

Biochemistry/ MCB 400 Instructor: Hardison
Final Examination
Fall 1995
December 15, 1995

This examination has 37 questions for a total of 200 points. The first 36 are multiple choice and are worth 5 points each. Please answer these on the enclosed answer sheet. **BE SURE TO ENCODE YOUR STUDENT NUMBER AND TEST FORM ON THE ANSWER SHEET!**

THIS IS FORM A, and has answers.

The final question requires you to write answers on the exam. I will grade this and encode the score on the answer sheet, which will then be machine graded.

PLEASE TURN IN YOUR WHOLE EXAM AND ANSWER SHEET. Put your name and student number below. The exams can be picked up outside 206 Althouse Lab after they are graded. Good luck!

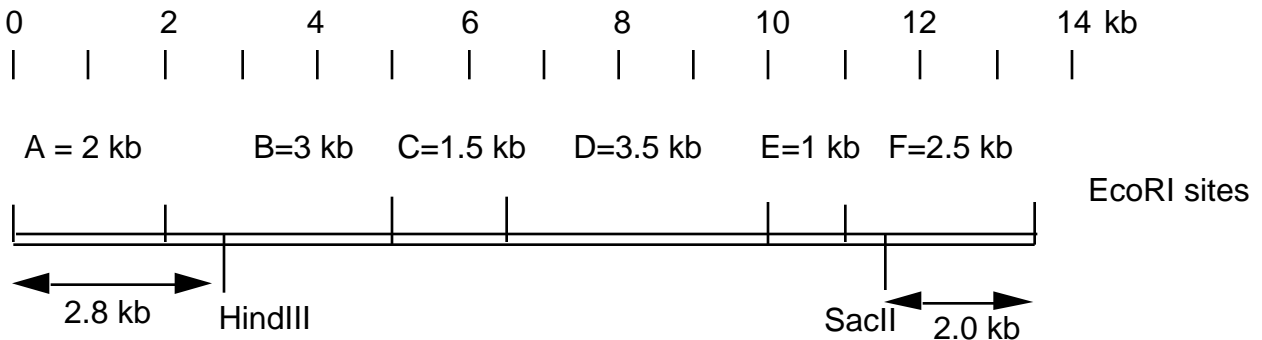
Name: _____

Student Number: _____

For a series of 25 problems, imagine the following scenario. After your graduation from Penn State, you are working for Amgen, as part of a group of scientists who want to understand expression of the OB gene in humans. Loss-of-function mutations in the homolog of this gene in mice cause obesity, and it is likely that the OB protein is involved in regulation of fat metabolism. You, your colleagues, the management and the shareholders of Amgen figure that there are big profits to be made in the exploitation of this protein. I have made up a restriction map of this region, from which you should be able to map the gene and various cis-acting regulatory sites. A copy of this map is on one of the back pages in the exam, as well as other useful information, like the genetic code, binding sites for a few transcription factors, and some equations. I suggest that you (neatly) tear off these last 2 sheets to use throughout the exam; that will give you some continuity as you solve these problems.

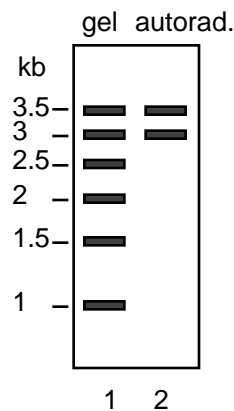
The map of EcoRI (above the line), HindIII and SacII sites in the region of the human OB gene is given below.

Restriction map of the cloned genomic DNA containing the OB locus



Problems 1-3 concern mapping the exons of the OB gene. These were mapped using two different approaches. First, an EcoRI digest of the cloned genomic OB locus was separated on an agarose gel, blotted onto a nylon membrane, and probed with a labeled OB cDNA. The pattern of all the EcoRI fragments from the OB locus is shown in lane 1, and the autoradiograph of the probed membrane is shown in lane 2.

Digest with EcoRI, probe with OB cDNA:

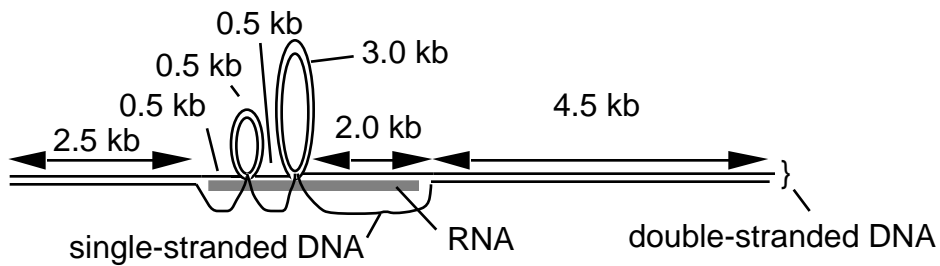


A1 = B12. (5 pts) Which EcoRI fragments contain exons?

- a. A, B, C, D, E and F
- b. B and D
- c. A and B
- d. E and F

b is correct

In the other approach, R-loops were formed between the genomic OB DNA and OB mRNA and visualized in the electron microscope. An interpretation of the results is shown below.



