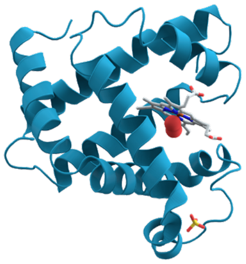
# The 1st Week Question List of *Molecular Biology*



*2024-pre-iGEMer@OUC-Haide*

Assignments: 'Short Course' & Question List

Due: Monday, 17 July 2023, 11:59 PM

| 7-10 | Protein Composition and Structure<br>Basic Concepts of Enzymes |
|---|---|
| Short Course | ☐ Chapter 3 Amino Acids<br>☐ Chapter 4 Protein Three-Dimensional Structure<br>☐ Chapter 5 Techniques in Protein Biochemistry<br>☐ Chapter 6 Basic Concepts of Enzyme Action |

## Instruction

- Answers in any form will be acceptable unless a specific rule is required in the question. Try your best to make your answers clear, solid and logical. Additional figures, screenshots and photos of your handwriting are recommended to do this.
- Submit your assignment tasks via email and we will give you constructive feedback. Student emails like sunjunteng@stu.ouc.edu.cn are the best because we can immediately know who you are.
- File naming rules: Week[number]_[Your serial number]_[Your Chinese name]. E.g., Week1_0_孙俊腾.doc/pdf
- Questions and answers are ALL in English.
- Looking forward to your wonderful answer. 😀

# Part I. General Question

---

*This part is for reviewing your involvement in the 'Short Course' and your understanding of molecular biology. Try to emphasize the necessary details.*

1. What three amino acids have aromatic components in their side chains? Please write their full names 1-letter abbreviates, and 3-letter abbreviates.
2. List some of the differences between an alpha helix and a beta strand.
3. Many of the loops on the proteins are composed of hydrophilic amino acids. Why?
4. What are the common types of weak bonds important in biochemistry? How does water affect these bonds? And briefly describe the hydrophobic effect.
5. Distinguish among ion-exchange chromatography, hydrophobic chromatography, gel-filtration chromatography, and affinity chromatography in terms of the column material and the basis for the separation of a mixture of proteins.

# Part II. Bonus

*This part is for testing your capacity for analysis ability, data handling, and modelling in molecular biology. Try to get the bonus and show off your talent.*

1. Like α helices, β sheets often have one side facing the surface of the protein and one side facing the interior, giving rise to an amphiphilic sheet with one hydrophobic surface and one hydrophilic surface. From the sequences listed below, pick the one that could form a strand in an amphiphilic β sheet. Think about the way side chains are arranged in a strand of a β sheet.

a) A L S C D V E T Y W L I

b) D K L V T S I A R E F M

c) D S E T K N A V F L I L

d) T L N I S F Q M E L D V

e) V L E F M D I A S V L D

2. Alpha Helices are often embedded in a protein so that one side faces the surface and the other side faces the interior. Such helices are often termed amphiphilic because the surface side is hydrophilic and the interior side is hydrophobic. A simple way to decide whether a sequence of amino acids might form an amphiphilic helix is to arrange the amino acids around what is known as a “helix-wheel projection” (**Figure 1**). The helix is amphiphilic if the hydrophobic and hydrophilic amino acids are segregated on opposite sides of the wheel. Using the helix-wheel projection, decide which of the three peptides in **Figure 1** might form an amphiphilic helix. (The mnemonic “`FAMILY VW`” will help you recognize hydrophobic amino acids, which is a convenient way to remember hydrophobic amino acids.)

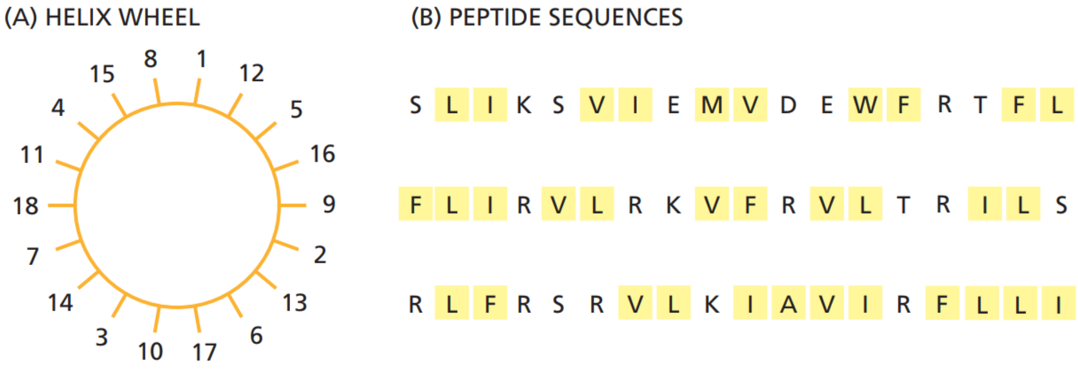
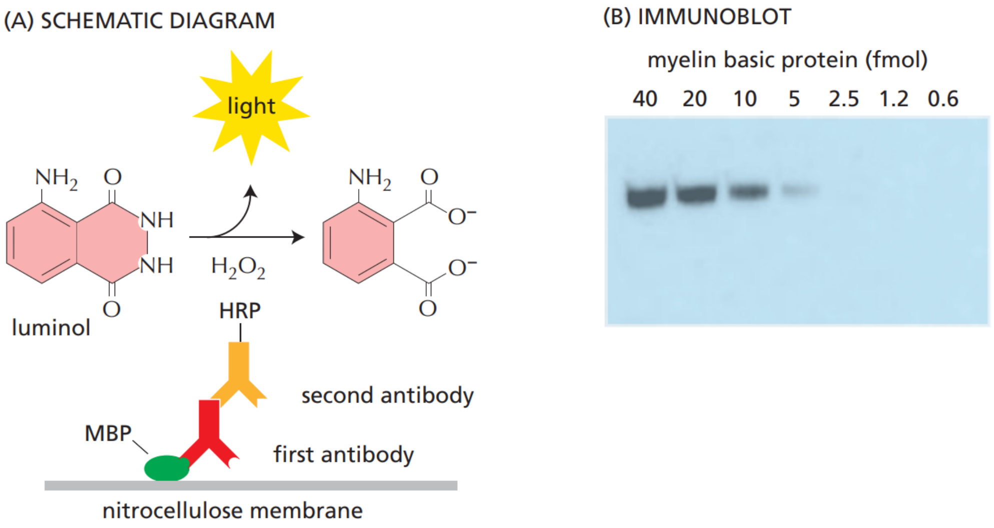


**Figure 1** Helix-wheel projection. (A) Helix wheel. The circle (wheel) represents the helix as viewed from one end. Numbers show the positions of the amino acid side chains, as projected on the wheel. The positions of the first 18 amino acids are shown; amino acid 19 would occupy the same position as amino acid 1. Amino acid 1 is closest to the reader; amino acid 18 is farthest away. (B) Peptide sequences. The N-termini are shown at the left; hydrophobic amino acids are highlighted in *yellow*; hydrophilic amino acids are *unmarked*.

3. You want to know the sensitivity for the detection of immunoblotting (Western blotting), using an enzyme-linked second antibody to detect the antibody directed against your protein (**Figure 2A**). You are using the mouse monoclonal antibody 4G10, which is specific for phosphotyrosine residues, to detect phosphorylated proteins. You first phosphorylate the myelin basic protein *in vitro* using a tyrosine-protein kinase that adds one phosphate per molecule. You then prepare a dilution series of the phosphorylated protein and subject the samples to SDS-PAGE. Next, the protein is transferred (blotted) onto a nitrocellulose filter, incubated with the 4G10 antibody, and washed to remove the unbound antibody. The blot is then incubated with a second goat anti-mouse antibody that carries horseradish peroxidase (HRP) conjugated to it, and any excess unbound antibody is again washed away. You place the blot in a thin plastic bag, add reagents that chemiluminescence when they react with HRP (**Figure 2A**), and place the bag against a sheet of x-ray film. When the film is developed, you see the picture shown in **Figure 2B**.

   a) Given the amounts of phosphorylated myelin basic protein indicated in each lane in Figure 2B, calculate the detection limit of this method in terms of molecules of protein per band.

   b) Assuming that you were using monoclonal antibodies to detect proteins, would you expect that the detection limit would depend on the molecular mass of the protein? Why or why not?


