Solution key- 7.013 EXAM 2

Question 1 (24 points)

Inflammation is a common manifestation of many infections and is associated with the **synthesis and secretion** of small peptides called cytokines.

a) Circle ALL the correct option(s) for each of the following. (***6pts or 1 each)***

i. Which protein functions as a transcription factor: **Cytokine/ CR/ JAK/ STAT/ SOCS**?

ii. Location(s) of cytokine synthesis: **cytoplasm/ mitochondria/ ER membrane/ lysosomes**?

iii. The highest order of protein structure for **ACTIVE STAT**: **Primary/ secondary/ tertiary/ quaternary**?

iv. Proteins that act as kinases: **Cytokine/ CR/ JAK/ STAT/ SOCS?**

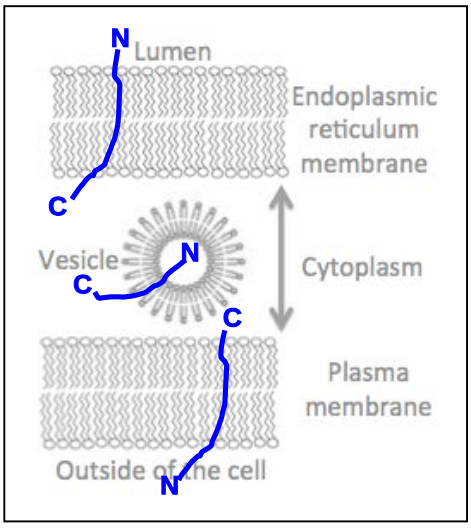
b) Consider the following **homozygous mutations** in different components of the signaling pathway.

- #1: CR lacks its extracellular domain that binds to cytokine ligand
- #2: STAT protein lacks its nuclear localization sequence
- #3: JAK protein is constitutively (always) **phosphorylated**
- #4: SOCS promoter sequence is heavily methylated

Complete the table for each of the following **mutations** in the **presence** of **cytokines**. ***(9pts or 3 each)***

Mutations	**CR active (*Yes/No*?)**	**JAK protein active (*Yes/No*?)**	**STAT protein dimerized (*Yes/No*?)**	**SOCS expressed (Yes/ No)?**	**Cell division (Yes/ No)?**
1	No	No	No	No	No
2	Yes	Yes	Yes	No	No
Both 3 & 4	Yes	Yes	Yes	No	No

c) Assuming that CR has **ONE transmembrane domain**, on the schematic below, draw its orientation in the ER membrane, vesicle and plasma membrane and label the N and C ends of the CR on each drawing. ***(5pts, 2 for ER, 1 for vesicle and 2 for cell membrane)***



d) You identify a mutant cell, which produces a misfolded SOCS.

i. Propose a cellular mechanism that can refold the SOCS protein into its functional 3D- conformation.

*The chaperone proteins guide the folding of the proteins in a cell. They can be: Molecular chaperons or Folding chaperones that they isolate the protein from other components in the cytoplasm by acting as lids. The protein comes out only once it is folded properly. **(2pts)***

ii. Briefly **explain** why the misfolded proteins tend to aggregate within the cytoplasm of a cell.

*If the proteins do not fold properly, their inner hydrophobic core is exposed. This causes it to aggregate with other proteins to form large aggregates through hydrophobic interactions. These aggregates are the cause of neurodegenerative diseases. **(2pts)***

Question 2 (15 points)

The following is the DNA sequence for the transcription initiation region of the **CR gene**. ***Note:*** *Part of the* ***promoter region*** *is boxed and the direction of transcription is shown as an →. Transcription* ***begins at*** *and includes the underlined* ***T/A*** *base pair.*

```
                  #1    ————————→      #2        #3
5'-|--CGGACTGCTAAAATA|GCTTAATCGATGTGCTGCACGCAAGGCCATACGCTTGTGGTTCTT----3' TOP
3'-|--GCCTGACGATTTTAT|CGATTAGCTACACGACGTGCGTTCCGGTATGCGAACACCAAGAA----5' BOTTOM
```

a) Identify the template strand for transcription: **Top / Bottom**? ***Bottom*** ***(3pts)***

b) Write the **first 6 nucleotides** of the newly transcribed CR mRNA: **5'-*UAAUCG*-3'** (***3pts, 2 if T instead of U)***

c) Write the **first 2 amino acids** of the newly synthesized CR protein: **N-*Met-Cys*-C** (***3pts, 1 for Met)***

d) You create mutants 1 - 3 by substituting the C/ G base pairs (bold and shaded) at positions 1- 3 by a T/A base pair as shown on the right.