To establish an orientation for the R-loops relative to the restriction map, the genomic double-stranded DNA was digested with *Sac*II prior to forming the R-loops. The resulting structure looked like that above, except that the segment of double-stranded DNA that is 4.5 kb in the figure above was shortened to 2.5 kb.

A2=B13. (5 pts) How many exons are present in the OB gene?

- a. two
- b. two in one gene; one in an adjacent gene
- c. three
- d. four

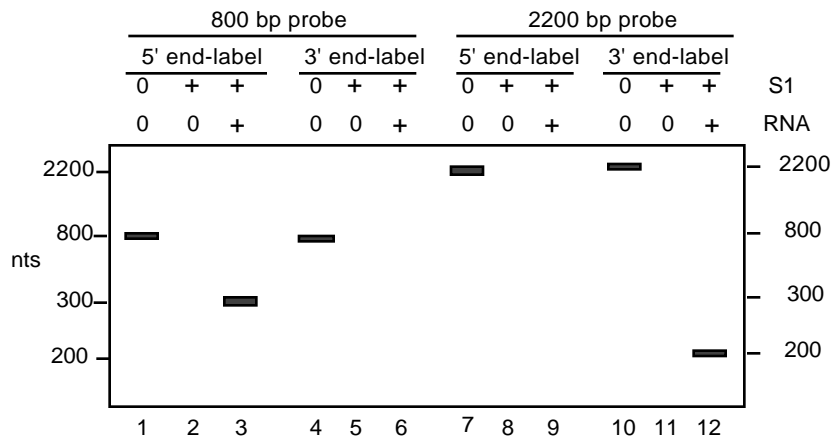
c is correct

A3=B14. (5 pts) Which of the following statements is correct about the relationship between exons and *Eco*RI fragments of the OB gene?

- a. Two exons are in fragment B, one is in fragment D.
- b. One exon is in fragment E, one exon is in fragment D, and another covers fragments B and C.
- c. One exon is in fragment B and one exon is in fragment C.
- d. One exon is in fragment E, two are in fragment D and one is in fragment B.

a is correct

The 5' end of the OB gene was examined more precisely using an nuclease S1-protection assay. The clone of genomic DNA was digested with HindIII (which is located 2.8 kb from the left side of the restriction map), radiolabeled on the 5' end using [³²P]ATP and polynucleotide kinase, and digested with EcoRI. An 800 bp fragment extending from the EcoRI site at the A/B junction to the 5' end-labeled HindIII site was isolated, as was a 2200 bp fragment extending from the 5' end-labeled HindIII site to the EcoRI site at the B/C junction. In a parallel experiment, the same DNA fragments were labeled on the 3' end at the HindIII site. The labeled DNA fragments were denatured, mixed with RNA from adipocytes (which actively express the OB gene), and RNA-DNA heteroduplexes were formed by annealing the labeled DNA and the RNA. The heteroduplexes were treated with nuclease S1, and the protected fragments run on a denaturing polyacrylamide gel and visualized by autoradiography. In the figure below, the rows above the autoradiograph indicate whether RNA was present in the annealing mixture, and whether the sample was treated with nuclease S1. Lanes 3, 6, 9 and 12 show the results of treating the RNA-DNA heteroduplexes with nuclease S1.



A4=B15. (5 pts) Relative to the restriction map of the OB locus, does the direction of transcription of the gene(s) correspond to left to right or right to left?

- There is one gene, transcribed from left to right.
- There is one gene, transcribed from right to left.
- There are two genes, one transcribed right to left and one transcribed left to right.
- There are two genes, both transcribed left to right.

a is correct

A5=B16. (5 pts) What do the results using the 800 bp probe tell you?

- The 5' end of the OB gene is located 0.3 kb from the left of the restriction map.
- The 5' end of the OB gene is located 3.1 kb from the left of the restriction map.
- The 5' end of the OB gene is located 2.5 kb from the left of the restriction map.
- The 3' end of the OB gene is located 2.8 kb from the left of the restriction map.

c is correct

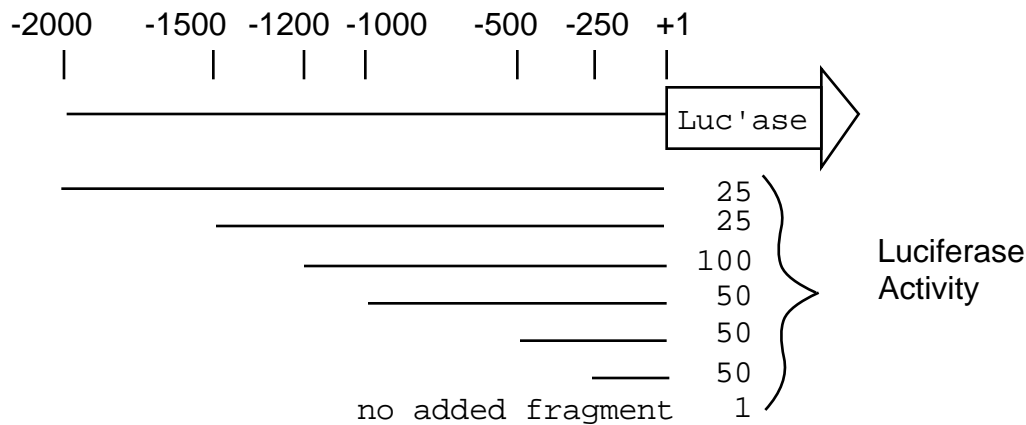
A6=B17. (5 pts) What do the results using the 2200 bp probe tell you?