(A) SCHEMATIC DIAGRAM

(B) IMMUNOBLOT

**Figure 2** Sensitivity of detection of immunoblotting. (A) Schematic diagram of the experiment.

MBP stands for myelin basic protein. In the presence of hydrogen peroxide, horseradish peroxidase (HRP) converts luminol to a chemiluminescent molecule that emits light, which is detected by exposure to an X-ray film. (B) Exposed film of an immunoblot. The number of femtomoles of myelin basic protein in each band is indicated.

4. Phosphoglucose isomerase catalyzes the interconversion of glucose 6-phosphate (G6P) and fructose 6-phosphate (F6P):

$$\mathrm{G6P} \rightleftharpoons \mathrm{F6P}$$

The ΔG for this reaction is given by the equation

$$\Delta G = \Delta G^{\circ} + 2.3RT\log(\frac{[\mathrm{F6P}]}{[\mathrm{G6P}]})$$

where R = $8.3 \times 10^{-3}$ kJ/ K mol and T = 310 K. A useful number to remember is that 2.3 RT = 5.9 kJ/mol at 37°C, which is body temperature.

a) At equilibrium, [F6P]/[G6P] is equal to the equilibrium constant (K) for the reaction. Rewrite the above equation for the reaction at equilibrium.

b) At equilibrium, the ratio of [F6P] to [G6P] is observed to be 0.5. At this equilibrium ratio, what are the values of ΔG and ΔG°?

c) Inside a cell, the value of ΔG for this reaction is –2.5 kJ/mole. What is the ratio of [F6P] to [G6P]? What is ΔG°?

5. The polymerization of subunits into a pentameric ring is shown in **Figure 3**. The equilibrium constants for the association of a subunit at each step in the assembly of the tetramer (that is, $K_1$, $K_2$, and $K_3$) are approximately equal at $10^6\ \mathrm{M}^{-1}$. The equilibrium constant for association of the final subunit in the ring ($K_4$), however, is > $10^{12}\ \mathrm{M}^{-1}$. Why is an association of the final subunit so much more highly favoured than association of the initial subunits? Why do you suppose the equilibrium constant for the association of the final subunit is approximately the square of the equilibrium constants for the earlier steps?

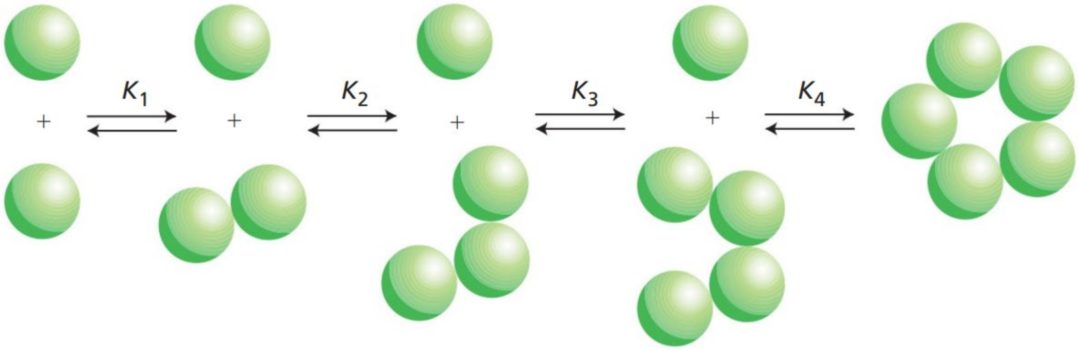


**Figure 3** Polymerization of subunits into a pentameric ring.

# The 2nd Week Question List of *Molecular Biology*

*2024-pre-iGEMer@OUC-Haide*

**Assignments: 'Short Course', Review & Question List**

Due: Monday, 24 July 2023, 11:59 PM

| 7-17 | Kinetics of Enzymes<br>Nucleic Acid Structure and DNA Replication |
|---|---|
| Short Course | ☐ Chapter 7 Kinetics and Regulation<br>☐ Chapter 33 The Structure of Informational Macromolecules: DNA and RNA<br>☐ Chapter 34 DNA Replication |
| A Review of Golden Gate Assembly | A User's Guide to Golden Gate Cloning Methods and Standards, 2022, ACS Synthetic Biology |

## Instruction

---

- Answers in any form will be acceptable unless a specific rule is required in the question. Try your best to make your answers clear, solid and logical. Additional figures, screenshots and photos of your handwriting are recommended to do this.
- Submit your assignment tasks via email and we will give you constructive feedback. Student emails like sunjunteng@stu.ouc.edu.cn are the best because we can immediately know who you are.
- File naming rules: Week[number]_[Your serial number]_[Your Chinese name]. E.g., Week1_0_孙俊腾.doc/pdf
- Questions and answers are ALL in English.
- Looking forward to your wonderful answer. 😀

# Part I. General Question

*This part is for reviewing your involvement in the 'Short Course' and your understanding of molecular biology. Try to emphasize the necessary details.*

1. Briefly describe 1) what is molecular cloning? 2) the principle of Golden Gate Cloning.
2. Why is $k_{cat}/K_m$ a measure of catalytic efficiency? And explain what is allosteric effect.
3. Explain: 1) why histones are positively charged whereas DNA is negatively charged? 2) why DNA replication is semicontinuous? 3) why do eukaryotes need multiple origins?
4. The nucleotide sequence of one DNA strand of a DNA double helix is

   `5′-GGATTTTTGTCCACAATCA-3′`

   What is the sequence of the complementary strand?
5. List major proteins (>5) involved in the synthesis of lagging strands in bacteria and briefly describe their functions.
6. SSB proteins bind to single-strand DNA at the replication fork and prevent the formation of short hairpin helices that would otherwise impede DNA synthesis. What sorts of sequences in single-strand DNA might be able to form a hairpin helix? Write out an example of a sequence that could form a five-nucleotide hairpin helix, and show the helix.
7. The DNA fragment in **Figure 1** is double-stranded at each end but single-stranded in the middle. The polarity of the top strand is indicated. Is the phosphate ($PO_4^-$) shown on the bottom strand at the 5′ end or the 3′ end of the fragment to which it is attached?

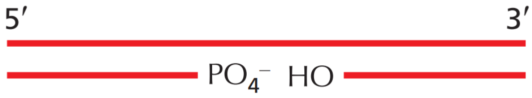


**Figure 1** A DNA fragment with a single-stranded gap on the bottom strand.

# Part II. Bonus

*This part is for testing your capacity for analysis, data handling, and modelling in molecular biology. Try to get the bonus and show off your talent.*

1. Examine Figure 2, which compares the energetics of a catalyzed and uncatalyzed reaction during the progress of the reaction from the substrate (S) to product (P). The highest peak in such a diagram corresponds to the transition state, which is an unstable, high-energy arrangement of substrate atoms that is intermediate between substrate and product. The free energy required to surmount this barrier to the reaction is termed the activation energy. Enzymes function by lowering the activation energy, thereby allowing a more rapid approach to equilibrium.

