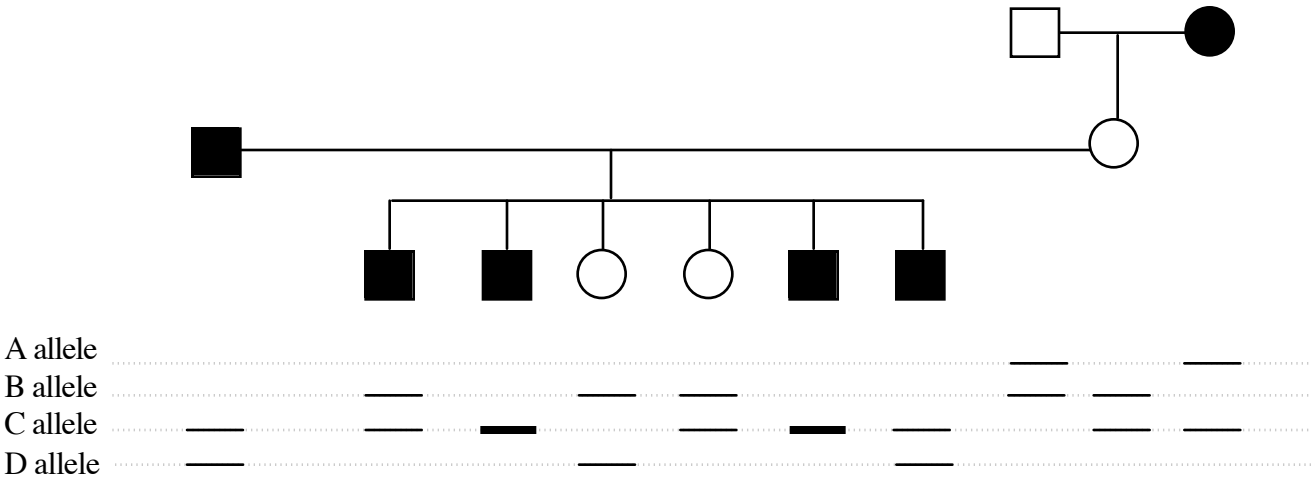
KEY

# 2005 7.03 Problem Set 7 KEY

NO DUE DATE. This problem set is to provide practice on concepts from lectures 30-36.

**1.** The following pedigree shows the inheritance of an autosomal recessive trait in a specific family. This trait is caused by a specific allele "g" at the G/g locus. You have some reason to suspect that the G/g locus is linked to an SSR on chromosome 6 called SSR41. You obtain blood samples from each member of the family, and perform a PCR reaction on the DNA of each individual that allows for the genotyping of SSR41. The results of the PCR reactions are shown below each family member in the pedigree, in a schematic of an agarose gel in which you have loaded the PCR reactions from each family member into a separate well in the gel.




| | | | | | | |
|---|---|---|---|---|---|---|
| paternally inherited allele at SSR41 | **C** | **C** | **D** | **C** | **C** | **D** |
| maternally inherited allele at SSR41 | **B** | **C** | **B** | **B** | **C** | **C** |

| | | | | | | |
|---|---|---|---|---|---|---|
| paternally inherited allele at G/g locus | **g** | **g** | **g** | **g** | **g** | **g** |
| maternally inherited allele at G/g locus | **g** | **g** | **G** | **G** | **g** | **g** |

**(a)** Fill in the tables above to indicate which alleles have been passed on to each child from their mother and father.

**(b)** Whose alleles (the mother's or the father's or both) should you follow to calculate the LOD score for the linkage of the SSR to the G/g locus?

**You want to follow the alleles of the parent who is heterozygous at both the disease locus and the SSR locus (so you can tell which of the children are parental and which are recombinant). In this case that parent is the mother.**

**(c)** Draw all possible phases for the parent(s) you listed in part **(b)**.

**G B inherited from grandpa**
**g C inherited from grandma**

**Since you know the genotypes of the mother's parents, there is only one possible phase for the mother.**

**(d)** For each phase you drew you drew in part **(c)**, state how many children are recombinants and how many children are parentals given that phase.

