

7.28 Spring 2005

Exam Three

Name _____

Question 1 ____/30 points

Question 2 ____/20 points

Question 3 ____/21 points

Question 4 ____/29 points

Total ____/100 points

Name _____

Question 1 (30 points total). You are studying the genetic control of muscle formation during mouse development. Comparisons of heart muscle vs skeletal muscle revealed several differences in key structural proteins. For example one protein, called p78 (with an apparent molecular weight of 78kDa) in skeletal muscle appears to be absent in heart muscle, and replaced with a slightly larger protein, (called p80).

To investigate the source of these differences in protein expression, you request a cDNA clone for p78 from another lab. Your first hypothesis is that p78 and p80 are encoded by alternatively spliced mRNAs from the same gene.

1A (6 points). Describe a *simple* experiment that could provide support for or against this model (no sequencing needed). You have available both heart and skeletal muscle cells in sufficient quantity to isolate RNA and/or DNA. Describe a result from this experiment that would strongly argue *against* this alternative splicing model.

Do a Northern blot: □

Isolate mRNA from both cell types using an oligo dT column (use cytoplasmic fractions to get mRNA). □

Use the p78 cDNA to make a radiolabeled probe. □

Run mRNA on a gel, transfer, and probe. □

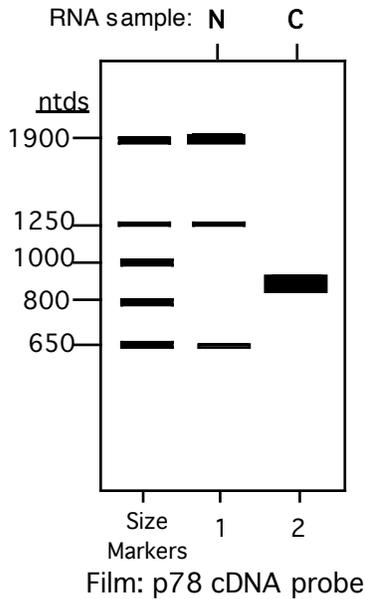
If p78 and p80 are alternatively spliced mRNAs from the same gene, you will see a band for both cell types, most likely of different sizes. □

A result that would argue strongly against alternative splicing is if the probe does not recognize anything in the heart muscle cell lane, suggesting that these two proteins are encoded by different genes. □

1B (4 points). Below is a Northern blot. The RNA samples are from skeletal muscle cells and the p78 cDNA has been labeled and used as a probe. The two samples are: nuclear RNA (N), and cytoplasmic RNA (C). Based on these data, draw a schematic of the *pre-mRNA* that encodes p78. Label the length of segments to the best of your ability.

5'CAP--UTR+exon1(650 nt)—intron(1000 nt)—exon2+UTR(250 nt)—3'polyA tail

Name _____



1C (4 points). Which of the RNA molecule(s) on this Northern blot would you expect to interact with the splicing factor U2-AF? Circle the band(s) on the figure, and briefly explain your answer.

U2-AF binds the Py tract and 3'SS. It will bind the band at 1900 nt because this represents the unspliced pre-mRNA. It may bind the band at 1250 nt because this is the intronic lariat + exon2. However, U2-AF normally is off after the first trans-esterification reaction and the structure of the RNA has changed. U2-AF may not be able to bind this structure even in an *in vitro* binding experiment.

