Name_____**Rec. Section**_____ **TA**_____

7.012 Fall 2018: Problem Set 4

Due: Fri 10/26/2018

The solutions to these problems must be submitted electronically to your TA through the 7.012 Stellar site. All submissions must be received before 9:50 AM on October 26, 2018. Check your file to ensure it was successfully submitted. Only the material that is received prior to the deadline will be graded, no additional material will be accepted after the deadline.

Question 1 (4 points)

Fusion proteins are proteins that are created by joining two genes that code for two different proteins together in a way that results in the two proteins being linked together as a single protein (a single polypeptide chain). You want to make a fusion protein to detect the Protein X by adding the green fluorescent protein (GFP) sequence to the C-terminus of Protein X. The stop codon of Gene X (the gene that encodes Protein X) and start codon of GFP gene are shown in bold font for your convenience. (Some amino acids can be lost or gained near the fusion junction without affecting the function of either protein.)

Partial sequence of Gene X around its stop codon (shown in bold)

```
531-AAAGAGATCTGCCTGCAGGATTCGCTTTAAAAAATGCATTTAG
    TTTCTCTAGACGGACGTCCTAAGCGAAATTTTTACGTAAATC
         BglII   PstI          DraI    NsiI
```

Partial sequence of GFP around its start codon (shown in bold)

```
                  DraI
CTGCAGAGGGAGATCTTTAAAATGCATCCAAGCGGC
GACGTCTCCCTCTAGAAATTTTACGTAGGTTCGCCG
 PstI      BglII      NsiI
```

*Bgl*II: 5′-AGATCT-3′
3′-TCTAGA-5′

*Pst*I: 5′-CTGCAG-3′
3′-GACGTC-5′

*Dra*I: 5′-TTTAAA-3′
3′-AAATTT -5′

*Nsi*I : 5′-ATGCAT-3′
3′-TACGTA-5′

A.) Given the choice among restriction enzymes BglII, PstI, DraI and NsiI, which of these four enzymes would be best to cut the Protein X sequence and the GFP sequence to make this fusion protein?

B.) For each enzyme you did not use, explain why you would not use it.

Question 1, continued

C.) Assume that you use the restriction enzyme you chose in part A to paste the two genes to produce the fusion protein between Protein X and GFP. (i) Write out the DNA sequence of the coding strand that corresponds to the fusion starting with nucleotide position 531 (shown) and ending with the start codon of GFP. Underline the part of the sequence that came from protein X. (ii) Write out the amino acids of the junction including the last amino acid from Protein X, new amino acids added in the fusion between Protein X and GFP protein and the first amino acid of GFP. The genetic code can be found on the last page.

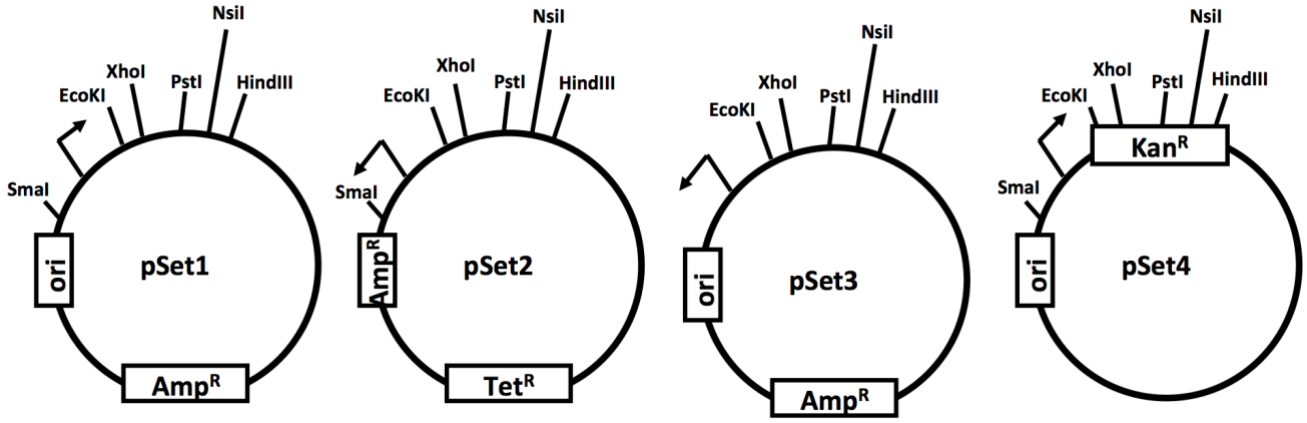
D.) Below is the recognition site of a new restriction enzyme, NooBI that has just come to market. As a promotion your first purchase of NooBI is free! Does this enzyme produce ends that are compatible with the ends produced by any of the enzymes shown above? If so could you use NooBI for this cloning strategy? Explain your answer.