**GB from mother = parental**
**gC from mother = parental**
**GC from mother = recombinant**
**gB from mother = recombinant**

**Therefore, child 1 is recombinant, the other 5 are all parental.**

**(e)** Calculate the LOD score for this family at theta = 0.04 for the linkage of the SSR to the G/g locus.

**For the numerator of the odds ratio, you assume that the SSR and G/g locus are linked. Since theta =0.04, this means you'd expect 4% recombinants and 96% parentals. There are two types of recombinants, GC and gB. Each type of recombinant will occur at equal frequencies, so there is a 2% chance of any child getting GC from the mother and a 2% chance of any child getting gB from the mother. Each type of parental will occur at equal frequencies, so there is a 48% chance of any child getting GB from the mother and a 48% chance of any child getting gC from the mother.**
**Chances of the mom creating each type of gamete:**
**GB-48%**
**gC-48%**
**GC-2%**
**gB-2%**

**Child one got Bg. The chance of this is 2%.**
**Child one got Cg. The chance of this is 48%.**
**Child one got BG. The chance of this is 48%.**
**Child one got BG. The chance of this is 48%.**
**Child one got Cg. The chance of this is 48%.**
**Child one got Cg. The chance of this is 48%.**
**The chance that you saw these six kids would be 2% * 48% * 48% * 48% * 48% * 48% .**

**For the denominator of the odds ratio, you assume that the SSR and G/g locus are NOT linked. UN-linkage corresponds to theta = 0.5, because UN-linkage gives 50% parentals and 50% recombinants.**

**Since theta =0.5, this means you'd expect 50% recombinants and 50% parentals.**
**There are two types of recombinants, GC and gB. Each type of recombinant will occur at equal frequencies, so there is a 25% chance of any child getting GC from the mother and a 25% chance of any child getting gB from the mother. Each type of parental will occur at equal frequencies, so there is a 25% chance of any child getting GB from the mother and a 25% chance of any child getting gC from the mother.**
**Chances of the mom creating each type of gamete:**
**GB-25%**
**gC-25%**
**GC-25%**
**gB-25%**
**Child one got Bg. The chance of this is 25%.**
**Child one got Cg. The chance of this is 25%.**
**Child one got BG. The chance of this is 25%.**
**Child one got BG. The chance of this is 25%.**
**Child one got Cg. The chance of this is 25%.**
**Child one got Cg. The chance of this is 25%.**
**The chance that you saw these six kids would be 25% * 25% * 25% * 25% * 25% * 25% .**

**LOD = log of the odds ratio**

$$\log_{10}\frac{(0.48)^5(0.02)^1}{(0.25)^6} = 0.32$$

**(f)** At what theta value would you achieve the maximal LOD score for this family, knowing everything you know about them?

**You'd expect to achieve the maximal LOD score when theta = the fraction of recombinant children that you actually have in your family. In this case 1 in 6 children were recombinant so you'd expect the maximum LOD score when theta = 1/6 = 0.17.**

**(g)** What is the LOD score value for the theta value you listed in part **(f)**?

**If theta = 0.17, you'd expect 83% total parentals and 17% total recombinants.**

$$\log_{10}\frac{(0.42)^5(0.08)^1}{(0.25)^6} = 0.63$$

**(h)** If you had never seen the genotyping results for this family, and only had their pedigree available, what would have been the theoretical maximum LOD score value that you could have ever calculated for this family? (**Hint:** Start by thinking about which theta value could give you the maximum possible LOD score.)