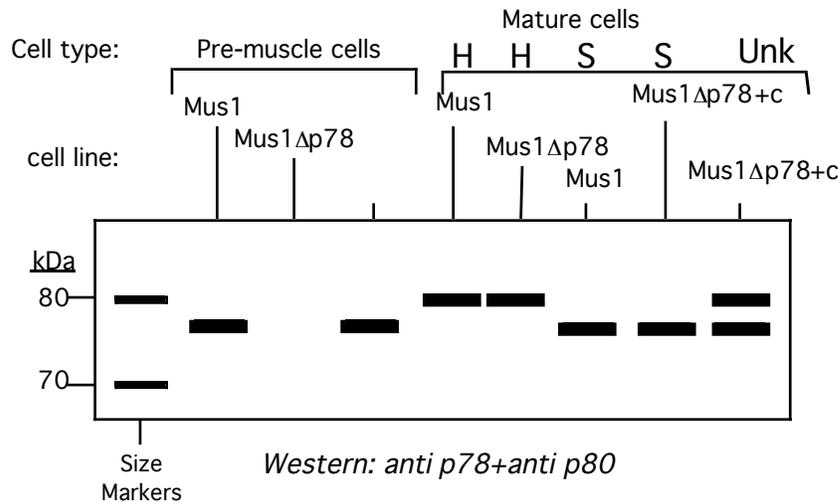
To further probe the mechanisms used to control the alternative expression of p78 vs. p80 you develop a tissue culture system for studying muscle cell development. Starting with one immature pre-muscle cell line (*Mus1*), you are able to grow both mature "heart-like" cells, which express p80 but not p78 and mature "skeletal-like" cells, which express p78 but not p80. Prior to maturing, the pre-muscle cells (*Mus1*) express the mRNA for p78 and the p78 protein, but do not express p80.

Starting with *Mus1* cells, you knock out the complete p78 gene by homologous recombination to make *Mus1Δp78*. As expected, this *Mus1Δp78* cell line can mature to form the heart-like cells, but fails to generate any mature skeletal-like cells. Next you introduce the p78 cDNA expressed from a constitutive promoter into the chromosomes of these cells to make *Mus1Δp78+c*. To your surprise, this new cell line is defective in developing into the mature heart-like cell type.

1D (4 points). Below is a Western blot of samples from the various cell lines, probed simultaneously with antibodies to p78 and p80. The samples from mature cells are labeled H for heart-like, S for skeletal-like or Unk, for unknown

Name _____

morphology. Based on these data, what is a simple explanation for the observation that the *Mus1Δp78+c* cells fail to develop properly into the mature heart-like cell type? (No model needed here—see below).



It is clear from the blot that mature heart and skeletal muscle cells only express p78 or p80 but not both. The presence of both proteins in the *Mus1Δp78+c* cells probably prevents them from following either developmental pathway and results in a mixed or unknown morphology.

1E 6 points). Propose a model to explain the changes in protein expression patterns observed in between the *Mus1* and the *Mus1Δp78+c*.

Key observations:

Only p78 is expressed in immature pre-muscle cells. (lanes 1-3)

When p78 is deleted its expression is lost (lane 2). When the p78 cDNA is added back in, the expression of p78 is restored (lane 3).

The expression of p80 seems to be turned on when cells develop into mature heart muscle cells (lanes 4 and 5). In addition, p78 expression seems to be repressed in mature heart cells—this could be due to p80.

We know that p78 and p80 proteins are encoded by different genes by lane 5—when the entire p78 gene is deleted, p80 is still expressed.

Mature skeletal cells only express p78 (lane 6). Cells that lack the p78 gene cannot mature into skeletal muscle cells (the *Mus1Δp78* cells only produce mature heart muscle cells).

***The last two lanes hold crucial information. p78 cDNA is able to rescue the development of mature skeletal cells (lane 7). This shows that the only thing you need for skeletal cell formation is a p78 protein encoded by its mRNA.

However, when the *Mus1Δp78+c* try to mature into heart cells (lane 8, we know they have begun to mature into heart cells because p80 is expressed), they are blocked because p78 expression off the cDNA cannot be repressed.

Repression of p78, then, must require introns (since these are not included in the cDNA). In addition, the cDNA has a constitutive promoter so the endogenous genomic promoter could be required for repression.