- The 3' end of the OB gene is located 3.0 kb from the left of the restriction map.
- The 3' end of the OB gene is located 2.8 kb from the left of the restriction map.
- The 5' end of the OB gene is located 5.0 kb from the left of the restriction map.

d. The 3' end of the first exon of the OB gene is located 3.0 kb from the left of the restriction map.

d is correct

After determining the 5' end of the OB gene, a DNA fragment that extends 2000 bp 5' to the cap site was isolated and placed upstream of the gene encoding luciferase on a plasmid. This new construct gave activity when transferred into adipocytes (a process called transfection), as indicated in the figure below. To better localize *cis*-acting control regions, this 2000 bp DNA fragment driving expression of the luciferase was progressively deleted from the 5' end. The effects on the level of expression after transfection into adipocytes is shown below.



A7=B18 (5 pts) Does one of these DNA segments appear to have a proximal promoter activity?

- a. No.
- b. Yes, it is in the -2000 to -1500 segment.
- c. Yes, it is in the -1200 to -1000 segment.
- d. Yes, it is in the -250 to +1 segment.

d is correct

A8=B19. (5 pts) Do any of the upstream segments have a positive regulatory effect?

- a. Yes, the -1200 to -1000 segment does.
- b. Yes, the -1500 to -1200 segment does.
- c. Yes, both the -1500 to -1200 segment and the -1000 to -5000 segment do.
- d. No.

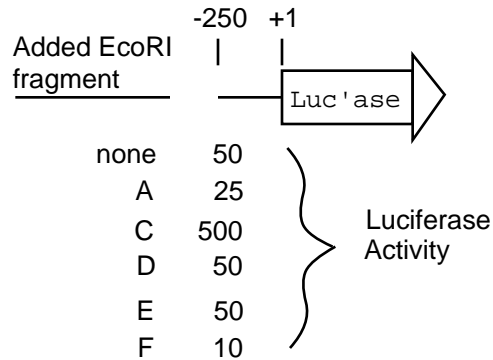
a is correct

A9=B20. (5 pts) Do any of the upstream segments have a negative regulatory effect?

- a. Yes, the -1200 to -1000 segment does.
- b. Yes, the -1500 to -1200 segment does.
- c. Yes, both the -2000 to -1500 segment and the -500 to -250 segment do.
- d. No.

b is correct

The search for other *cis*-acting regulatory elements was extended through the rest of the OB region by adding EcoRI fragments at the 5' end of a construct with 250 bp of OB promoter driving expression of the luciferase gene. The effects of these added fragments on the level of expression in transfected adipocytes is shown below.



A10=B21. (5 pts) Do any of the EcoRI fragments have a negative effect on expression from the OB promoter?

- Yes, only fragment A.
- Yes, both fragments A and F.
- Yes, only fragment C.
- No.

b is correct

A11=B22. (5 pts) Which EcoRI fragment(s) has (have) a positive effect, and where is (are) these fragments relative to the gene? Note: "5' to the gene" means that the element is located 5' to the cap site on the nontemplate strand (top or message-synonymous strand); this is also referred to as "upstream."

- Fragment A has a positive effect and it is 5' to the gene.
- Fragment C has a positive effect and it is 3' to the gene.
- Fragment C has a positive effect and it is in an intron of the gene.
- Fragments D and E have positive effects, D has an exon and E is 3' to the gene.