**The theoretical maximum LOD score for an SSR and a trait locus from a 6-child family in which there is one relevant parent whose phase you do indeed know occurs if all children are parentals. If all children are parentals, then the optimal LOD score value would occur at a theta = 0. Thus the theoretical maximum is:**

$$\log_{10}\frac{(0.5)^6}{(0.25)^6} = 1.8$$

**(i)** If you had never seen the genotyping results for this family, and only had their pedigree available, what is the minimum number of kids that the family would have had to have contained in order to reach a theoretical maximum LOD score that is > 3?

$$\log_{10}\frac{(0.5)^x}{(0.25)^x} \geq 3$$ **Solving for x, x = at least 10 children.**

**2.** A tumor results when a cell in the body loses control over cell growth and division such that the cell divides many times, forming a ball of cells. Cancer can be extremely harmful to the organism when these balls of cells either physically interfere with function of an essential organ, or begin to steal the nutrients away from cells of essential organs. Cells become capable of growing and dividing inappropriately when they have accumulated multiple mutations in genes (such as oncogenes and tumor suppressor genes) whose normal functions are to control cell growth and division (i.e. to control the cell cycle).

**(a)** Why is the notion of there being "a cure for cancer" unreasonable?

**As we discussed in class, we use the word "cancer" to describe a variety of diseases with a variety of causes. Each different specific type of cancer is caused by an accumulation of roughly 6 different mutations in a collection of oncogenes and tumor suppressor genes. No two types of cancer are caused by the same collection of 6 mutations. Therefore the notion that there will be one cure that will be effective for all types of cancers is unreasonable.**

**(b)** What is the wild-type function of an oncogene?

**The wild-type function of an oncogene is to promote cell growth and cell division when it is an appropriate (and in the appropriate cell types).**

**(c)** What phenotype may result if an oncogene gets mutated so that it becomes over-active?

**If an oncogene gets mutated so that it becomes over-active, you'd expect the phenotype of uncontrolled growth. A wild-type oncogene stimulates the cell cycle whever appropriate, so an overactive mutant version of an oncogene would stimulate the cell cycle even when not appropriate. Cancer is the result of uncontrolled cell proliferation, so stimulating the cell cycle at inappropriate times leads to cancer.**

**(d)** Would an over-active allele of an oncogene cause a dominant or a recessive phenotype?

**An over-active allele of an oncogene would cause a dominant phenotype. If you have a cell that has one overactive mutant allele and one wild-type allele that is properly regulated, the sum total activity that you would see would be overactivity. Adding back a wild-type version of an oncogene does not help a cell whose problem is that it has a mutant overactive version of an oncogene.**

**(e)** Would an over-active allele of an oncogene be the result of a loss-of-function mutation or a gain-of-function mutation?

**An over-active allele of an oncogene would be the result of a gain of function mutation. The mutated oncogene has gained the ability to stimulate the cell cycle all of the time (when really it is supposed to only stimulate the cell cycle when appropriate, i.e. when the cell receives growth signals from the environment).**

**(f)** Could a mutation in an oncogene (that caused the gene to become over-active) have occurred in the cis regulatory regions of the gene? If so, give an example of how a change in the cis regulatory regions of an oncogene could lead to an over-active mutant allele of an oncogene.

**Yes. One example would be if the cis-acting region of an oncogene was no longer able to bind to a transcriptional repressor. In this case, the repressor would no longer regulate the oncogene, so the oncogene would be transcribed constitutively.**

**(g)** What is the wild-type function of a tumor-suppressor gene?

**The wild-type function of a tumor-suppressor gene is to inhibit cell growth and division whenever it is not appropriate for cells to be going through the cell cycle. (Some tumor suppressor genes' wild-type roles are to promote apoptosis, or programmed cell death.)**

**(h)** What phenotype may result if a tumor-suppressor gene is mutated so that it no longer functions?

**If a tumor suppressor gene gets mutated so that it becomes inactive, you'd expect the phenotype of uncontrolled growth. A wild-type tumor suppressor gene inhibits the cell cycle whever appropriate, so an inactive mutant version of a tumor suppressor gene would allow cells to go through the cell cycle even when not appropriate. Cancer is the result of uncontrolled cell proliferation, so allowing the cell cycle to progress at inappropriate times leads to cancer.**

**(ALTERNATELY, FOR OTHER KINDS OF TUMOR SUPPRESSOR GENES: If a tumor suppressor gene gets mutated so that it becomes inactive, you'd expect the phenotype of uncontrolled growth. A wild-type tumor suppressor gene promotes apoptosis whever appropriate, so an inactive mutant version of a tumor suppressor gene would prevent cells from dying even when they should be dying. Cancer is the result of uncontrolled cell proliferation, so allowing cells to survive and proliferate even when they are supposed to die leads to cancer.)**