Possible Models:

1. p80 protein binds to a sequence in a p78 intron that disrupts splicing. This could lead to the inclusion of that intron which contains an in frame stop codon. The p78 protein is truncated and rapidly degraded in the cell leading to no p78 expression.
2. p80 protein turns on expression of a miRNA that is encoded within a p78 intron. When the miRNA is expressed in heart cells, it is processed, base pairs with p78 mRNA and leads to the p78 mRNA cleavage or inhibition of its translation.
3. p80 protein binds to a sequence in the p78 genomic promoter which turns off its expression (this model works although not exactly what we were looking for since this section was on splicing and non-coding RNAs, not transcription!)

1F (6 points). Design an experiment that tests an important prediction for your model. Remember, the complete genomic sequence of the mouse is available.

We gave partial credit for anything that fit the model that you proposed in part E, even if it wasn't correct.

For the models above, experiments such as testing p80 binding to p78 promoter or intron by gel shift assay would have worked.

Question 2 (20 points total). You have isolated a new strain of an *E. coli*-like bacteria from hospital patients. Upon sequencing the genome, you discover that this bacteria has a group of ~30 genes, absent from the standard lab strains, that you suspect are involved in pathogenesis. You initiate a project to understand the expression and function of these genes.

By comparing mRNA sizes with the DNA sequence, you realize that one of these genes, *pukC*, appears to contain an intron. Since bacteria lack the spliceosome components found in higher cells, you hypothesize that you have isolated a new self-splicing intron.

Sequence analysis of the *pukC* intron failed to reveal clear homology to either the group I or group 2 self-splicing intron families.

To study the mechanism of splicing you clone the *pukC* gene on to a plasmid behind a good promoter. This construct allows you to synthesize the *pukC* pre-mRNA using *in vitro* transcription. When you run the products of this first transcription experiment on a gel, you see RNA molecules with lengths suggesting that splicing is occurring (although very slowly) during the transcription reactions.

Name _____

2A (5 points). Given this result, suggest a modification of the *in vitro* reaction used above that could give a quick indication of whether the *pukC* intron uses a mechanism similar to the group I or similar to the group II introns.

Group I introns require a G nucleotide or nucleoside, which are around in the original experiment. Modify the original experiment by isolating the RNA away from the rNTPs and determine whether splicing continues to occur. If splicing still occurs, the mechanism is probably similar to Group II, if splicing no longer occurs, it could be similar to Group I.

2B (5 points). Now describe a more definitive experiment that will allow you to show that the *pukC* intron uses either a group I or group II-like mechanism. What results do you expect to observe for each of the two mechanisms?

The two splicing mechanisms have distinctly different intermediates, which can be distinguished on a gel (eg have radiolabeled rUTP in original transcription reaction so that mRNA is visible by autoradiography). Group I intermediates are linear, whereas Group II intermediates contain a lariat structure which will run differently on a gel as compared to a linear piece of mRNA of the same length.

2C (10 points). Next you mutagenize the intron sequence using a PCR-mutagenesis protocol that gives you multiple changes, on average, per intron. You sequence random clones arising from this procedure, and assay for the splicing activity of the intron using the *in vitro* assay developed above. A segment of the intron sequence from the wild type (WT) and five mutant clones (m1-m5) is shown below. Clones marked with a "+" sign are competent for *in vitro* splicing, whereas those marked with a "-" sign are defective in splicing. For each of the mutants, propose why you think the sequence changes lead to the observed splicing phenotype. Briefly explain each answer.

atgcatcatt ccaacggatt ctacctaag aatccgttg tagtcctat tccggacttt	WT +
atgcatcatt ccaacggatt GtcacctaaC aatccgttg tagtcctat tccggacttt	m1 +
atgcatcatt ccaacggatt GtcacctaaC aatccgttg Atagtcctat tccggacttt	m2 -
atgcatcatt ccaacggatt ctacctaag aatccgttg Ctagtcctat tccggacttt	m3 -
atgcatcatt ccaacggatt ctacctaag aatccgttgC Gtagtcctat tccggacttt	m4 -
atgcatcatt ccaacggatt ctacGtaaC aatccgttg tagtcctat tccggacttt	m5 -