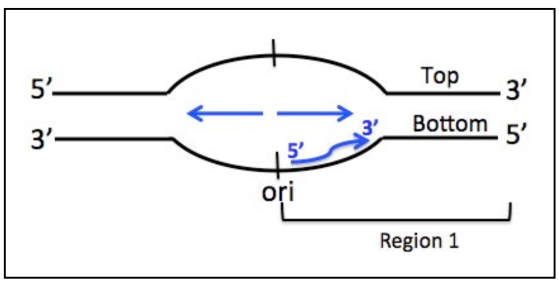
C → T (Top strand)
G → A (Bottom strand)

Which mutant(s) (*1/ 2/ 3*) would produce a truncated CR protein?
Explain why you selected this option and **NOT** the others. ***(6pts, 2for each)***

Mutation 1 *is within the promoter sequence of CR gene. it can regulate the rate of CR gene expression but Not the CR amino acid sequence. So it will not generate an hydrophobic pocket on the protein surface to cause sickle cell anemia.* ***Mutation 2*** *is a silent mutation within the reading frame which neither changes the rate of CR gene expression nor the CR amino acid sequence.* ***Mutation 3*** *generates an in-frame premature stop codon resulting in a truncated CR protein.*

Question 3 (14 points)

Below is a segment of replicating DNA from an ori site in the CR expressing skin cells.



a) On the drawing, show the direction of movement of **each replication fork** by drawing an arrow (→). ***(2pts)***

b) In **Region 1**, draw the primer for the leading (continuous) strand as a dashed arrow (**=>**) and label its 5' and 3' ends. (***2pts with correct end labels)***

b) The following is a segment of **a newly replicated DNA segment**. Which base was wrongly inserted during replication: The bold and shaded **"G" or "A"**? **Explain** why you selected this base. ***Note:*** *The methylated Cytosine bases are bold and underlined.*

```
5'ATCGGGTACCGATTCTGAAG3'  TOP
3'TAGCCCATGGCTAAGAATTC5'  BOTTOM
```

The shaded "G" is the wrongly incorporated base since it is a part of the newly synthesized DNA strand, which is hypo-methylated. The BOTTOM strand is methylated and therefore the template strand. ***(4pts, 2 for explain)***

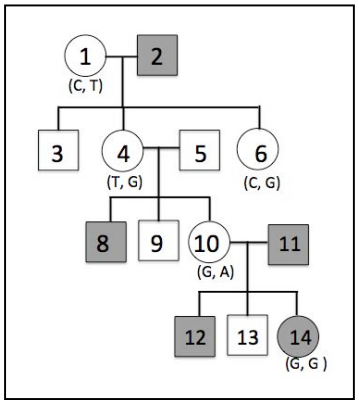
c) You grow the skin cells on plates that contain the following compounds.

- **Plate 1** includes arylstibonic acid, a compound that promotes DNA supercoiling.
- **Plate 2** includes a drug that blocks the 3'->5' exonuclease activity of DNA polymerase.
- **Plate 3** includes a protease that degrades the single stranded DNA binding protein.

Which plate(s) will have cells that are **NOT** dividing and **why**: **1/ 2/ 3**? **(6pts, 3 for each correct plate)**

In plate 1, arylstibonic acid will inhibit the topoisomerase to promote supercoiling and thus inhibit replication. In plate 2, DNA will replicate but will have more errors since the newly replicating DNA will not be proofread. In Plate 3,in the absence of SSDBP the genome will be unstable and will likely degrade. So Cells in plates 1 and 3 will not replicate their genome and hence will not divide.

Question 4 (14 points)
The pedigree below shows the inheritance of hyper-inflammation due to the mutations in the CR gene. ***Note:*** #5 does not have a disease related allele. *Affected individuals are shaded. The CR-associated SNPs for some individuals are indicated.*



a) Give the **mode of inheritance** of this disease and identify the SNP(s) associated with the disease phenotype.

i. **Mode of inheritance:** ***X-linked (2pts) recessive (2pts)***

ii. **Disease associated SNP(s):** ***G (2pts)***

b) You observe that the mature CR mRNA in affected individuals is longer (in bases) than that in normal healthy individuals due to disease-associated SNP. Circle **TWO** possible locations of SNP: **Promoter/ Exons/** ***Splice donor site/ Splice acceptor site*****/ 5'UTR/ 3'UTR**? (***4pts)***

c) You create a model of the disease by using the CRISPR-Cas9 endonuclease complex. The double stranded nicks of the target sequence by CRISPR-Cas9 are an example of **hydrolysis/ condensation** reaction, which breaks the **covalent/ ionic/ hydrogen** bond. (***4pts, 2 each)***