**(i)** Would the inactivation of a tumor suppressor gene cause a dominant or a recessive phenotype?

**The inactivation of a tumor suppressor gene would cause a recessive phenotype. If you have a cell that has one inactive mutant allele and one wild-type allele that is properly regulated, the sum total activity that you would see would be normal activity. Adding back a wild-type version of a tumor suppressor gene does help a cell whose problem is that it has a mutant inactive version of a tumor suppressor gene.**

**(j)** Would the inactivation of a tumor suppressor gene be the result of a loss-of-function mutation or a gain-of-function mutation?

**The inactivation of a tumor suppressor gene would be the result of a loss-of-function mutation. The mutated tumor suppressor gene has lost the ability to inhibit the cell cycle.**

**(k)** Could a mutation in a tumor suppressor gene (that caused the gene to become inactive) have occurred in the coding region of the gene? If so, give an example of how a change in the coding region of a tumor suppressor gene could lead to an inactive mutant allele of a tumor suppressor gene.

**Yes, a premature stop codon (nonsense mutation) could lead to a non-functional and therefore inactive tumor suppressor gene. A missense mutation that led to a change in an amino acid necessary for the tumor suppressor protein to perform its function would also lead to an inactive tumor suppressor gene. Finally, a frameshift mutation could cause the reading frame to shift, thus changing all codons and potentially exposing a premature stop codon that is normally not read.**

**(l)** Could a mutation in a tumor suppressor gene (that caused the gene to become inactive) have occurred in the cis regulatory regions of the gene? If so, give an example of how a change in the cis regulatory regions of a tumor suppressor gene could lead to an inactive mutant allele of a tumor suppressor gene.

**Yes, one example would be if the promoter of a tumor suppressor gene was no longer able to bind to a transcriptional activator or to RNA polymerase. In this case, the activator would no longer be able to recruit RNA polymerase, or RNA polymerase would not be able to bind, leading to a lack of transcription of the tumor suppressor gene's mRNA.**

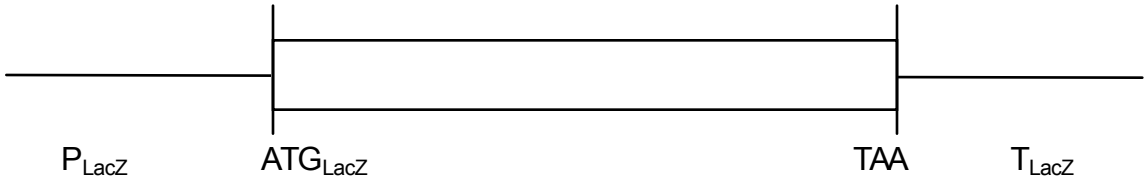
**3.** You are studying a yeast gene (Act1) which encodes a transcriptional activator protein; Act1 activates the Yst1 gene, which encodes a yeast enzyme that helps yeast cells deal with high salt conditions. Act1 is normally transcribed only when yeast cells are grown in high salt concentrations. You create two different DNA constructs that will allow you to visualize when and/or where Act1 is expressed in cells. Each DNA construct is a fusion of part of the Act1 gene to part of the *E. coli* LacZ gene. You make these two fusion constructs because you want to visualize when and/or where Act1 is expressed in yeast cells, but you don't have a good assay for measuring the presence or activity of Act1 protein. You do, however, have a good assay for measuring the presence and activity of *E. coli* beta-galactosidase, because you know that this enzyme cleaves X-gal and releases a blue-colored compound.

The first DNA construct you make is called an "Act1-LacZ transcriptional fusion." To make this construct, you fuse the cis regulatory region ("$P_{Act1}$") that lies upstream of the Act1 open reading frame to the LacZ coding sequence and terminator. You then place this hybrid gene on a yeast plasmid.