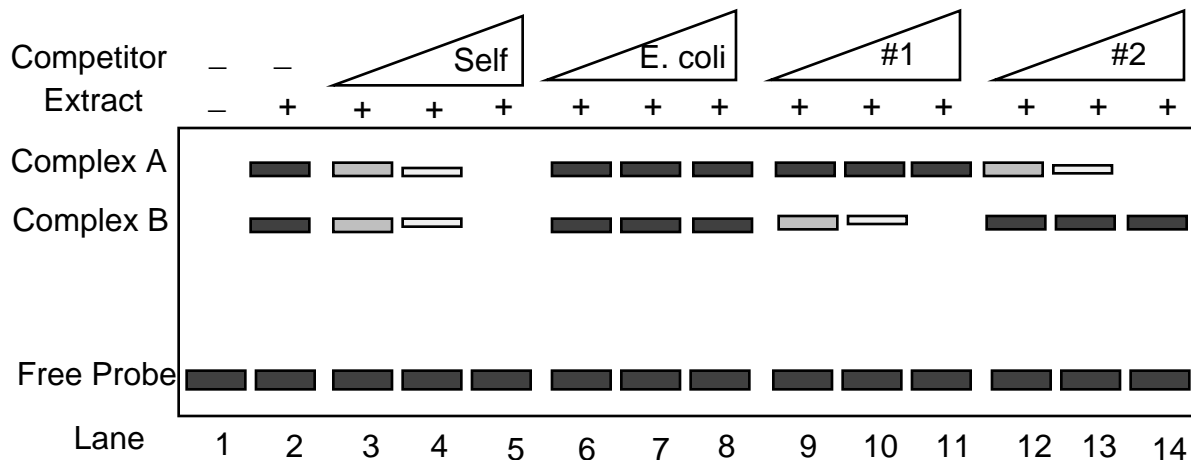
c is correct

A12=B23. (5 pts) If you postulated that EcoRI fragment(s) with a positive effect contained an enhancer, how would you test this?

- Delete the fragment from the construct; then assay whether the level of expression changed.
- Add multiple copies of the fragment to the construct; then assay whether the level of expression changed.
- Flip the fragment to the opposite orientation and, in a separate experiment, move it to the other side of the reporter gene in the construct; then assay whether the level of expression changed.
- Assay for alterations in chromatin structure in the vicinity of the construct in the transfected cells.

c is correct

In an effort to determine what types of proteins might be binding to the positive regulatory element, a short segment of duplex DNA from this region was synthesized (we'll call this the self oligonucleotide), end-labeled to make a probe, mixed with an extract containing adipocyte nuclear proteins and analyzed by an electrophoretic mobility shift assay. Lane 1 is the probe alone, and lane 2 shows the complexes formed by adipocyte nuclear proteins binding to the probe. The sequence of the top strand of the self oligonucleotide is given in the figure below. Lanes 3-14 show the results of adding increasing amounts of cold DNA (E. coli DNA or particular duplex oligonucleotides) in excess to the binding reaction as competitors. Oligonucleotides #1 and #2 have altered the sequence of individual motifs in the self oligonucleotide, as shown by the underlining in the figure.



Top strand of self oligonucleotide: TGCACCTATTTGCATCGTGACTCAGGCATG

Top strand of oligonucleotide #1: TGCACCTATTTGCATCGAGTAAGTGGCATG

Top strand of oligonucleotide #2: TGCACCTGCTGCAGTCGTGACTCAGGCATG

A13=B24. (5 pts) What type of protein is generating complex A? The protein has a similar binding site to that of:

- a. Oct1
- b. Sp1
- c. GAL4
- d. AP1

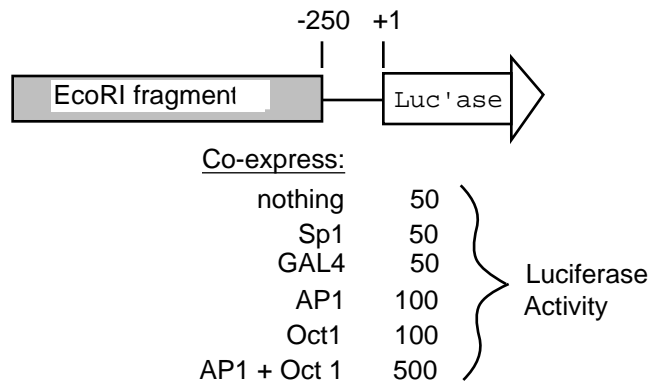
d is correct

A14=B25. (5 pts) What type of protein is generating complex B? The protein has a similar binding site to that of:

- a. Oct1
- b. Sp1
- c. Glucocorticoid receptor
- d. AP1

a is correct

The role of these proteins in regulation of the OB gene was examined more directly by co-expression assays in cultured insect cells. For this problem, assume that insect cells have none of the proteins that activate the OB gene (such as those detected in the electrophoretic mobility shift assay), but they are capable of basal transcription of the OB gene. In this reporter construct, an EcoRI fragment with a positive regulatory element has been placed upstream of the OB promoter, which is driving expression of the luciferase gene. The insect cells carrying this reporter construct were additionally transfected with plasmids that direct the synthesis of several transcription factors, as listed below. The effects of expressing these factors on the level of luciferase activity are given in the figure below.



A15=B26. (5 pts) What information is provided by the co-expression data?

- GAL4 negatively regulates expression from the OB promoter.
- Oct1 and AP1 interact synergistically to increase expression from the OB promoter.
- Oct1 negatively regulates expression until it can form a complex with AP1, and that Oct1-AP1 complex positively regulates the OB promoter.
- Sp1 represses activity until it is in a complex with Oct1, and Oct1-Sp1 will increase expression from the OB promoter.

b is correct

A16=B27. (5 pts) The sequence from part of one of the exons on the nontemplate (also called "top" or message-synonymous) strand is