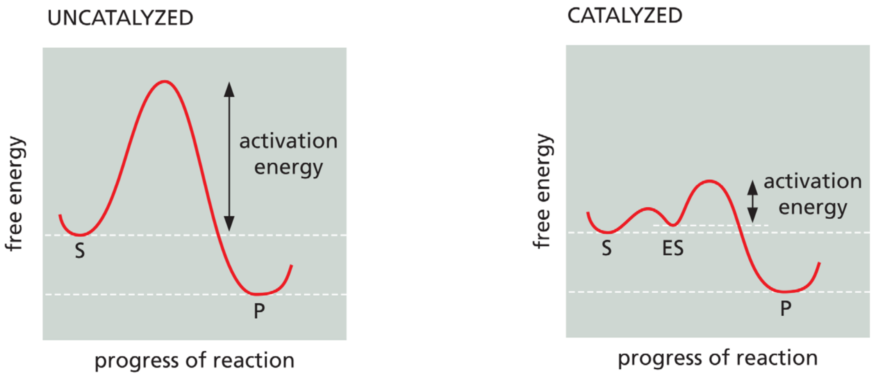


**Figure 2** Catalyzed and uncatalyzed reactions show the free energy at various stages in the progress of the reaction.

With this diagram in mind, consider the following question. Suppose the enzyme in the diagram was mutated so that its affinity for the substrate was increased by a factor of 100. Assume that there was no other effect beyond increasing the depth of the trough labelled ES (enzyme–substrate complex) in Figure 2. Would you expect the rate of the reaction catalyzed by the altered enzyme to be faster, slower, or equal to the reaction rate catalyzed by the normal enzyme?

2. The Michaelis constant, $K_\mathrm{m}$, is often spoken of as if it were a measure of the affinity of the enzyme for the substrate: the lower the $K_\mathrm{m}$, the higher the binding affinity. This would be true if $K_m$, were the same as $K_\mathrm{d}$ (the equilibrium constant for the dissociation reaction), but it is not. For an enzyme-catalyzed reaction

$$\mathrm{E + S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \overset{k_\mathrm{cat}}{\rightarrow} \mathrm{E + P}$$

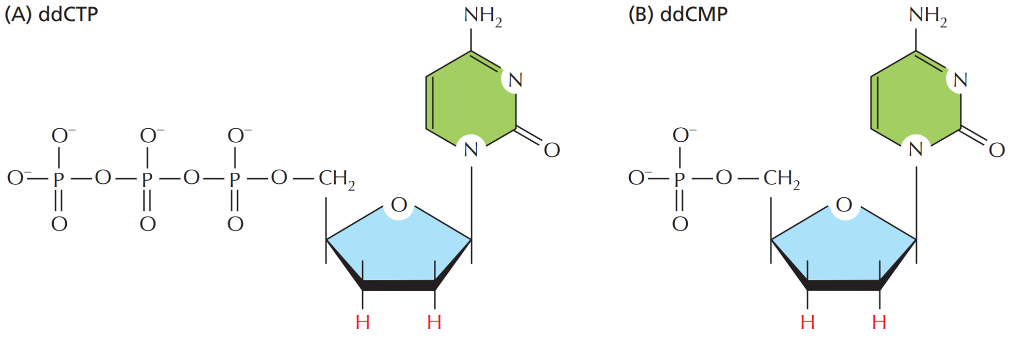
$$K_\mathrm{m} = \frac{k_{-1} + k_\mathrm{cat}}{k_1}$$

   a) In terms of these rate constants, what is $K_\mathrm{d}$ for dissociation of the ES complex to E + S?

   b) Under what conditions is $K_\mathrm{m}$ approximately equal to $K_\mathrm{d}$?

   c) Does $K_\mathrm{m}$ consistently overestimate or underestimate the binding affinity? Or does it sometimes overestimate and sometimes underestimate the binding affinity?

3. Look carefully at the structures of the molecules in Figure 3.


(A) ddCTP

(B) ddCMP

**Figure 3** Potential replication substrates. (A) Dideoxycytidine triphosphate (ddCTP). (B) Dideoxycytidine monophosphate (ddCMP).

a) What would you expect to happen if dideoxycytidine triphosphate (ddCTP) were added to a DNA replication reaction in large excess over the concentration of deoxycytidine triphosphate (dCTP)? Would it be incorporated into the DNA? If it were, what would happen after that? Give your reasoning.

b) What would happen if ddCTP were added at 10% of the concentration of dCTP?

c) What effects would you expect if dideoxycytidine monophosphate (ddCMP) were added to a DNA replication reaction in large excess, or at 10% of the concentration of dCTP?

4. The laboratory you joined is studying the life cycle of an animal virus that uses circular, double-strand DNA as its genome. Your project is to define the location of the origin(s) of replication and to determine whether replication proceeds in one or both directions away from an origin (unidirectional or bidirectional replication). To accomplish your goal, you isolated replicating molecules, cleaved them with a restriction nuclease that cuts the viral genome at only one site to produce a linear molecule from the circle, and examined the resulting molecules in the electron microscope. Some of the molecules you observed are illustrated schematically in Figure 4. *Note that it is impossible to distinguish the orientation of one DNA molecule from another in the electron microscope.*

   You must present your conclusions to the rest of the lab tomorrow. How will you answer the two questions your advisor posed for you?

   a) Is there a single, unique origin of replication or several origins?

   b) Is replication unidirectional or bidirectional?

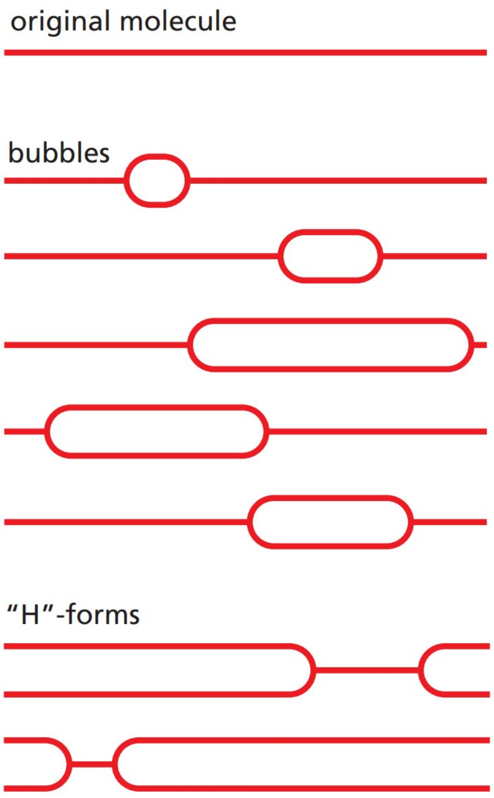
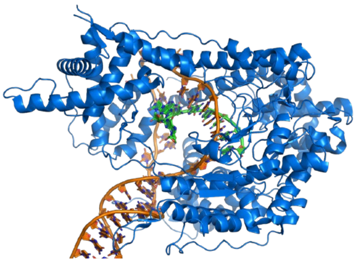


Figure 4 Parental and replicating forms of an animal virus.