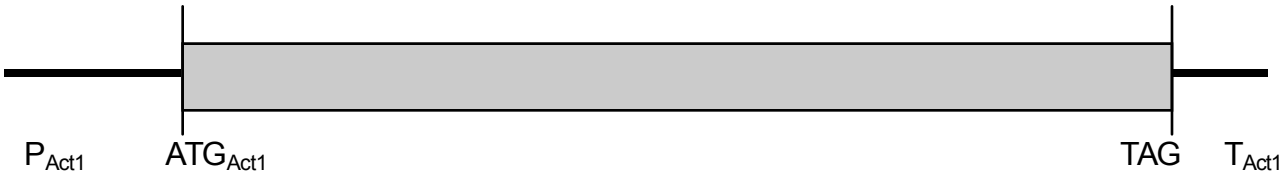
The second DNA construct you make is called an "Act1-LacZ translational fusion." To make this construct, you fuse almost the entire Act1 gene (beginning with its promoter and ending right before its stop codon) directly upstream of a portion of the LacZ gene (from the start codon through the terminator). You then place this hybrid gene on a yeast plasmid.

The beta-galactosidase enzyme (which is encoded by the lacZ gene) is found in the cytoplasm of *E. coli* bacterial cells. When beta-galactosidase is expressed in yeast cells, it is also found in the cytoplasm. The transcriptional activator protein Act1 is found in the nucleus of yeast cells.

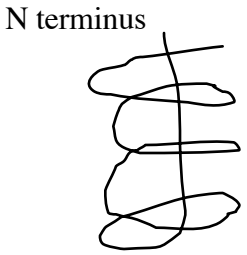
The gene for beta-galactosidase (which is *lacZ*) looks like this: (T = transcription terminator)




The gene for Act1 looks like this:



The gene for beta-galactosidase produces a protein that looks like this:




The gene for Act1 produces a protein that looks like this:

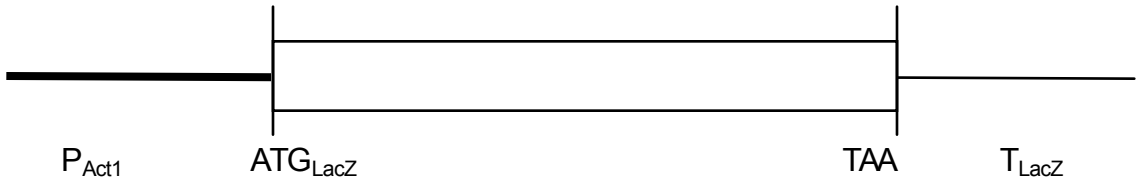

**(a)** Why does it make sense that the transcriptional activator protein Act1 is found in the nucleus of yeast cells?

**In eukaryotes, transcription happens in the nucleus, therefore it makes sense that a transcriptional activator would also be found in the nucleus. The job of a transcriptional activator protein is to bind to DNA and activate transcription. DNA lives in the nucleus of eukaryotic cells, so that is where activators are found in the cell.**

**(b)** What might be the functions of the two different protein domains possessed by the Act1 protein?

**The job of a transcriptional activator protein is to bind to DNA and activate transcription. The way that activator proteins activate transcription is by recruiting the transcriptional machinery (RNA polymerase and its associated factors). Thus activator proteins have two different domains: 1) the DNA binding domain (the transcriptional activator must bind to DNA in order to activate transcription) and 2) the activation domain that recruits the transcriptional machinery.**

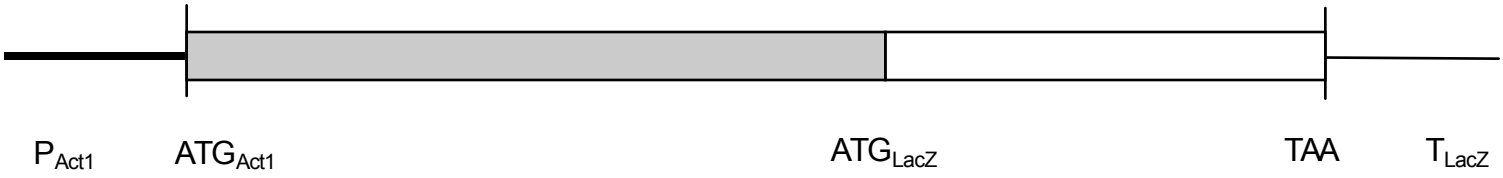
**(c)** Based on the diagrams above, draw a schematic of the DNA construct that would result when you made the transcriptional Act1-LacZ fusion.