NooBI: 5′-AGA↓TCT-3′
3′-TCT↑AGA-5′

E.) Assume that now you want to make a new fusion protein. This time you would like to have GFP fused to <u>the N terminus of Protein X</u> instead of the C-terminus of Protein X. Which parts of each gene sequence should you look at for potential restriction enzyme recognition sites? Consider start and stop codon regions for each gene.

Question 2 (4.5 points)

Find below potential plasmids to clone and express a gene of interest in E. coli. Please note that the promoter is shown with an arrow and that the promoter indicates the direction of transcription. Also shown are restriction enzyme recognition sites to be used to insert the gene of interest. Assume all restriction enzyme sites are in close proximity to the promoter to allow efficient transcription. Amp^R, Tet^R and Kan^R are ampicillin-, tetracycline-, and kanamycin-resistance genes, respectively. Please note that all antibiotic resistance genes have their own promoter (not shown).



A.) Answer the following questions about cloning genes into plasmids.
i) Why do you need an antibiotic resistance gene in a plasmid?

ii) Where does replication begin in a plasmid?

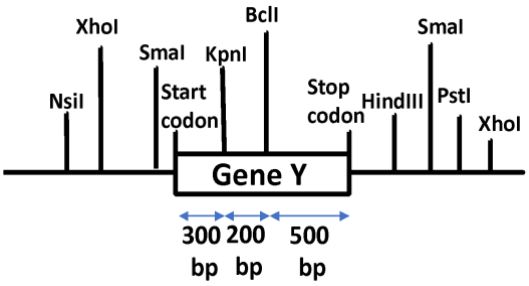
iii) Where would you want to insert your gene of interest in relationship to the promoter? Does the orientation of the gene matter? Explain.

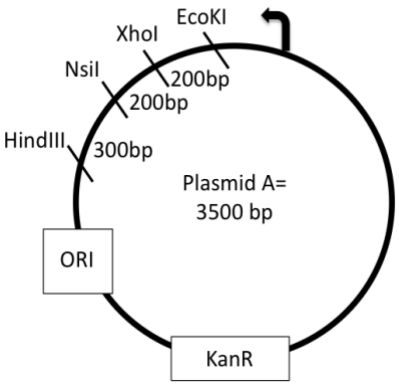
iv) State which plasmid above would be the best choice for cloning your gene of interest. Explain why you chose this plasmid and not the other three.

B.) Assume that you want to clone a human protein into the plasmid of your choice from part A(iv) and use it to express the protein in E. coli. Your roommate says that the promoter on your plasmid should be an E. coli promoter and not a human promoter. Do you agree with your roommate? Explain your answer.

C.) Which of the following sequences would be the best choice when you want to insert your gene of interest into the plasmid to produce a protein: (i) the human genomic sequence of your gene of interest or (ii) a DNA sequence matching the mature human mRNA of your gene of interest (cDNA)? Explain your answer.

Question 3 (3.5 points)
Below is the sequence for Gene Y. You would like to clone Gene Y into Plasmid A, mapped below.