# The 3rd Week Question List of *Molecular Biology*



*2024-pre-iGEMer@OUC-Haide*

**Assignments: ‘Short Course’, Review & Question List**

Due: Monday, 31 July 2023, 11:59 PM

| 7-24 | RNA Synthesis and Regulation<br>Recombinant DNA Techniques |
|---|---|
| Short Course | ☐ Chapter 36 RNA Synthesis and Regulation in Bacteria<br>☐ Chapter 37 Gene Expression in Eukaryotes<br>☐ Chapter 41 Recombinant DNA Techniques |
| A Review of Aptamer Screening: SELEX | SELEX—A (r)evolutionary method to generate high-affinity nucleic acid ligands, 2007, Molecular Engineering |

## Instruction

- Answers in any form will be acceptable unless a specific rule is required in the question. Try your best to make your answers clear, solid and logical. Additional figures, screenshots and photos of your handwriting are recommended to do this.
- Submit your assignment tasks via email and we will give you constructive feedback. Student emails like sunjunteng@stu.ouc.edu.cn are the best because we can immediately know who you are.
- File naming rules: Week[number]_[Your serial number]_[Your Chinese name]. E.g., Week1_0_孙俊腾.doc/pdf
- Questions and answers are ALL in English.
- Looking forward to your wonderful answer. 😀

# Part I. General Question

---

*This part is for reviewing your involvement in the 'Short Course' and your understanding of molecular biology. Try to emphasize the necessary details.*

1. Briefly answer: a) what is aptamer? b) the principle of SELEX?
2. Explain: a) what is the function of the σ subunit in prokaryotic transcription? b) what is the function of TBP in eukaryotic transcription? c) the difference between the transcription initiation in prokaryotes and eukaryotes.
3. Which of the following techniques would serve best as the basis for a rapid, highly sensitive blood test to detect circulating cancer cells in the bloodstream that carry the *Bcr-Abl* fusion gene?
   a) DNA sequencing
   b) Flow cytometry
   c) PCR analysis
   d) Western blotting
4. Explain the difference between the coding strand and the template strand in DNA. An RNA polymerase is transcribing a segment of DNA that contains the sequence

   ```
   5′-GTAACGGATG-3′
   3′-CATTGCCTAC-5′
   ```

   If the polymerase transcribes this sequence from left to right, what will the sequence of the RNA be? What will the RNA sequence be if the polymerase moves right to left?
5. To start with, briefly describe the negative regulation in the *lac* operon. Then check the **Figure 1**, the bacterial activator protein CAP and the *Lac* repressor have been placed in the four possible combinations on their binding sites in the promoter for the *lac* operon. Each combination of transcription regulators corresponds to the expected binding in a particular mixture of glucose and lactose. For each of the four combinations, indicate on the left-hand side of the figure which sugars are present and, on the right-hand side, whether the operon is expected to be turned on or off.

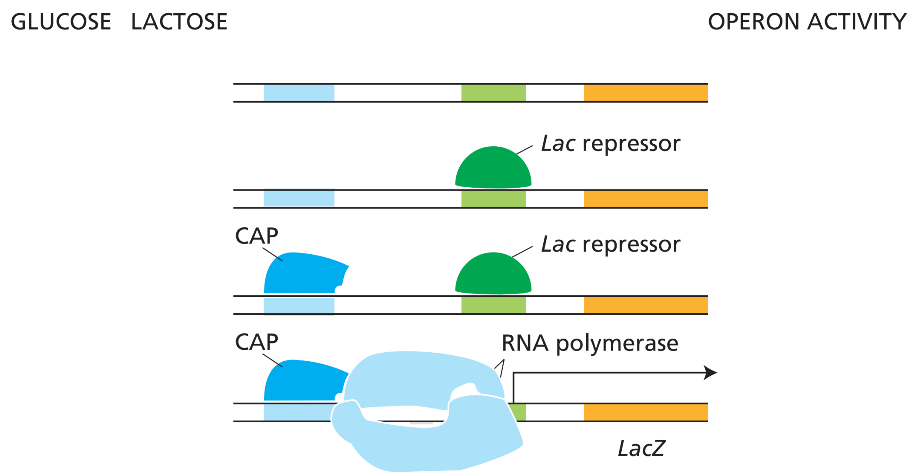


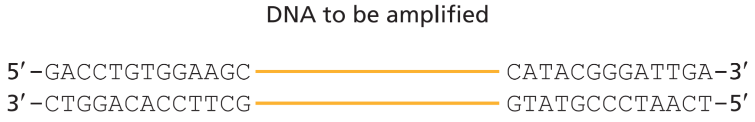
**Figure 1** Arrangement of binding sites and the four possible combinations of transcription regulators on the promoter for the *lac* operon

6. Which, if any, of the restriction nucleases listed in **Table 1** will definitely cleave a segment of cDNA that encodes the peptide `KIGPACF`? (See ‘Short Course’ for the genetic code). *Hint: <u>N</u> stands for any <u>N</u>ucleotide.*

**Table 1** A set of restriction nucleases and their recognition sequences.

| Restriction nuclease | Recognition sequence |
|---|---|
| AluI | AGCT |
| Sau96I | GGNCC |
| HindIII | AAGCTT |

7. You want to amplify the DNA between the two stretches of sequence shown in **Figure 2**. Of the listed primers, choose the pair that will allow you to amplify the DNA by PCR.



primers

(1) 5′-GACCTGTGGAAGC-3′
(2) 5′-CTGGACACCTTCG-3′
(3) 5′-CGAAGGTGTCCAG-3′
(4) 5′-GCTTCCACAGGTC-3′
(5) 5′-CATACGGGATTGA-3′
(6) 5′-GTATGCCCTAACT-3′
(7) 5′-TGTTAGGGCATAC-3′
(8) 5′-TCAATCCCGTATG-3′

**Figure 2** DNA to be amplified and potential PCR primers.

## Part II. Bonus

*This part is for testing your capacity for analysis, data handling, and modelling in molecular biology. Try to get the bonus and show off your talent.*

1. In principle, a eukaryotic cell can regulate gene expression at any step in the pathway from DNA to the active protein.

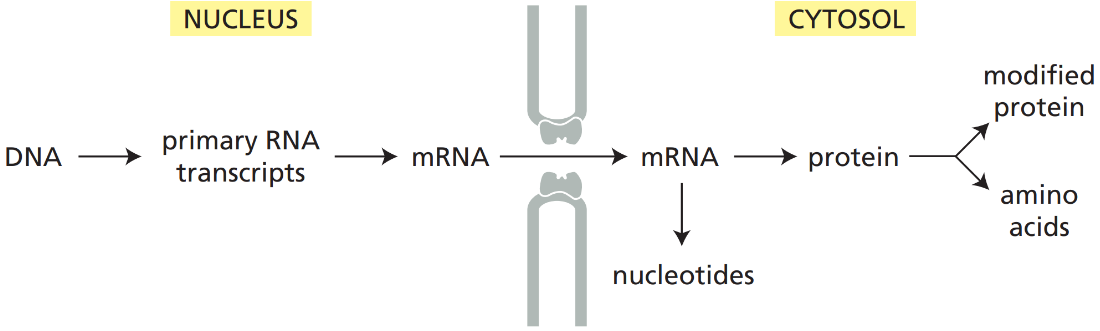


**Figure 3** Six steps at which the pathway for eukaryotic gene expression can be controlled.

a) Place or describe the types of control listed below at appropriate places on the diagram in **Figure 3**

   i. mRNA degradation control
   ii. protein activity control
   iii. RNA processing control
   iv. RNA transport and localization control
   v. transcriptional control
   vi. translational control

b) Which of the types of control listed above are unlikely to be used in bacteria?