Question 5 (16 points)
Your fellow classmate is studying families where the affected individuals do not express the CR gene. Further analysis reveals that the enhancer sequence corresponding to CR gene in these patients cannot bind to specific transcription factors. She wants to further characterize the enhancer sequence.

a) Which library should she use to identify the bacterial clone carrying the enhancer sequence specific to CR gene in affected and healthy individuals: The **genomic/ skin cell cDNA/ skin cell expression** library? **Why?**
You will use genomic library since enhancer is a regulatory DNA sequence that is a part of the genome but not the gene. It regulates transcription of gen(s) by allowing the binding of specific transcription factors to it in order to form the transcription initiation complex. ***(4pts, 2 for explanation)***

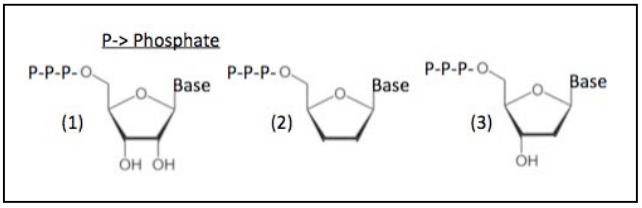
b) She isolates the plasmid that has the enhancer sequence insert and PCR amplifies the enhancer sequence (shown below) from affected and healthy individuals. Give the sequence of the 6-bases long primer to make the…. ***(4pts, 2 for each)***

```
5'GCTAGG__________CGTCAA3' TOP
3'CGATCC__________GCAGTT5' BOTTOM
          Enhancer
```

I. **The Top strand:** 5'*GCTAGG3'*

II. **The Bottom strand:** 5'*TTGACG*3'

c) She sequences the PCR amplified enhancer sequence. Which of the following nucleotides is used in DNA sequencing but NOT in PCR and **why**? You would also use 2'3'ddNTP (shown as 2 in the drawing to the left) to terminate the reaction. ***(4pts, 2 for explanation)***



d) She finds that the sequence of the CR enhancer in affected and healthy individuals is the SAME. Select an alternative mechanism from below that explains why CR gene is **not transcribed** in patients. ***(4pts, 1 point for the first two and 2 for the 3rd)***

i. Mature mRNA corresponding to CR gene lacks the 7- Methyl-Guanine at its 5' end *(Incorrect)*
ii. DNA demethylase removes the methyl group from the bases in the enhancer sequence *(Incorrect)*
iii. Histone proteins bound to the CR- enhancer region are de-acetylated *(Correct)*

Question 6 (17 points)
You would like to understand where the CR protein is localized in the cell, and whether its location changes with inflammation. To do this you plan to ligate the cDNA sequence corresponding to the C-terminus of CR gene with the cDNA sequence corresponding to the N- terminus of GFP gene to make a **CR-GFP fusion cDNA** that encodes the **CR-GFP fusion protein**.

.

a) What part of the **CR-GFP fusion protein** can inform you of CR location? *GFP through its fluorescence* ***(1pt)***

The following is the partial cDNA sequence encoding the C- terminus of the CR gene. ***Note:*** *The DNA corresponding to the stop codon is bold and underlined. The sequence specifically recognized by each restriction enzyme is shown in gray. Each codon is separated from the next by a space.*

```
            1                       3                5
CR:  5'AAA ATT CTG CAG AAT ACA ATT CCG CTG CAG TAG TTT GAA TTC ATC3'
     3'TTT TAA GAC GTC TTA TGT TAA GGC GAC GTC ATC AAA CTT AAG TAG5'
```

The following is the partial cDNA sequence encoding the N- terminus of GFP gene. ***Note:*** *The DNA corresponding to the start codon is bold and underlined. The recognition sequence for each restriction enzyme is shown in gray. Each codon is separated from the next by a space.*

```
            2           5           4
GFP: 5'ATG TGC AGG GCG GAA TTC GGG TTG CAA ATG CCA CTC GAG GAA TTC…3'
     3'TAC ACG TCC CGC CTT AAG CCC AAC GTT TAC GGT GAG CTC CTT AAG…5'
```

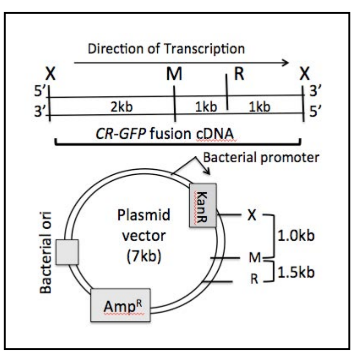
The recognition sequences and the cleavage sites (*indicated by /*) for each enzyme are given below.