**FROM THE INTRODUCTION TO THE QUESTION: To make this construct, you fuse the cis regulatory region ("$P_{Act1}$") that lies upstream of the Act1 open reading frame to the LacZ coding sequence and terminator.**

**(d)** Based on the diagrams above, draw a schematic of the DNA construct that would result when you made the translational Act1-LacZ fusion.

**(The scale of this drawing is 70% so that it fits on the page.)**




**FROM THE INTRODUCTION TO THE QUESTION: To make this construct, you fuse almost the entire Act1 gene (beginning with its promoter and ending right before its stop codon) directly upstream of a portion of the LacZ gene (from the start codon through the terminator).**

**(e)** Under what cellular conditions is Act1 normally transcribed and translated?

**Act1 is expressed under the condition of high salt in yeast cells. This makes sense, because the job of Act1 is to activate genes that are necessary for yeast cells to deal with high salt conditions (osmotic stress).**

**(f)** Under what cellular conditions is beta-galactosidase normally transcribed and translated?

**LacZ is expressed when lactose is present AND when glucose is NOT present in *E. coli* cells. This makes sense, because the job of the LacZ gene product is to metabolize lactose for energy.**

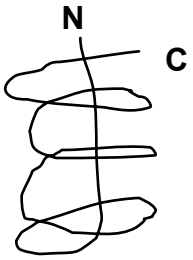
**(g)** Under what conditions would the protein be made that is produced from the transcriptional Act1-LacZ fusion?

**Since the expression of this construct is controlled by the Act1 promoter, you'd expect the product of this gene to be made under the condition of high salt.**

**(h)** Under what conditions would the protein be made that is produced from the translational Act1-LacZ fusion?

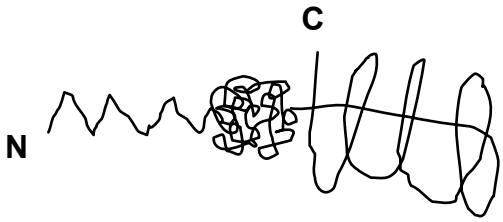
**Since the expression of this construct is controlled by the Act1 promoter, you'd expect the product of this gene to be made under the condition of high salt.**

**(i)** Based on the diagrams above, draw the protein that would be produced from the transcriptional Act1-LacZ fusion.

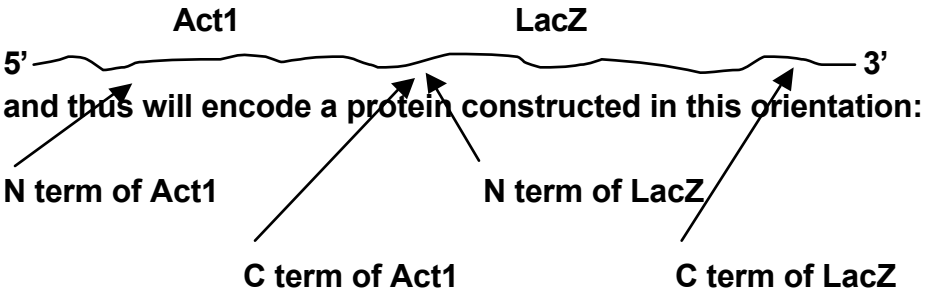



**The open reading frame in this reporter gene construct is LacZ. Thus the protein produced by this construct will be identical to LacZ.**