2. In **Figure 4**, the sequences of 13 promoters recognized by the $\sigma^{70}$ factor of RNA polymerase have been aligned. Deduce the consensus sequences for the -10 and -35 regions of these promoters.

```
TCTCAACGTAACACTTTACAGCGGCG--CGTCATTTGATATGATGC-GCCCCGCTTCCCGATAAGGG
GATCAAAAAAATACTTGTGCAAAAAA--TTGGGATCCCTATAATGCGCCTCCGTTGAGACGACAACG
ATGCATTTTTCCGCTTGTCTTCCTGA--GCCGACTCCCTATAATGCGCCTCCATCGACACGGCGGAT
CCTGAAATTCAGGGTTGACTCTGAAA--GAGGAAAGCGTAATATAC-GCCACCTCGCGACAGTGAGC
CTGCAATTTTTCTATTGCGGCCTGCG--GAGAACTCCCTATAATGCGCCTCCATCGACACGGCGGAT
TTTTAAATTTCCTCTTGTCAGGCCGG--AATAACTCCCTATAATGCGCCACCACTGACACGGAACAA
GCAAAAATAAATGCTTGACTCTGTAG--CGGGAAGGCGTATTATGC-ACACCCCGCGCCGCTGAGAA
TAACACCGTGCGTGTTGACTATTTTA-CCTCTGGCGGTGATAATGG--TTGCATGTACTAAGGAGGT
TATCTCTGGCGGTGTTGACATAAATA-CCACTGGCGGTGATACTGA--GCACATCAGCAGGACGCAC
GTGAAACAAAACGGTTGACAACATGA-AGTAAACACGGTACGATGT-ACCACATGAAACGACAGTGA
TATCAAAAAGAGTATTGACTTAAAGT-CTAACCTATAGGATACTTA-CAGCCATCGAGAGGGACACG
ACGAAAAACAGGTATTGACAACATGAAGTAACATGCAGTAAGATAC-AAATCGCTAGGTAACACTAG
GATACAAATCTCCGTTGTACTTTGTT--TCGCGCTTGGTATAATCG-CTGGGGGTCAAAGATGAGTG
```

**Figure 4** Sequences recognized by $\sigma^{70}$ factor. Different σ factors are designated by their molecular masses; $\sigma^{70}$ has a mass of 70 kilodaltons. Dashes represent spaces that have been added to maximize alignment of sequences in the -10 and -35 regions.

3. Genetic analyses in bacteria in the 1950s provided the first evidence for the existence of transcription regulators. The lambda repressor, one such regulator, is encoded by the bacterial virus, bacteriophage lambda. The lambda repressor binds as a dimer to critical sites on the bacteriophage lambda genome to keep the lytic genes turned off, which allows the bacteriophage lambda genome to be maintained as a silent resident in the bacterial genome. Each molecule of the repressor consists of an N-terminal DNA-binding domain and a C-terminal dimerization domain (Figure 5). Upon induction (for example, by irradiation with ultraviolet light), the genes for lytic growth are expressed, bacteriophage lambda progeny are produced, and the bacterial cell lyses to release the viral progeny. Induction is initiated by cleavage of the lambda repressor at a site between the DNA-binding domain and the dimerization domain. In the absence of bound repressor, RNA polymerase initiates transcription of the lytic genes, triggering lytic growth. Given that the number (concentration) of DNA-binding domains is unchanged by cleavage of the repressor, why do you suppose its cleavage results in its removal from the DNA?

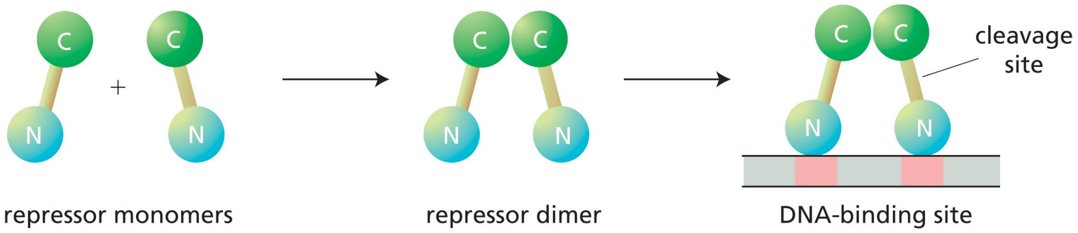


**Figure 5** Domains of the lambda repressor and the binding of repressor dimers to DNA.

4. When *cis*-regulatory sequences were initially found to influence activity at distant promoters, two principal models were invoked to explain this action at a distance. In the “DNA looping” model, direct interactions between transcription regulators bound at *cis*-regulatory sequences and the distant promoters were proposed to stimulate RNA polymerase. In the “scanning” model, RNA polymerase (or a transcription regulator) was proposed to bind at the regulatory sequence and then slide along the DNA until it reached the promoter. These two models were distinguished using a *cis*-regulatory sequence on one piece of DNA and the *β-globin* gene with its promoter on a separate piece of DNA (**Figure 6**). The *β-globin* gene was not expressed from the mixture of pieces. However, when the two segments of DNA were joined via a protein linker, the *β-globin* gene was expressed.

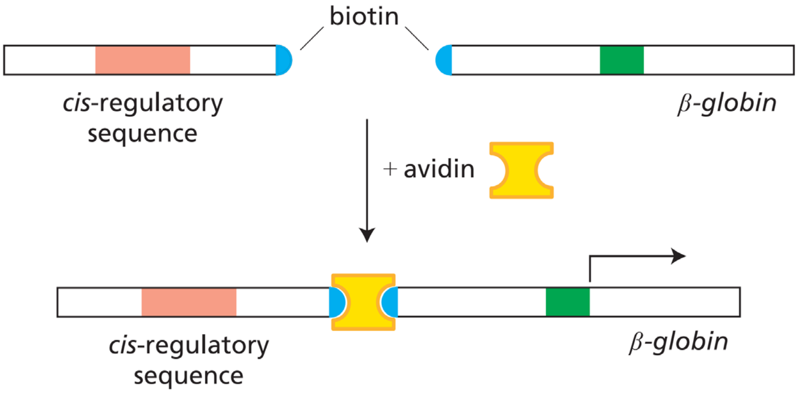
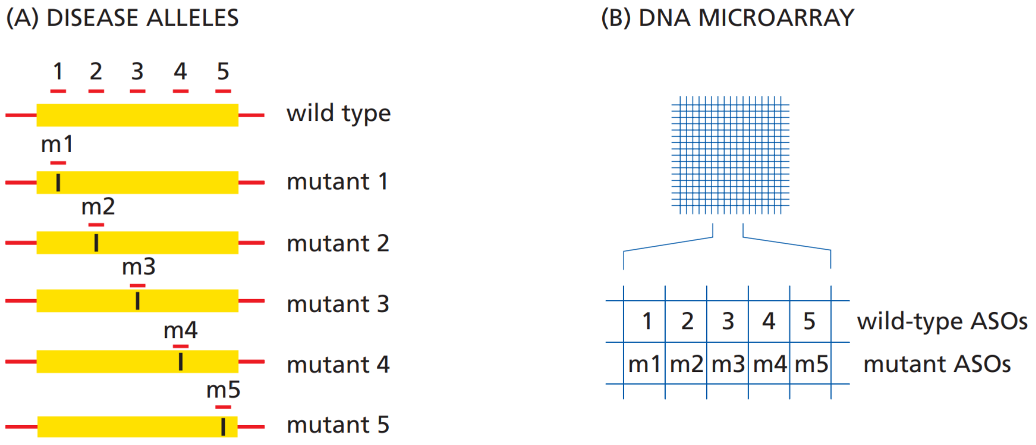


**Figure 6** Stimulation of *β-globin* gene expression by a *cis*-regulatory sequence linked via a protein bridge. Each DNA molecule carries biotin attached to one end, as shown. In the presence of the protein avidin, the two molecules are linked together and transcription occurs, as shown by the arrow above the *β-globin* gene.

How does this experiment distinguish between the DNA looping model and the scanning model? Explain your answer.

5. Now that news of your disease-specific DNA microarray has gotten around, you are being inundated with requests to analyze various samples. Just today you received requests from four physicians for help in the prenatal diagnosis of the same disease. Each of the pregnant mothers has a family history of this disease. You included on your array the five alleles known to cause this disease (Figure 7). Each of these alleles is recessive. You agree to help. You prepare samples of fetal DNA gotten by amniocentesis and hybridize them to your microarrays. Your data are shown in Figure 7C. Assuming that the five disease alleles shown in Figure 7A are the only ones in the human population, decide for each sample of DNA whether the individual will have the disease or not. Explain your reasoning (ASOs: antisense oligonucleotides).