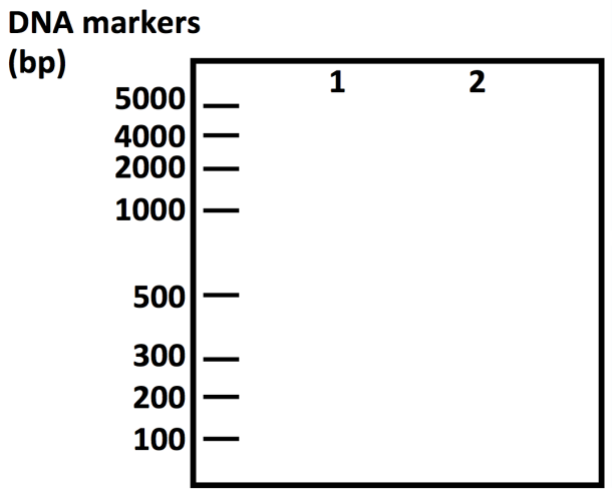
You cut Gene Y and Plasmid A with XhoI restriction enzyme and then ligate Gene Y into Plasmid A. You transform bacteria with the ligated Gene Y and Plasmid A mix and allow colonies to grow on a Petri plate containing Kanamycin. Then you purify plasmid from various colonies on your plate.

A.) A postdoc in the lab says you should check the size of your construct (plasmid + Gene Y) by using a restriction enzyme like EcoKI that makes a single cut in the plasmid and run it on an agarose gel that separates DNA fragments by size. How could this gel indicate whether or not your ligation of Gene Y into Plasmid A was successful?

B.) A grad student in the lab says you should also use plasmid purified from a few different colonies from your plate and check whether Gene Y was inserted in the correct orientation. She gives you EcoKI and NsiI restriction enzymes. Explain if using these two enzymes to cut plasmid could indicate the orientation of Gene Y.

C.) On the agarose gel image shown on the next page (DNA size markers are shown in nucleotides), mark which end the (+) and (-) electrodes go on.

D.) A UROP in the lab suggests one final check of your plasmid. She recommends that you digest the plasmid with HindIII to check if Gene Y was inserted in the correct orientation, namely with the start→stop codon aligned with the direction of the promoter arrow. Draw the expected results of this digestion on the gel if (i) Gene Y was inserted in the correct orientation or (ii) Gene Y was inserted in the incorrect orientation. Explain your answer. (Assume that the restriction enzyme recognition sites around Gene Y are very close to Gene Y and ignore their contribution to the size of the Gene Y fragment.)



Question 4 (3 points)

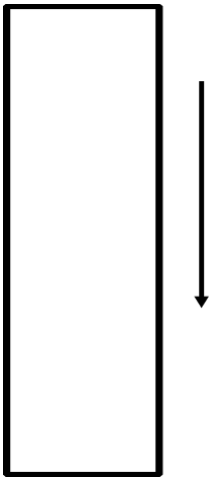
You would like to sequence the short DNA sequence denoted in bold below:

5′ AATTGG**GGACTAACGA**GGGCCT 3′
3′ TTAACC**CCTGATTGCT**CCCGGA 5′

Primer 1: 5′-GGGCCT-3′
Primer 2: 5′-TTAACC-3′
Primer 3: 5′-AGGCCC-3′
Primer 4: 5′-CCAATT-3′

A.) Which primer (1-4) will allow you to sequence the bolded sequence?

B.) For the sequencing gel shown, draw what the sequencing output would look like with fluorescence sequencing using the primer you chose in part A. (The fluorophores linked to detect each base are: A-red, C- green, G-blue and T-orange. If you want, you may write the names of colors instead of using colors in your answer.)



C.) Circle the true statement(s) below.

(i) ddNTPs lack 2’ OH group.
(ii) ddNTPs lack 3’ OH group
(iii) ddNTPs lack a phosphate group

Question 4, continued

D.) Circle which of the following is present in a sequencing reaction:

(i) ddNTPs (ii) RNA polymerase (iii) primase (iv) dNTPS (v) forward primer only
(vi) forward and reverse primers (vii) DNA polymerase (viii) NTPS

Question 5 (2 points)

Your four best friends have decided to help you set up a PCR reaction and put it in the thermocycler so you can watch the Red Sox game. Despite their best intentions, they have each inadvertently made modifications to the protocol below. For each friend state if they will be able to get the correct PCR product and explain your answer.

