + X-gal | No glucose + lactose + X-gal |
| Wild type | White colonies | White colonies | White colonies | Dark blue colonies |
| Mutants 1-7 | White colonies | White colonies | White colonies | White colonies |

c) A single loss-of-function mutation in which component or components could produce the phenotype seen in these mutants? List all that apply. Choose from:

I, P_i, CAP binding site, O, lacZ, or P_{lac}

CAP binding site, P_{lac}, lacZ

You also find three mutants with the following phenotype. Each mutant has a single loss-of-function mutation.

| Cell Type | Media | | | |
|-----------|---------------------------------|--------------------------------|----------------------------------|---------------------------------|
| | + glucose No lactose + X-gal | + glucose + lactose + X-gal | No glucose No lactose + X-gal | No glucose + lactose + X-gal |
| Wild type | White colonies | White colonies | White colonies | Dark blue colonies |
| Mutants | White colonies | White colonies | Dark blue colonies | Dark blue colonies |

d) A single loss-of-function mutation in which component or components could produce the phenotype seen in these mutants? List all that apply. Choose from:

I, P_i, CAP binding site, O, lacZ, or P_{lac}

I, P_i, O

MIT OpenCourseWare
<http://ocw.mit.edu>

7.01SC Fundamentals of Biology
Fall 2011

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.