(A) DISEASE ALLELES

(B) DNA MICROARRAY

(C) HYBRIDIZATION DATA

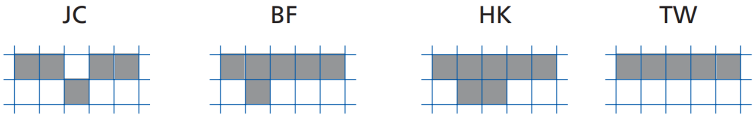


**Figure 7** DNA microarray analysis of alleles present in prenatal samples. (A) Wild-type and disease alleles. *Vertical lines* indicate the sites of the mutations in the disease-causing alleles. The ASOs specific for the disease mutations are shown as *red lines* and labeled m1, m2, etc. The corresponding ASOs that hybridize to the wild-type gene at sites that correspond to the position of the mutations are labeled 1, 2, etc. (B) DNA microarray. The arrangement of wild-type and mutant ASOs is indicated. (C) Hybridization data. Samples of fetal DNA were hybridized to DNA arrays. *Dark spots* indicate sites where hybridization occurred. *Letters* identify the patients.

# The 4th Week Question List of *Molecular Biology*

*2024-pre-iGEMer@OUC-Haide*

**Assignments: 'Short Course', Review & Question List**

Due: Monday, 7 August 2023, 11:59 PM

| 7-31 | RNA Processing<br>Protein Synthesis |
|---|---|
| Short Course | ☐ Chapter 38 RNA Processing in Eukaryotes<br>☐ Chapter 39 The Genetic Code<br>☐ Chapter 40 The Mechanism of Protein Synthesis |
| A Review of the CRISPR-Cas System | The Biology of CRISPR-Cas: Backward and Forward, 2018, Cell |
| Congratulation 🎉 See you in summer term! ⛱ | |

## Instruction

- Answers in any form will be acceptable unless a specific rule is required in the question. Try your best to make your answers clear, solid and logical. Additional figures, screenshots and photos of your handwriting are recommended to do this.
- Submit your assignment tasks via email and we will give you constructive feedback. Student emails like sunjunteng@stu.ouc.edu.cn are the best because we can immediately know who you are.
- File naming rules: Week[number]_[Your serial number]_[Your Chinese name]. E.g., Week1_0_孙俊腾.doc/pdf
- Questions and answers are ALL in English.
- Looking forward to your wonderful answer. 😀

# Part I. General Question

*This part is for reviewing your involvement in the 'Short Course' and your understanding of molecular biology. Try to emphasize the necessary details.*

1. Briefly answer: a) what is the role of CTD of RNA Pol II in coupling RNA synthesis and splicing? b) what are the conserved motifs involved in mRNA splicing via spliceosome (provide at least two examples)?
2. Describe: a) what are the two reaction steps for the formation of an aminoacyl-tRNA? b) how does the ribosome identify the AUG specifying initiation in prokaryotes? c) the differences of initiation process during protein synthesis in bacteria vs. eukaryotes.
3. For the RNA sequence below, indicate the amino acids that are encoded in the three reading frames. If you were told that this segment of RNA was in the middle of an mRNA that encoded a large protein, would you know which reading frame was used? How so?

   ...AGUCUAGGCACUGA...

4. Imagine the two situations shown in Figure 1. In cell 1, a transient signal induces the synthesis of protein A, which is a transcription activator that turns on many genes including its own. In cell 2, a transient signal induces the synthesis of protein R, which is a transcription repressor that turns off many genes including its own. In which, if either, of these situations will the descendants of the original cell "remember" that the progenitor cell had experienced the transient signal? Explain your reasoning.

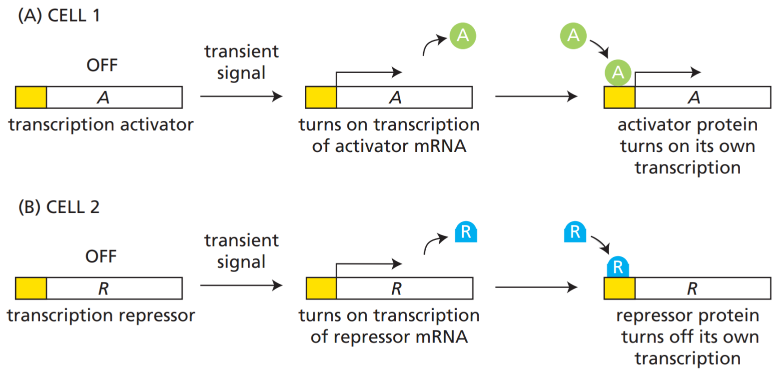


**Figure 1** Gene regulatory circuits and cell memory. (A) Induction of synthesis of transcription activator A by a transient signal. (B) Induction of synthesis of transcription repressor R by a

transient signal.

5. The rules for wobble base-pairing in bacteria and eukaryotes are shown in Table 1. On the left side of the table, the rules are expressed as a wobble codon base and its recognition by possible anticodon bases. [The anticodon base I (inosine) is a common modification in tRNAs; it is generated by deamination of A.] Reformulate these rules as particular anticodon bases and their recognition by possible codon bases, as suggested by the partial information on the right side of the table.

**Table 1** Rules for wobble base-pairing between codon and anticodon.

| | Wobble codon base | Possible anticodon base | Wobble anticodon base | Possible codon base |
|---|---|---|---|---|
| Bacteria | U<br>C<br>A<br>G | A, G, or I<br>G or I<br>U or I<br>C or U | U<br>C<br>A<br>G<br>I | |
| Eukaryotes | U<br>C<br>A<br>G | G or I<br>G or I<br>U<br>C | U<br>C<br>A<br>G<br>I | |

6. You have just gotten back the results from an RNA-seq analysis of mRNA from liver. You had anticipated counting the number of reads of each mRNA to determine the relative abundance of different mRNAs. But you are puzzled because many of the mRNAs have given you results like those shown in Figure 2. How is it that different parts of an mRNA can be represented at different levels?

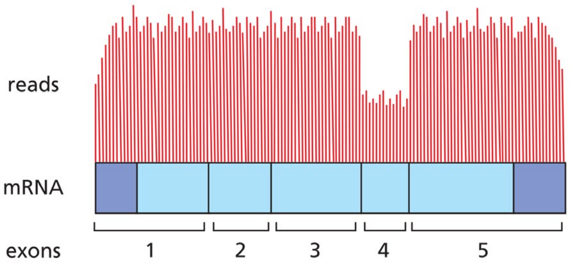


**Figure 2** RNA-seq reads for a liver mRNA. The exon structure of the mRNA is indicated, with protein coding segments indicated in *light blue* and untranslated regions in *dark blue*. The numbers of sequencing reads are indicated by the heights of the *vertical lines* above the mRNA.

## Part II. Bonus

*This part is for testing your capacity for analysis, data handling, and modelling in molecular biology. Try to get the bonus and show off your talent.*

1. CRISPR/Cas9-guide RNA complexes hold enormous promise as aids for genome engineering in plants and animals. The CRISPR system is almost too good to be true. You want to test just how specific the Cas9-guide RNA complexes are; that is, whether they really recognize individual sites in the genome, which is the basis for their touted actions. You realize that you can test their specificity using the "DNA curtain" assay you have developed. You make the DNA curtain by tethering single molecules of bacteriophage lambda DNA (about 50,000 nucleotides) at one end, and then stretching them in the same direction by flowing buffer across the slide. You incubate the DNA curtain with a highly fluorescent version of Cas9—either loaded with the guide RNA or free—and visualize the distribution of the Cas9 by sensitive fluorescence microscopy, as shown in **Figure 3**.