A large putative hairpin structure is observed in the wildtype sequence encompassing sections 2 and 4 (plus two nucleotides on either side of section 3). In the m1 mutant, it is possible that the phenotype is still wildtype because the mutations maintain the hairpin structure. In the m2 mutant, these same changes

can be seen, but there is an additional A at the beginning of section 5. This is outside the area of the putative hairpin and so is probably disrupting nucleotides that are important in catalysis. The m3 mutant contains a mutation in the same spot as m2 and so again, catalytic function is probably disrupted. The m4 mutant has a mutation disrupting the large hairpin and a mutation in the putative catalytic site – either of these mutations could probably disrupt function. And lastly, in the m5 mutant, the hairpin is disrupted by the g->c mutation and the loop of the hairpin also has a mutation, either or both of which could disrupt function by disrupting secondary structure.

Question 3. (21 points). You would like to make a large number of 30 amino acid peptides of different sequences to identify peptides that bind to and inhibit an enzyme you are characterizing. Although you consider making the peptides chemically, your advisor is concerned about unwanted chemical modifications and suggests that you make the peptides using an *in vitro* translation extracts.

You initially use a plasmid construct with a promoter driving transcription of an mRNA with a 90 base open reading frame (ORF). To generate many different peptides you use a PCR-based mutagenesis technique to create random substitution mutations in the ORF. To generate many different peptides you use very strong mutagenesis conditions such that approximately one out of every three bases is mutated.

3A (3 points). When you make mRNA from the mutagenized plasmid and add it to the *in vitro* transcription reaction you find that the majority of the peptides are shorter than the 30 amino acids that you expect. Provide an explanation for this observation.

You have probably created premature STOP codons during your mutagenesis experiment.

You decide to abandon the plasmid mutagenesis approach. Instead you decide to reduce the specificity of the translation machinery. Your first approach is to target the tRNA charging machinery. Because there are 20 aminoacyl-tRNA-synthetases you do not want to have to carefully select for precise mutations to reduce their specificity.

3B (4 points). Of the three active sites in an aminoacyl-tRNA-synthetase, which would you target for mutation? Explain your choice.

The best active site to target in aminoacyl-tRNA synthetase is the amino acid hydrolysis domain/ editing pocket. You can easily mutate this site and lose specificity without worrying about disrupting the main catalytic activity of the enzyme. In contrast, if you chose to mutate the sites involved in amino acid adenylation or coupling of the amino acid to the tRNA, you would lose specificity but could easily lose catalytic activity as well.

3C (4 points). Based on your knowledge of the mechanisms used by aminoacyl-tRNA-synthetase to charge each tRNA with the correct amino acid, would you expect your proposed mutation to lead to dramatic or subtle changes in the type of amino acid incorporated at each position in the ORF. Explain your reasoning.

Mutation of the editing pocket would only cause subtle changes because the active sites involved in adenylating and coupling the amino acid to the tRNA already contain mechanisms to ensure specificity.

Although your aminoacyl-tRNA-synthetase approach is working, you want to explore a different approach. You decide you would like to reduce the specificity of codon-anticodon pairing.

Your first approach is to make a form of EF-Tu that no longer requires the factor binding center of the ribosome to stimulate its GTPase activity. Instead, the mutant EF-Tu hydrolyzes its bound GTP approximately every 20 seconds independent of ribosome binding.

Using your mutant EF-Tu in the *in vitro* translation assay, you find that the products are all 30 amino acids in length. When you analyze the sequences of the peptides you find that all the substitutions you observe can be explained by mismatches in the base pairing of the “wobble-pairing” of the anticodon-codon base pair.

3D (4 points). Based on your knowledge of the mechanisms that control the specificity of codon-anticodon recognition, propose a reason that the EF-Tu mutant never allows amino acid substitutions that are the result of mismatches in the Watson-Crick base pairs of an codon-anticodon pair. Explain your answer.