1	2	3	4	5
5'C/TGCA G3' 3'G ACGT/C5'	5'G/TGCA G3' 3'C ACGT/C5'	5'C TGCA/G3' 3'G/ACGT C5'	5'T TGCA/A3' 3'A/ACGT T5'	5'G/AATT C3' 3'C TTAA/G5'

b) Complete the table below for each pair of restriction enzyme. ***(9pts, 3 each)***

Restriction enzyme pair used to digest CR and GFP cDNAs	Can you clone and express the CR-GFP fusion cDNA in the bacteria? **Why or why not?**
1 & 2	*No, it puts GFP cDNA sequence in CR-GFP cDNA out of frame*
3 & 4	*Yes, it keeps both CR and GFP cDNA sequences in frame*
5 & 5	*No, although it keeps both CR and GFP cDNA sequences in frame it does not remove the in-frame stop codon between CR and GFP.*

You clone the CR-GFP fusion gene into the following plasmid and use it to transform the bacteria. ***Note:*** *Both the CR-GFP fusion cDNA and the plasmid have the sequence for restriction enzymes X, M & R. The plasmid also has the ampicillin resistance (amp^R) and kanamycin resistance (Kan^R) genes.*

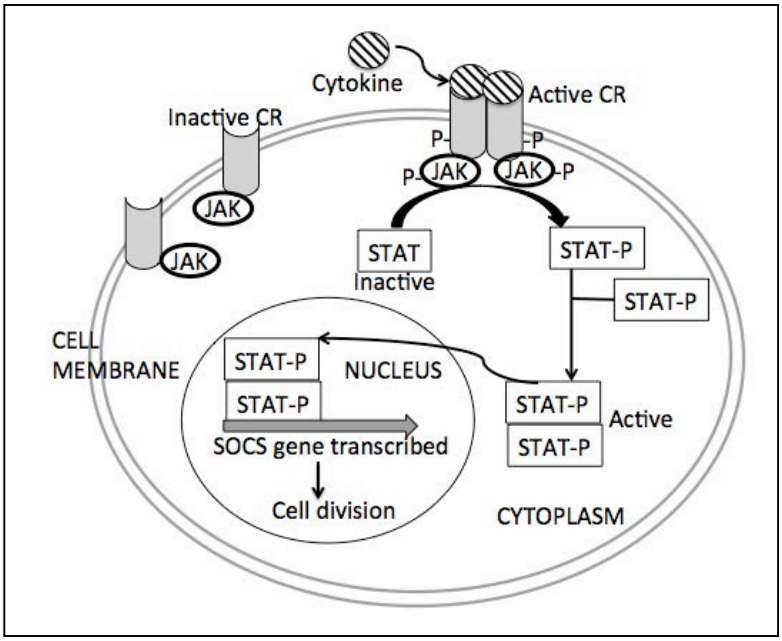


c) How would you **select and screen** for bacterial colonies that have the recombinant plasmid? *You would plate the bacteria in amp containing plate (Plate 1) and replica plate them on kan containing plate (Plate 2) and look for amp^{R}kan^S colonies that will grow on Plate 1 but not in plate 2. (4pts, 2 for kan and 2for amp) GFP screening also accepted*

d) You analyze two bacterial colonies that have the recombinant plasmid with the CR-GFP insert. Which restriction enzyme would you use to determine the orientation of the CR-GFP insert within the recombinant plasmid**: X/ M/ R**? **Explain,** why you selected this option. *Enzyme R is assymetrically located in the CR-GFP CDNA unlike enzyme M. The recombinant plasmid cut with R will give DNA fragments of size 8.5kb, 3.5kb if oriented correctly and 5.5kb and 6.5kb if inserted opposite to the orientation of the promoter.* ***(3pts)***

Signaling pathway for Question 1

Step 1: Cytokine receptors (CR) remain bound to JAK proteins and they are both dephosphorylated when inactive.



Step 2: Binding of cytokines to CR causes the dimerization of CR.

Step 3: JAK proteins bound to the cytoplasmic domains of the CR phosphorylate each other (shown as –P). They also phosphorylate and activate the dimerized CR.

Step 4: Activated CR phosphorylates and dimerizes STAT proteins and this results in its activation.

Step 5: Active **STAT dimer** moves to the nucleus and promotes the transcription of the *SOCS* gene by binding to the *SOCS* promoter sequence. This promotes cell proliferation.

Codon Chart

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

7.013 Introductory Biology

Spring 2018