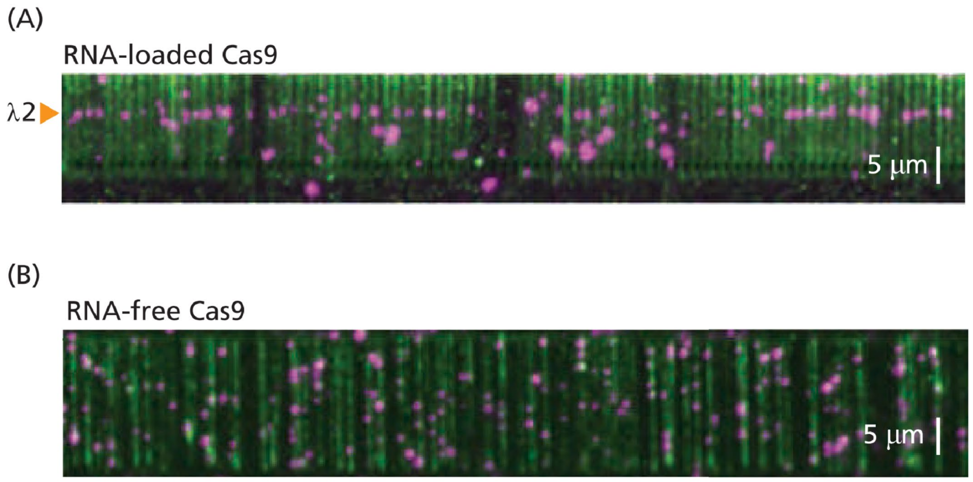


**Figure 3** DNA curtain assay for target binding by Cas9-guide RNA. (A) DNA curtain incubated with Cas9–guide RNA that matches one site in the lambda genome. Pink spots indicate the sites where Cas9–guide RNA is located. (B) DNA curtains incubated with Cas9 in the absence of guide RNA.

a) Phage lambda DNA has a single site that perfectly matches the guide RNA. Does your experiment support the existence of a single site that is recognized by the Cas9-guide RNA complex? Explain your answer.

b) One of the main concerns for using the CRISPR system is that the Cas9-guide RNA will bind to other sites that are related to the intended target (so-called off-target effects). Is there any evidence for off-target binding in your experiment? Do you think the results would be any different if you had a DNA curtain made from human chromosome 1 (about 250,000,000 nucleotides)?

2. Bacteriophage lambda can replicate as a prophage or lytically. In the prophage state, the viral DNA is integrated into the bacterial chromosome and is copied once per cell division. In the lytic state, the viral DNA is released from the chromosome and replicates many times. This viral DNA then produces viral coat proteins that enclose the replicated viral genomes to form many new virus particles, which are released when the bacterial cell bursts.

   These two states are controlled by the transcription regulators cI and Cro, which are encoded by the virus. In the prophage state, cI is expressed; in the lytic state, Cro is expressed. In addition to regulating the expression of other genes, cI is a repressor of transcription of the gene that encodes Cro, and Cro is a repressor of the gene that encodes cI (**Figure 4**).

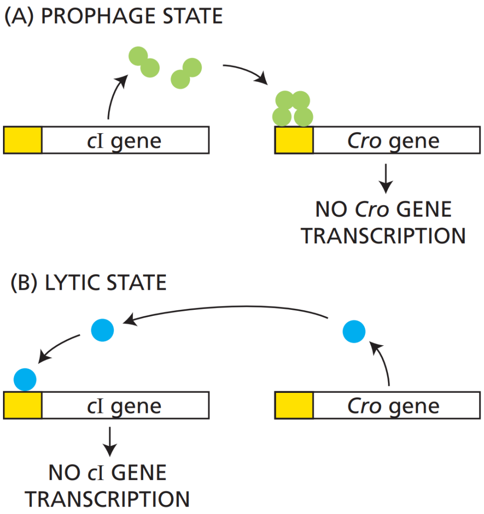


**Figure 4** Regulation of bacteriophage lambda replication by cI and Cro. (A) The prophage state. (B) The lytic state.

When bacteria containing a lambda prophage are briefly irradiated with ultraviolet (UV) light, cI protein is degraded.

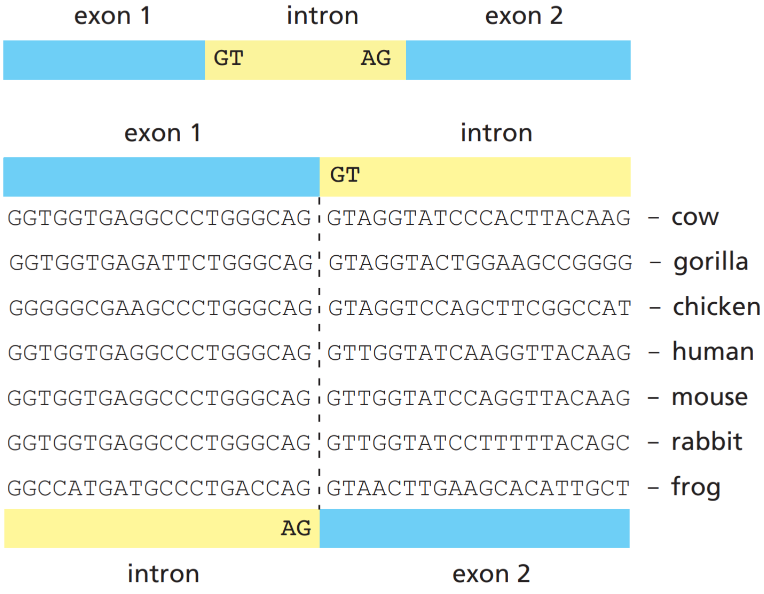
a) What will happen next?

b) Will the change in question A be reversed when the UV light is switched off?

c) How is this mechanism beneficial to the virus?

3. You have printed out a set of DNA sequences around the intron/exon boundaries for genes in the β-globin family, and have taken the thick file to the country to study for the weekend. When you look at the printout, you discover to your annoyance that there's no indication of where in the gene you are. You know that the sequences in Figure 5 come from one of the exon/intron or intron/exon boundaries and that the boundaries lie on the dotted line, but you don't know the order of the intron and exon. You know that introns begin with the dinucleotide sequence GT and end with AG, but you realize that these particular sequences would fit *either* as the start *or* the finish of an intron.

   If you cannot decide which side is the intron, you will have to cut your weekend short and return to the city (or find a neighbor with Internet access). In desperation, you consider the problem from an evolutionary perspective. You know that introns evolve faster (suffering more nucleotide changes) than exons because they are not constrained by function. Does this perspective allow you to identify the intron, or will you have to pack your bags?



```
GGTGGTGAGGCCCTGGGCAG GTAGGTATCCCACTTACAAG – cow
GGTGGTGAGATTCTGGGCAG GTAGGTACTGGAAGCCGGGG – gorilla
GGGGGCGAAGCCCTGGGCAG GTAGGTCCAGCTTCGGCCAT – chicken
GGTGGTGAGGCCCTGGGCAG GTTGGTATCAAGGTTACAAG – human
GGTGGTGAGGCCCTGGGCAG GTTGGTATCCAGGTTACAAG – mouse
GGTGGTGAGGCCCTGGGCAG GTTGGTATCCTTTTTACAGC – rabbit
GGCCATGATGCCCTGACCAG GTAACTTGAAGCACATTGCT – frog
```

**Figure 5** Aligned DNA sequences from the *β-globin* genes in different species. As indicated by the gene structures shown above and below, the DNA sequences could come from the boundary of exon 1 with the intron, or from the boundary of the intron with exon 2.