Normally, GTP hydrolysis of EF-Tu-GTP in the factor binding center is a mechanism to ensure correct anti-codon-codon pairing. In this mutant, accuracy ensured by EF-Tu is lost because it constitutively hydrolyzes its bound GTP. However, there are other mechanisms which ensure correct anti-codon-codon pairing. Accomodation, the rotation of the tRNA, also selects for correct binding because an incorrect pairing places too much strain on the tRNA, and it will be released before catalysis. The wobble-pairing base pair is able to get through accomodation even when mismatched, but the W-C base pairs must be paired correctly in order to withstand accomodation. This is why you only see peptides with substitutions resulting from mismatches of the wobble-pair but not the W-C pairs.

Name _____

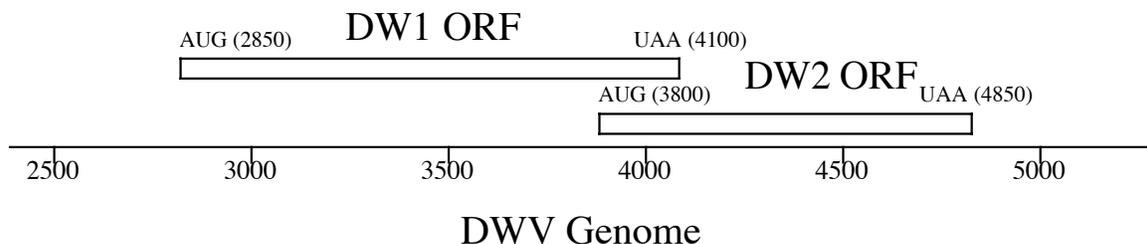
You also take a second approach, you mutate the two A-residues in the 16S rRNA that are involved in increasing the specificity of anticodon-codon pairing.

3E (6 points). You test ribosomes containing the mutant 16S rRNA in your in vitro assay and find that there is no peptide synthesis. Propose two steps in the translation elongation process that could be inhibited by the 16S rRNA mutation. Explain your reasoning.

1. □ **anti-codon-codon pairing: Additional H bonds are formed between the 2 adenines of the 16S rRNA and the minor groove of the Watson-Crick base pairs to stabilize the interaction.**
2. □ **accomodation: Rotation of the tRNA puts strain on the anti-codon-codon interaction. Stabilization of the interaction by additional H bonds ensures the efficiency of accomodation.**

Question 4 (29 points). You are studying a virus that infects and kills old growth Oak trees that you call DWV (for Deadwood Virus). Using a tissue culture model for infection, you identify two proteins (DW1 and DW2) that are strongly-expressed after the virus infects cells (remember, plants are eukaryotes!).

Suspecting these proteins are responsible for cell killing, you purify both proteins, obtain partial protein sequence, and mapped the coding sequences to two ORFs in the viral genome. Upon further investigation, you find that the DW1 ORF overlaps (but is in a different reading frame) with the DW2 ORF (see diagram below).



4A (6 Points). You decide that either there must be a RNA Pol II promoter in the DW1 ORF to drive transcription of the DW2 mRNA or the two proteins are being expressed from the same mRNA. Propose an experiment that could distinguish between these two possibilities. You have access to any radiolabeled DNA that you need.

There are several possible answers to this question – here is an example: To determine whether DW1 and DW2 are expressed from the same mRNA, do a Northern blot with labeled complementary DNA to both DW1 and DW2. Do two separate blots with the two probes – if there are two mRNAs, they should be distinguishable by size (the DW1 blot would give a band that is larger than the band

Name _____

on the DW2 blot), but if the two genes are expressed from the same mRNA, the two blots will give bands that look to be the same length.

Your findings indicate that both proteins are expressed from the same mRNA. You decide that there must be an unusual method of translational initiation occurring to drive DW2 translation.