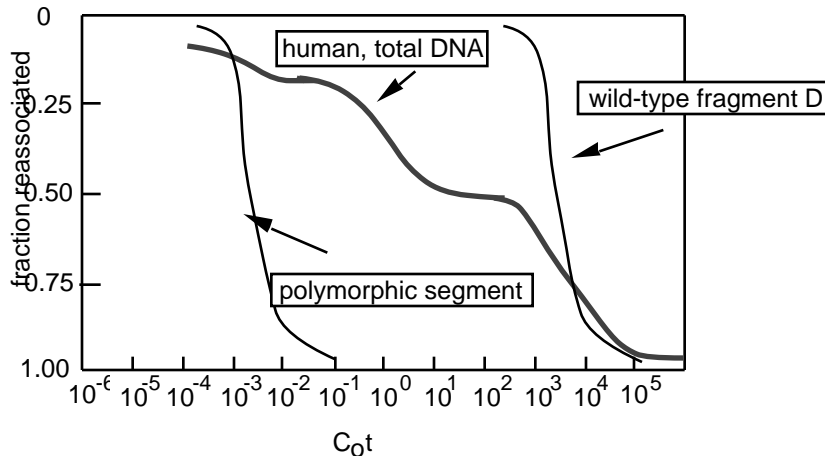
5' ... GGTTGATTCACGAATAACG ... 3'

You know that the protein-coding portion of the gene extends through this region, i.e. there are no stop codons in the correct reading frame for this segment. What is the sequence of amino acids encoded by this segment? Use the "universal" genetic code for your answer.

- Val-Asp-Ser-Arg-Ile-Thr
- Val-Ile-Arg-Glu-Ser-Thr
- Leu-Phe-Val-Asn-Gln-Pro
- Leu-Ile-His-Glu-Gly-Asn

a is correct. Find the stop codons and translate the frame without stop codons.

In screening through families with a strong genetic predisposition to obesity, searching for large changes in the region of the OB gene, you found one individual whose fragment D is 3.8 kb rather than the usual 3.5 kb. The DNA segment that accounted for this difference in size was isolated and labeled, and the kinetics with which it annealed to total genomic DNA was measured (indicated as "polymorphic segment" below). For comparison, the profile for human total genomic DNA reannealing to itself is shown as the thicker line with several inflections, and the results of using the wild-type fragment D as the labeled probe annealing to human genomic DNA is also shown below.



A17=B30. (5 pts) To the nearest order of magnitude (or power of 10), what is the repetition frequency of the polymorphic segment?

- a. 10⁶
- b. 10⁴
- c. 10²
- d. 1

a is correct

A18=B31. (5 pts) Which of the following is the most likely cause of this polymorphism?

- a. An L1 repeat inserted in an enhancer located 3' to the OB gene.
- b. A single nucleotide substitution altered the sequence of the encoded protein.
- c. An Alu repeat inserted in an exon, altering the sequence of the encoded protein.
- d. A deletion between direct repeats causes a loss of an enhancer in an intron.

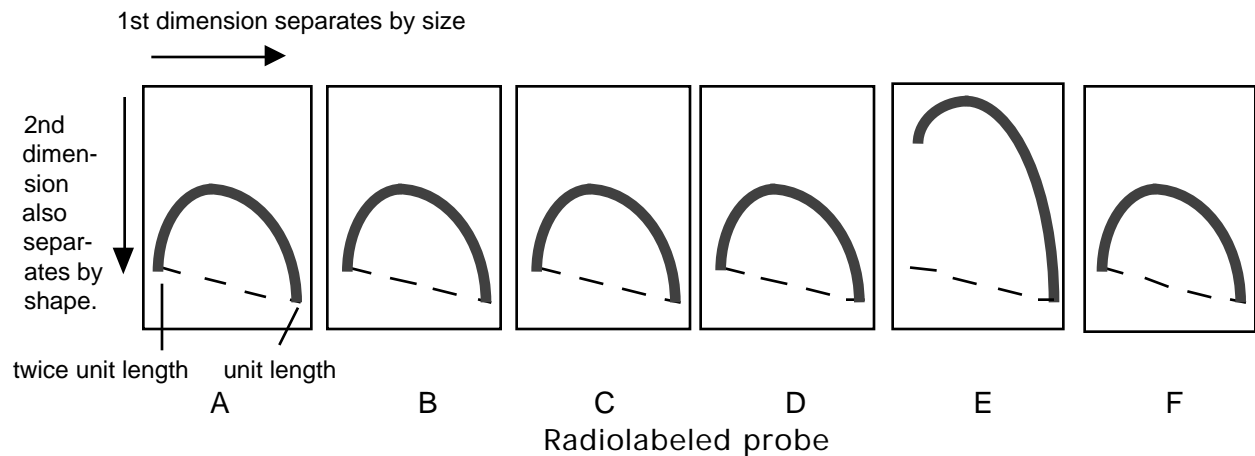
c is correct

A19=B32. (5 pts) Both parents of the affected individual were homozygous for the 3.5 kb D band. What does this tell you about the source of the DNA causing the polymorphism?