**(j)** Based on the diagrams above, draw the protein that would be produced from the translational Act1-LacZ fusion.

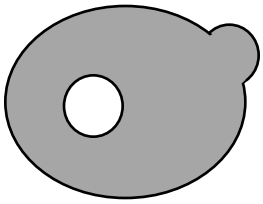



**The open reading frame in this reporter gene construct is a fusion of Act1 and LacZ. Thus the protein produced by this construct will be the Act1 protein fused (i.e. covalently linked) to the LacZ protein.**
**PLEASE NOTE that the directionality with which you draw this fusion protein is very important. ***The 5' end of an mRNA encodes the N terminus of a protein. ***
**The mRNA produced from this fusion will look like:**



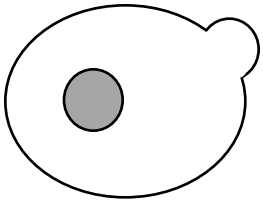
**(k)** Below is drawn a budding yeast cell with a nucleus inside the cell. Shade in where the beta-galactosidase enzyme would be found in a yeast cell expressing the transcriptional Act1-LacZ fusion. (To answer parts **(k)** and **(l)**, you must know that most signals used by the cell to direct intracellular protein localization are found at the very N terminus of the protein.)

**LacZ will be expressed in the cytoplasm only, since beta-gal is normally used in the cytoplasm and the signal for localizing the beta-gal would still be intact and at the N terminus of the protein produced from the transcriptional fusion. (Most signals used by the cell to direct intracellular protein localization are found at the very N terminus of the protein.)**



**(l)** Below is drawn a budding yeast cell with a nucleus inside the cell. Shade in where the beta-galactosidase enzyme would be found in a yeast cell expressing the translational Act1-LacZ fusion.

**In the nucleus only, since Act1 is normally used in the nucleus and the signal for localizing Act1 is at the very N terminus of the protein produced from this translational fusion. (Most signals used by the cell to direct intracellular protein localization are found at the very N terminus of the protein.)**



**(m)** Say that you had a haploid yeast strain that had the endogenous chromosomal copy of Act1 deleted. If you transformed the act1$^{-}$ haploid yeast strain with the transcriptional Act1-LacZ fusion plasmid, would the strain be able to cleave the compound X-gal?

**Yes. Under conditions when Act1 would normally be produced (high salt), now beta-gal would be produced and would be able to cleave X-gal.**

**(n)** If you transformed the act1$^{-}$ haploid yeast strain with the transcriptional Act1-LacZ fusion plasmid, would the strain be able to induce Yst1 gene expression?

**No. There is no Act1 protein present in these yeast cells, because the chromosomal copy of Act1 has been deleted, and no Act1 protein is produced from the transcriptional fusion. If there is no Act1 protein produced, then there is nothing to activate the Yst1 gene.**

**(o)** If you transformed the act1$^-$ haploid yeast strain with the translational Act1-LacZ fusion plasmid, would the strain be able to cleave the compound X-gal?

**It is possible, yes. LacZ is being produced in these cells, even though the LacZ being produced is fused to Act1. For LacZ to function in the cell, it must be true that having the Act1 protein fused to the LacZ protein doesn't affect the functionality of LacZ. Experimentally, we know that LacZ can function just fine EVEN when it are covalently tethered to other proteins. This may seem counterintuitive, but we know it to be true imperically. NOTE that there is an additional issue involved in answering this question, which is that LacZ protein will now be in the nucleus of the yeast cells. It is also only true that X-gal can be cleaved if X-gal is capable of gaining access to the nucleus. (X-gal normally only needs to gain access to the cytoplasm in normal cells that are expressing LacZ where it should reside, which is in the cytoplasm.)**

**(p)** If you transformed the act1$^-$ haploid yeast strain with the translational Act1-LacZ fusion plasmid, would the strain be able to induce Yst1 gene expression?

**It is possible, yes. Act1 is being produced in these cells, even though the Act1 being produced is fused to LacZ. For Act1 to function in the cell, it must be true that having the beta-gal protein fused to the Act-1 protein doesn't affect the functionality of Act1. Experimentally, we know that many proteins can function just fine EVEN when they are covalently tethered to other proteins. This may seem counterintuitive, but we know it to be true imperically.**