You wonder if any ORF downstream of DW1 would be expressed. To test this you insert the LacZ gene at various positions in the mRNA and assay for LacZ activity and DW2 expression. You observe the following results:

<u>Site of LacZ Insertion*</u>	<u>LacZ Activity</u>	<u>DW2 Expression**</u>
3001	-	+++
3500	-	+++
3601	-	+++
3700	-	+++
3751	+++	-
3850	-	-

* LacZ insertions are **not** in the same reading frame as DW1 or DW2.

** Only full length DW2 protein is measured.

4B (4 Points). Based on these observations what can you conclude about the number of ORFs that can be expressed from this mRNA? What restrictions are there on the location of the internally expressed ORF?

Two ORFs can be expressed from the mRNA. The internally expressed ORF must be the first ORF downstream of nucleotides 3701-3750.

You wonder whether the ability to initiate translation internally is intrinsic to the mRNA or if it requires other viral proteins.

4C (4 points). Propose an experiment to distinguish between these two possibilities. You have the ability to transform plant cells with any portion of the viral genome. Describe the results you would expect for each possibility.

Transform the plant cells with just the portion of the genome containing DW1 and DW2 (plus promoter). Test for expression of DW2 by Western Blot. In parallel, transform the rest of the genome along with the piece of DNA containing DW1 and DW2 (to make sure that how you are doing the experiment is not somehow affecting the expression of the genes). If expression of DW2 is not dependent upon other viral

proteins, it will be expressed when transformed into the plant cell without the rest of the viral genome. If other viral proteins are needed it should only be expressed when the rest of the genome is transformed into the plant cells.

You find that at least one viral gene is required for DW2 protein synthesis. After mutating each ORF in the viral genome you find one protein (that you call DW3) that is required for DW2 synthesis.

Interestingly, DW3 has a clear RNA binding motif and you find that it binds to a 18 base hairpin sequence located at position 3710 in the viral genome.

4D (5 points). You wonder whether DW3 might directly recruit the eukaryotic small ribosomal subunit to the mRNA. Propose an experiment to test this possibility.

To test for DW3 binding to the small ribosomal subunit, you can do a ribosome binding assay. In this assay, you would radiolabel DW3 and incubate the purified protein with purified ribosome (small subunit only). Put this mixture over a gel filtration column (ribosome too big for gel shift assay), and determine whether the radiolabeled protein elutes near where the ribosome elutes from the column. If so, it would suggest that DW3 is binding to the ribosome. You must run each protein separately to determine where they elute from the column by themselves. Look for ribosome by doing a Northern Blot looking for the 16S rRNA.

You find no evidence that DW3 binds to pure 40S ribosomal subunits. In contrast, you find that if you perform the same experiment using ribosomes in a crude translation extract, you observe a clear association between DW3 and the 40S subunit.

4E (4 points). Given your knowledge of the events of eukaryotic translation initiation, why does it make sense that the purified 40S subunit is unable to associate with DW3.

The initiator tRNA must bind the ribosome before DW3 can bring the mRNA to the ribosome. This experiment would work in crude translation extract because there is initiator tRNA present.

To try and narrow down the mechanism of translation initiation for DW2, you ask whether other translation initiation factors are required for DW2 protein synthesis. You are surprised to find that both DW1 and DW2 translation absolutely require eIF4F function.

4F (6 points). Propose a model for DW3 function that can explain the requirement of eIF4F to initiate DW2 translation. Describe a simple experimental test of your model.

Name _____

eIF4F normally binds to the 5` cap of the mRNA and helps bring the mRNA to the ribosome. DW2 is an internally expressed gene and so doesn't have a 5` cap near it, but DW3 could be acting like the cap by binding the hairpin of the mRNA then DW3 could bind eIF4F and this would recruit the ribosome. There are a number of possible ways to test this – for example, you could look for a direct interaction between DW3 and eIF4F by doing a two hybrid analysis or co-immunoprecipitation experiment.