- a. The deletion occurred in somatic tissues in the affected individual.
- b. The Alu repeat results from a transposition event in the last generation, i.e. during meioses that led to either the sperm or the egg that formed the affected individual.
- c. The single nucleotide substitution occurred in during gametogenesis in the father (i.e. formation of sperm).
- d. An L1 repeat present in the 3.5 kb D allele duplicated during the meioses in the parents that led to the gametes that formed the affected individual.

b is correct

Your boss suggests that it may be useful to make a construct of the OB gene that will replicate the DNA from a natural origin in human cells. You figure it would be wise to check for replication origins in the cloned OB locus, so you synchronize a culture of adipocytes in S phase (DNA synthesis) to enrich for chromosomes with replication intermediates, isolate the DNA, digest with EcoRI and then run the DNA on Brewer and Fangman 2-D gels. These gels separate by size in the first dimension, but in the second dimension the separation is largely on the basis of shape. In particular, deviations from linear DNA will cause a substantial slowing of migration in the second dimension. The 2-D gel is then blotted onto a nylon membrane and hybridized individually to labeled single-copy DNA from fragments A, B, C, D, E, and F. The results of autoradiography of the hybridized blots are shown below.



A20=B28. (5 pts) What fragment(s) has/have replication origins by this assay? Assume the fewest number of origins consistent with this pattern.

- F
- E
- B and E
- A, B, C, D and F

b is correct

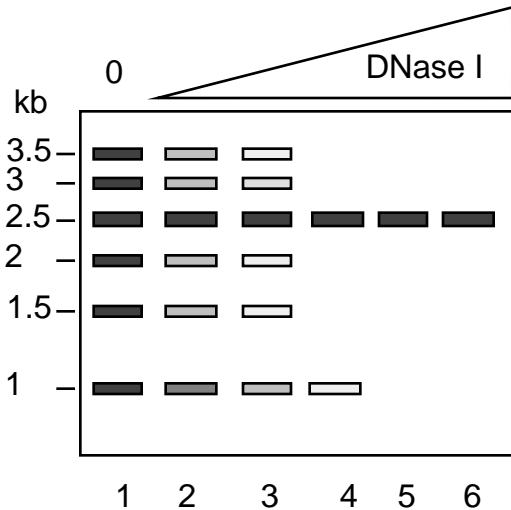
A21=B29. (5 pts) Which of the following experiments would tell you the direction of replication fork movement from this (these) origin(s)?

- Brewer and Fangman 2-D gels, using restriction fragments from the OB locus as probes in the blot-hybridization experiments.
- Monitor the kinetics of incorporation of [^{32}P]dATP into restriction fragments from the OB locus by DNA polymerase *in vitro*.
- Examine the replicative intermediates from the adipocytes in the electron microscope and measure the size of replication loops.
- Leading strand nascent chain assay (inhibit lagging strand synthesis, label replicating DNA and determine which strand it hybridizes to).

d is correct

In order to determine the approximate border to the active chromatin domain that contains the OB gene, nuclei from adipocytes were isolated and treated with increasing amounts of DNase I. DNA was purified from the digested nuclei, cut with EcoRI, separated on an agarose gel, blotted and hybridized to a panel of single-copy DNA probes covering the cloned OB region. The resulting autoradiograph is shown below, with the density of the fill proportional to the abundance of the fragment. What can you conclude about the chromatin structure of this region in adipocytes, based on the DNase-accessibility assays?

Digest nuclei with increasing amounts of DNase I, isolate DNA, digest with EcoRI, probe with single copy DNA from the OB region



A22=B33. (5 pts) Which restriction fragments are inside the active, or accessible, chromatin domain?

- a. F only
- b. E and F
- c. A, B, C, D, E
- d. A, B, C, D, E and F

c is correct

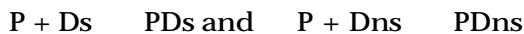
A23=B34. (5 pts) Where is the border between active and inactive (accessible versus inaccessible) chromatin?

- a. The data are not sufficient to address this issue.
- b. The border is between A and C.
- c. The border is close to the junction between D and E.
- d. The border is close to the junction between E and F.

d is correct

Use the following information for problems 24 and 25. Let's imagine that part of the regulation of expression of the OB gene is mediated by a protein we will call OBF1. There is one binding site for OBF1 in the OB gene, and let's assume that is the only specific binding site in the haploid genome, or 2 specific sites in a diploid genome. The haploid human genome has about 3×10^9 bp, or 6×10^9 bp in a diploid genome. If we assume that about 1/3 of the nuclear DNA is in an accessible chromatin conformation, that means that about 2×10^9 bp of DNA are available to bind OBF1 nonspecifically. The volume of a mammalian nucleus is about 5×10^{-13} L, which means the concentration of specific sites is about 6.6×10^{-12} M and the concentration of nonspecific sites is about 6.6×10^{-3} M. Binding of OBF1 to a specific site and to nonspecific sites is described by the following equations.