PCR reaction contains proper amounts of:
- Template DNA
- Forward primer
- Reverse primer
- Reaction buffer
- All four dNTPs
- Taq DNA polymerase

STEP		TEMP	TIME
Initial Denaturation		95°C	30 seconds
30 Cycles:	Denaturation	95°C	30 seconds
	Primer Annealing	60°C	60 seconds
	Extension	68°C	1 minute/kb
Final Extension		68°C	5 minutes

A.) Mookie used E. coli DNA polymerase instead of Taq polymerase.

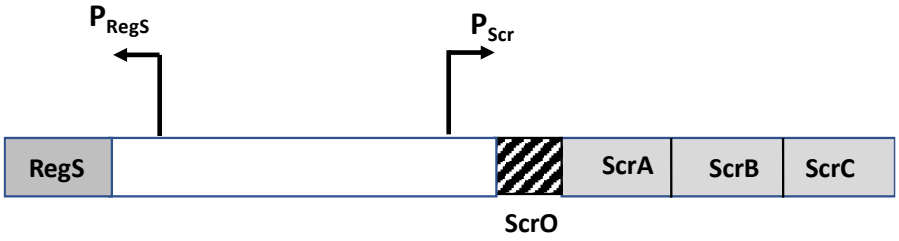
B.) Xander used 40 sec of 90^{o}C in the first step of PCR reaction instead of 30 sec.

C.) Jackie used ddNTPs instead of dNTPs in the reaction

D.) J.D. used just the forward primer instead of both forward and reverse primers.

Question 6 (3 points)
Find below the genomic region in a newly-found strain of *E. coli* that encodes genes that are important for sucrose metabolism.

RegS: encodes RegS, the transcriptional regulatory protein for the sucrose operon responsive to sucrose concentration
P_{Scr}: promoter for Scr operon that encodes enzymes to break down sucrose (ScrA, ScrB and ScrC)
P_{RegS}: promoter for the RegS gene
ScrO: operator for sucrose operon where RegS regulation occurs
ScrA: encodes enzyme ScrA
ScrB: encodes enzyme ScrB
ScrC: encodes enzyme ScrC



A.) Draw the mRNA generated for the ScrA, ScrB and ScrC genes. Label 5' and 3' ends.

B.) Do you expect the Scr operon to be inducible or constitutive?

C.) Which of the following proteins (RegS, RNA polymerase) bind the DNA elements listed below?

P_{Scr} :

P_{RegS} :

SrcO :

Consider the following table with mutants in genes responsible for sucrose metabolism.

Genotype	Sucrose breakdown activity (without Sucrose)	Sucrose breakdown activity (with Sucrose)
Wild type	low	high
$ScrA^-$	low	low
P_{Scr}^-	low	low
P_{RegS}^-	low	low
$RegS^-$	low	low

Question 6, continued

D.) Do you expect the operon to be on or off in the absence of sucrose? Explain your reasoning.

E.) What kind of transcriptional regulator is RegS, an activator or a repressor? Explain how you came to that conclusion.

F.) State if sucrose breakdown will be high or low if the ScrO region is deleted.

G.) You find a new mutant in the RegS gene, called RegS^{mut}. You observe the following results:

Genotype	Sucrose breakdown activity (without Sucrose)	Sucrose breakdown activity (with Sucrose)
RegS^{mut}	high	high

How could this mutation be affecting the RegS protein?

The Genetic Code

First letter \ Second letter	U	C	A	G	Third letter
U	UUU, UUC: Phe UUA, UUG: Leu	UCU, UCC, UCA, UCG: Ser	UAU, UAC: Tyr **UAA Stop** **UAG Stop**	UGU, UGC: Cys **UGA Stop** UGG Trp	U C A G
C	CUU, CUC, CUA, CUG: Leu	CCU, CCC, CCA, CCG: Pro	CAU, CAC: His CAA, CAG: Gln	CGU, CGC, CGA, CGG: Arg	U C A G
A	AUU, AUC, AUA: Ile **AUG Met**	ACU, ACC, ACA, ACG: Thr	AAU, AAC: Asn AAA, AAG: Lys	AGU, AGC: Ser AGA, AGG: Arg	U C A G
G	GUU, GUC, GUA, GUG: Val	GCU, GCC, GCA, GCG: Ala	GAU, GAC: Asp GAA, GAG: Glu	GGU, GGC, GGA, GGG: Gly	U C A G