4. In nematodes, the choice between spermatogenesis and oogenesis in the hermaphrodite germ line depends on translational regulation of the *Tra2* and *Fem3* genes, as shown in **Figure 6**. When they are expressed, *Tra2* promotes oogenesis and *Fem3* promotes spermatogenesis. Translation of each gene's mRNA is regulated by the binding of proteins to elements within their 3′ untranslated regions. In each case, the protein-bound mRNA, although stable, is not efficiently translated and has a short poly-A tail. In each case, the mRNA in its unbound form is translationally active and has a long poly-A tail. How do you suppose that the lengths of the poly-A tails might affect the efficiency of translation of these mRNAs?

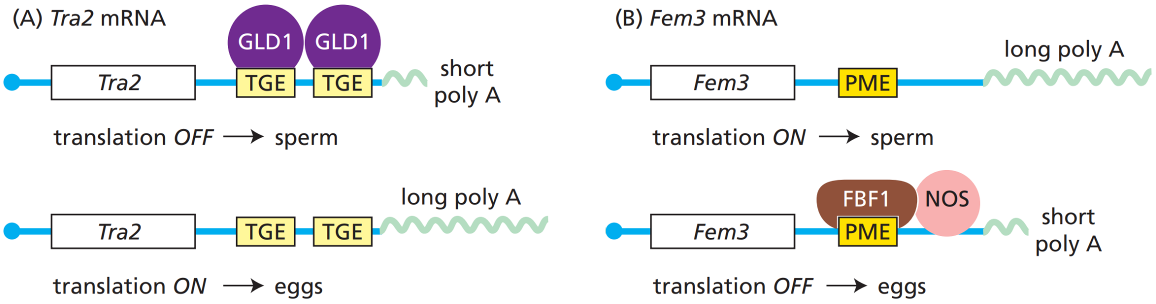


**Figure 6** Translational control of the choice between spermatogenesis and oogenesis. (A) *Tra2* mRNA. (B) *Fem3* mRNA. Control elements within the 3′ untranslated regions of the two genes are shown, along with the specific proteins that bind to those elements. TGEs: *Tra2* regulatory elements.

5. You are skeptical that internal ribosome entry sites (IRESs) really allow direct binding of the eukaryotic translation machinery to the interior of an mRNA. As a critical test of this notion, you prepare a set of linear and circular RNA molecules, with and without IRESs (**Figure 7A and B**). You translate these various RNAs in rabbit reticulocyte lysates and display the translation products by sodium dodecyl sulfate (SDS) gel electrophoresis (**Figure 7C**). Do these results support or refute the idea that IRESs allow ribosomes to initiate translation of mRNAs in a cap-independent fashion? Explain your answer. *If you want to learn more about IRESs and basics of genetic design:* *https://technology.igem.org/mammalian/guide#h-individual-parts*

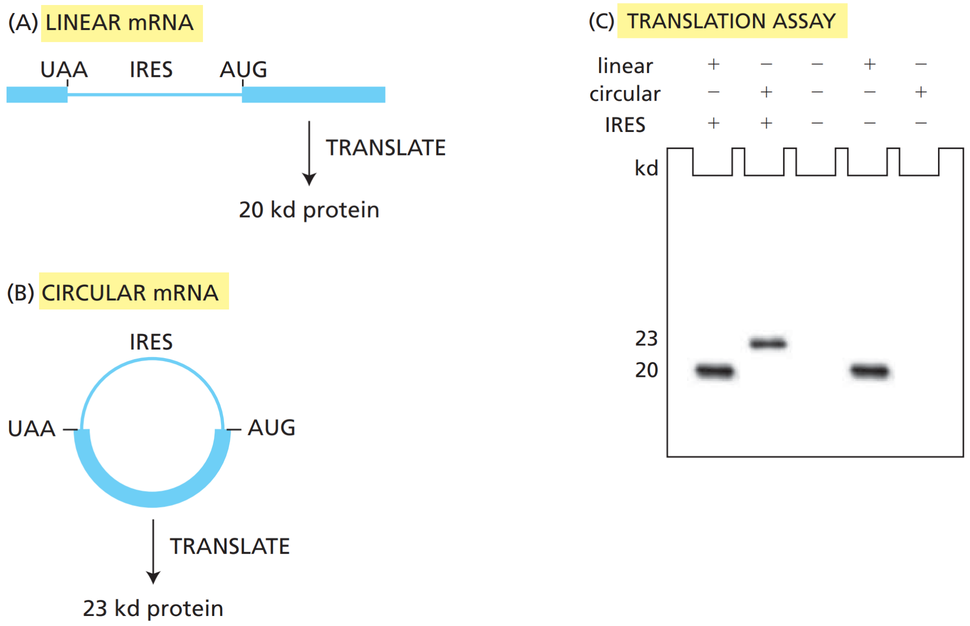


**Figure 7** Analysis of effects of IRESs on translation. (A) Linear mRNA that contains an IRES. The structure of the mRNA without the IRES was the same. (B) Circular mRNA that contains an IRES. The structure of the mRNA without the IRES was the same. Circular RNAs were prepared by ligating the ends together; they were purified from the linear starting molecules by gel electrophoresis. (C) Display of translation products from various species of mRNA.

6. You have just joined a laboratory that is analyzing the nuclear transport machinery in yeast. Your advisor, who is known for her extraordinarily clever ideas, has given you a project with enormous potential. In principle, it would allow a genetic selection for conditional-lethal mutants in the nuclear transport apparatus.

   She gave you two plasmids. Each plasmid consists of a hybrid gene under the control of a regulatable promoter (**Figure 8**). The hybrid gene is a fusion between a gene whose product is normally imported into the nucleus and the gene for the restriction nuclease EcoRI. The plasmid pNL+ contains a functional nuclear localization signal (NLS); the plasmid pNL- does not have an NLS. The promoter, which is from the yeast *Gal1* gene, allows transcription of the hybrid gene only when the sugar galactose is present in the growth medium.

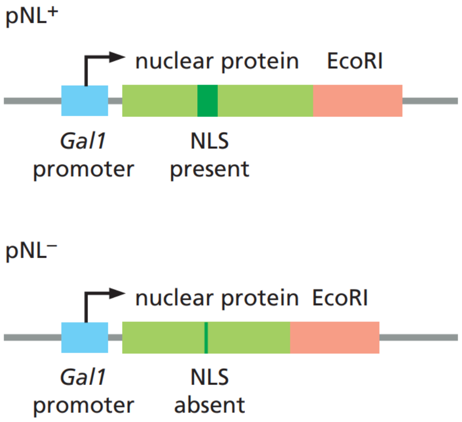


**Figure 8** Two plasmids for investigating nuclear localization in yeast. Plasmids are shown as linear molecules for clarity. Arrows indicate direction of transcription from the *Gal1* promoter.

Following her instructions, you introduce the plasmids into yeast (in the absence of galactose) and then assay the transformed yeast in medium containing glucose and in medium containing galactose. Your results are shown in **Table 2**. You don't remember what your advisor told you to expect, but you know you will be expected to explain these results at the weekly lab meeting.

**Table 2** Results of proliferation experiments with yeast carrying plasmids pNL+ or pNL-.

| Plasmid | Glucose medium | Galactose medium |
|---|---|---|
| pNL+ | proliferation | death |
| pNL- | proliferation | proliferation |

a) Why do yeasts with the pNL+ plasmid proliferate in the presence of glucose but die in the presence of galactose, whereas yeasts with the pNL– plasmid proliferate in both media?

b) How might you use this system for a selection assay to isolate cells defective in nuclear transport?