Let $P = \text{OBF1}$
 $D_s = \text{a specific binding site in DNA}$
 $D_{ns} = \text{a nonspecific binding site in the genomic DNA}$



$$K_s = \frac{[PD_s]}{[P][D_s]} = 10^{11} \text{ M}^{-1}$$

$$K_{ns} = \frac{[PD_{ns}]}{[P][D_{ns}]} = 10^5 \text{ M}^{-1}$$

A24=B35. (5 pts) What fraction of the OBF1 is **NOT** bound to either specific or nonspecific sites in the DNA? In other words, calculate $\frac{[P]}{[P \text{ total}]}$

- a. 660
- b. 0.66
- c. 0.015
- d. 0.0015

d is correct. Use the equation for K_{ns} , since the binding to nonspecific DNA will take up the bulk of the protein (there are 10^9 more nonspecific than specific sites). Plug in the values for K_{ns} and $[D_{ns}]$, and you calculate that $\frac{[P]}{[PD_{ns}]} = \frac{1}{660}$. Since $[P \text{ total}] = [P] + [PD_s] + [PD_{ns}]$ (note that $[PD_s]$ is extremely small relative to these other factors), the $\frac{[P]}{[P \text{ total}]} = \frac{1}{661} = 0.0015$.

A25=B36. (5 pts) How many molecules of OBF1 are needed per nucleus to maintain 90% occupancy of the specific sites? This condition means

$$\frac{[P]}{[P \text{ total}]} = 0.9$$

- a. 1.8 molecules of OBF1 per nucleus
- b. 1800 molecules of OBF1 per nucleus
- c. 18,000 molecules of OBF1 per nucleus
- d. 180,000 molecules of OBF1 per nucleus

c is correct. Use the equation for specificity (see the next to last page of the exam), which is the ratio between K_s and K_{ns} , which equals 10^6 in this case. Since $\frac{[P]}{[P \text{ total}]}$,

you can easily solve this equation for [P total], which is 60 nM. Then $(60 \times 10^{-9} \text{ moles L}^{-1})(5 \times 10^{-13} \text{ L})(6.02 \times 10^{23} \text{ molecules mole}^{-1}) = 18,000 \text{ molecules}$.

Use the following proteins to answer questions 26 and 27.

- [1] E. coli *lac* Repressor
- [2] E. coli CAP, or catabolite activator protein
- [3] cI protein, or repressor
- [4] cII
- [5] Cro

A26=B1. (5 pts) Which proteins can negatively regulate a target gene?

- a. 1, 3, and 5
- b. 2 and 4
- c. 3, 4 and 5
- d. 1 and 3

a is correct.

A27=B2. (5 pts) Which proteins can positively regulate a target gene?

- a. 1, 3, and 5
- b. 2 and 4
- c. 2, 3, and 4
- d. 1 and 4

c is correct. Recall that repressor can stimulate transcription from P_{RM} .

A28=B3. (5 pts) The N protein can relieve the polarity of certain amber mutations (which are substitutions leading to a UAG in the mRNA and premature termination of translation). Which of the following would you expect in a gene that is susceptible to such N-mediated relief of polarity?

- a. The gene should have binding sites for the host protein NusA.
- b. The gene should have *nut* sites.
- c. The gene should have binding sites for the host termination factor .
- d. Transcription of the gene should normally terminate at -independent sites.

b is correct.

A29=B4. (5 pts) Which of the following events listed as (1) through (4) occur in order for transcription to continue through the *trp* attenuator to express the *trp* operon in E. coli?

- (1) The leader peptide (*trpL*) is translated completely.
- (2) Translation of the leader peptide (*trpL*) stalls at *trp* codons.
- (3) A duplex forms between regions 2 and 3 in the RNA.
- (4) A duplex forms between regions 3 and 4 in the RNA.

- a. (1) and (4)
- b. (1) and (3)
- c. (2) and (4)
- d. (2) and (3)

d is correct.

Use the following proteins to answer questions 30 and 31.

- [1] *E. coli lac* Repressor
- [2] cI protein, or repressor
- [3] Cro
- [4] mammalian Oct1 protein
- [5] mammalian Sp1 protein
- [6] mammalian AP1 protein

A30=B5. (5 pts) Which proteins have a helix-turn-helix motif in their DNA binding domain?

- a. 1, 2, 3, and 4
- b. 1 and 3
- c. 2, 4 and 6
- d. 2, 3 and 5

a is correct.

A31=B6. (5 pts) Which protein forms a dimer through a leucine zipper?

- a. 2
- b. 4
- c. 5
- d. 6

d is correct.

For questions 32 to 34, from the following list, choose the enzyme(s) that carry out the designated process in *E. coli*.

- a. UvrA, UvrB, UvrC, UvrD
- b. MutH, MutL, MutS
- c. LexA
- d. RecBCD
- e. RecA

A32=B7. (5 pts) Mismatch repair

b is correct.

A33=B8. (5 pts) Excision repair

a is correct.

A34=B9. (5 pts) Assimilation of single strands to form a heteroduplex during recombination.

e is correct.

A35=BB10. (5 pts) Which RNA splicing event (listed as a-d) has all the following characteristics (listed as 1-4)?

- (1) Uses ATP to form a large spliceosome.
- (2) Uses a phosphoester transfer mechanism.
- (3) Introns begin with GU and end with AG.
- (4) Initiation of the splicing reaction uses the 2' OH of an internal adenine nucleotide.

- a. Splicing of group I introns (which include the precursor to rRNA in *Tetrahymena*.)
- b. Splicing of group II introns.
- c. Splicing of introns from nuclear pre-mRNA.
- d. Splicing of introns in pre-tRNA.

c is correct.

A36=B11. (5 pts) Which of the following DNA polymerases is best suited for highly processive copying of a DNA template with a very low error frequency?

- a. A thermostable DNA polymerase, such as Taq polymerase.
- b. The Klenow fragment of *E. coli* DNA polymerase I.
- c. An RNA-dependent DNA polymerase from a retrovirus.
- d. *E. coli* DNA polymerase III holoenzyme.

d is correct.

37. (20 points total)

A common route for the entry of galactose into the glycolytic catabolic pathway utilizes galactokinase to catalyze formation of galactose-1-phosphate and a transferase to convert galactose-1-phosphate to UDP-galactose. Subsequently, UDP-galactose is converted to UDP-glucose by an epimerase followed by release as glucose-1-phosphate. Imagine that you are studying a fungus that shows coordinate induction of these enzymes in response to galactose, and you have good assays for the galactokinase (referred to hereafter as kinase) and the transferase. Four genes or loci, *galA*, *galB*, *galC*, and *galD*, that affect the activity or regulation of these enzymes were studied in a series of haploid and diploid strains. In the following table, wild type alleles of the genes or loci are indicated by a + under the letter of the *gal* gene or locus and mutant alleles are indicated by a - under the letter. The activities of the kinase and transferase were assayed in the presence or absence of the inducer galactose. The units of enzyme activity are 100 = induced activity of the wild-type enzyme, 1 = repressed activity of the wild-type enzyme, and 0 = no measurable activity. In the diploid analysis, one copy of each operon is present in each cell.

Strain number	<i>gal</i>				kinase		transferase		
	A	B	C	D	- galactose	+ galactose	-galactose	+galactose	
Haploid 1	+	+	+	+	1	100	1	100	
2	-	+	+	+	1	100	0	0	
3	+	-	+	+	100	100	100	100	
4	+	+	-	+	0	0	1	100	
5	+	+	+	-	100	100	100	100	
Diploid 6	A	B	C	D/A	B	C	D		
6	+	+	-	+/-	+	+	+	1	100
7	+	-	-	+/-	+	+	-	100	100
8	-	-	+	+/+	+	-	-	1	100
9	-	+	+	+/+	+	-	-	1	100

a. (8 pts) Describe the phenotypes of the following the strains with respect to kinase and transferase activities. A single word will suffice for each phenotype.

Strain	kinase	transferase
2	inducible	inactive, null, deficient
3	constitutive	constitutive
4	inactive, null, deficient	inducible
5	constitutive	constitutive

b. (4 pts) The diploid analysis shows three examples of a dominant constitutive phenotype, either for the kinase or for the transferase activity. What do these three examples have in common? The answer can be succinctly stated by choosing the correct option for filling in the following 3 blanks.

Answer: In all three cases, the reporter gene (i.e. encoding either the kinase or the transferase) is in cis (*cis* or *trans*) to the mutant

(wild-type or mutant) allele of the galD (*galA*, *galB*, *galC*, *galD*) gene or locus.

c. (8 pts.) What is the role of each of these genes or loci in activity or regulation of galactose metabolism? Brief phrases will suffice.

galA

encodes the transferase

galB

encodes a repressor

galC

encodes the kinase

galD

is an operator

List of Restriction Endonucleases and their Cleavage Sites:

Enzyme	Site
<i>EcoRI</i>	G'AATTC
<i>HindIII</i>	A'AGCTT
<i>SacII</i>	CCGC'GG

The Genetic Code

1st	Position in Codon								3rd
	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	Term	UGA	Term	A
	UUG	Leu	UCG	Ser	UAG	Term	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

* Sometimes used as initiator codons.

Binding sites for transcription factors

Transcription factor	DNA binding site
Sp1	GGGCGG
AP1	TGASTCA
Oct1	ATTGTCAT

Transcription factor	DNA binding site
GAL4	CGGASGACWGTCTCCG
Glucocorticoid Receptor	TGGTACAAATGTTCT
MyoD	CAGCTG

S = G or C W = A or T

Equation for specificity of binding of proteins to DNA:

$$\text{specificity} = \frac{K_s}{K_{ns}} = \frac{[PDs]}{[Ds]} \times \frac{[Dns]}{[PDns]} = \frac{[PDs]}{[Ds]} \times \frac{[Dns]}{[P \text{ total}] - [Ds \text{ total}]}$$

Equations for complexity (N) and repetition frequency (R):

$$N_n = C_{ot1/2}^{\text{pure},n} \times \frac{N_{\text{std}}}{C_{ot1/2}^{\text{std}}}$$

$$R_n = \frac{f_n G}{N_n} = \frac{C_{ot1/2}^{\text{mix single copy}}}{C_{ot1/2}^{\text{mix } n}}$$

This is the basic answer to much of questions 1-25:

Restriction map of the cloned genomic DNA containing the OB locus

