

Biochemistry and Molecular Biology 400
Final Examination
Fall 1996
December 12, 1996

Instructor: Hardison

This examination has 35 questions for a total of 200 points. The first 34 are multiple choice and are worth 5 points each (except for one worth 10 points). 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 two questions 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.

Useful information and equations are on the last page. You may wish to tear this off and use it throughout the exam.

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 the next two questions, consider the following proteins involved in gene regulation:

- [1] *lac* repressor from *E. coli*
- [2] CAP protein (or catabolite activator protein) from *E. coli*
- [3] repressor, product of the *cI* gene
- [4] cII protein, product of the *cII* gene
- [5] mammalian glucocorticoid receptor
- [6] mammalian AP1

A1=B4. (5 pts) Which proteins can have a *negative* effect on expression of target genes?

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

a is correct

A2=B5. (5 pts) The activity or function of which of the proteins is regulated by binding of a small molecule ligand?

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

c is correct, 2 points for a

A3=B1. (5 pts) Which of the following statements about regulation of the *lac* operon in *E. coli* are correct?

- [1] The operon is negatively regulated by a repressor with the compound allolactose bound to it.
- [2] Transcription of the operon is stimulated by binding of the complex cAMP-CAP.
- [3] The *lac* repressor prevents binding of RNA polymerase.
- [4] Binding of the complex cAMP-CAP to DNA will bend the DNA.

Correct choices are:

- a. 1, 2, 3, 4 b. 2, 4 c. 2, 3 d. 1, 3 e. 2, 3, 4

b is correct

A4=B2. (5 pts) Which of the following events occur when *E. coli* is grown in the presence of the amino acid tryptophan?

- [1] The ribosome translates the leader peptide completely (to the UGA stop codon).
- [2] A G+C rich stem-loop structure forms in the nascent RNA (regions 3 and 4) at the attenuator site.
- [3] A stem-loop structure forms in the nascent RNA (regions 2 and 3) that precludes formation of the G+C rich stem-loop at the attenuator site.
- [4] Transcription reads through the attenuator into *trp* *EDCBA*.

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

a is correct, 2 points for d.

A5=B3. (5 pts) The mutants cI^- and cII^- produce no lysogens, so they make clear plaques. If they are coinfecting into *E. coli*, will they produce turbid plaques, and if so which phage will be found in the resulting lysogen?

- a. Turbid plaques will be produced, and they will contain only the cI^- phage.
- b. Turbid plaques will be produced, and they will contain only the cII^- phage.
- c. Turbid plaques will be produced, and they will contain both the cI^- and cII^- phage.
- d. No turbid plaques will be produced.

b. is correct

For the next two questions, consider a hybrid protein that fuses the DNA-binding domain of Sp1 with the activation domain of GAL4 and the hormone binding domain of the glucocorticoid receptor.

A6=B6. (5 pts) Under conditions favoring DNA binding, what would be expected for the DNA-binding properties of the hybrid protein? The protein will

- [1] Bind with high specificity to DNA containing the sequence CGGAGGACTGTCGTCCG.
- [2] Bind with high specificity to DNA containing the sequence GGGGCGGGG.
- [3] Bind with high specificity to DNA containing the sequence TGGTACAAATGTTCT.
- [4] Bind via a domain containing C_2H_2 Zn fingers.
- [5] Bind via a helix-turn-helix domain.
- [6] Bind via a leucine zipper domain.

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

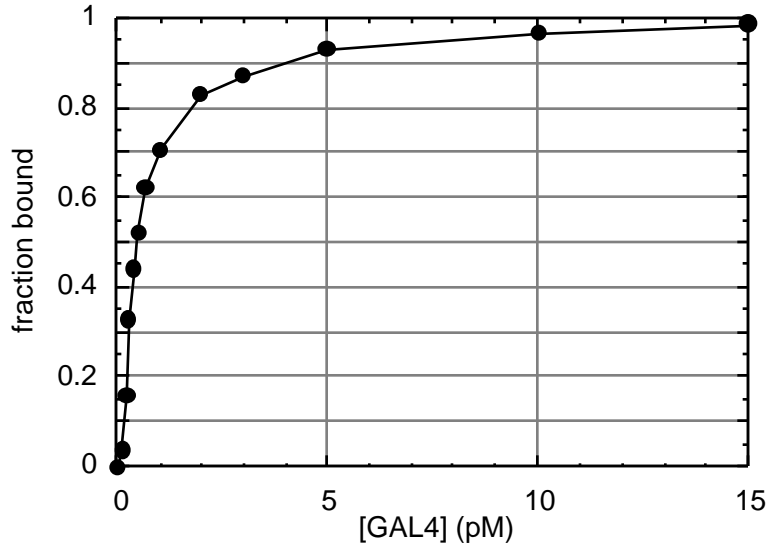
d is correct

A7=B7. (5 pts) What effects will the hybrid protein be expected to have on transcription of target genes (i.e. containing the appropriate DNA-binding site)?

- a. Activate target genes only in the presence of galactose.
- b. Activate target genes only in the presence of a glucocorticoid hormone.
- c. Activate all target genes containing the appropriate DNA-binding site.
- d. Activate only genes encoding enzymes involved in galactose metabolism.

b is correct, 3 pts for a.

For the next two questions, let's imagine that you mixed increasing amounts of GAL4 with a constant amount of a labeled duplex oligonucleotide containing UAS_G, and measured the fraction of the labeled DNA bound to GAL4 as a function of [GAL4]. You obtained the following data. The X-axis is the concentration of GAL4 in pM (1 pM = 1×10^{-12} M). The sequence of UAS_G is CGGAGGACTGTCGTCCG.



The equation that describes this curve is

$$\frac{[PD]}{[D_{tot}]} = \frac{K_s[P]}{1+K_s[P]}$$

where $[PD]$ is the concentration of the protein GAL4 bound to its specific site on DNA, $[D_{tot}]$ is the concentration of total DNA, i.e. $[D_{free}] + [PD]$, $[P]$ is the concentration of the protein GAL4, and K_s is the equilibrium constant for GAL4 binding to its specific site on DNA.

A8=B10. (5 pts) When half of the labeled DNA is bound to GAL4, what is the relationship between $[P]$ and K_s ?

- a. $[P] = \frac{K_s}{1+K_s}$
- b. $[P] = \frac{[D_{tot}]}{2K_s}$
- c. $[P] = K_s$
- d. $[P] = \frac{1}{K_s}$

d is correct

A9=B11. (5 pts) What is the measured value of K_s ?

- a. $15 \times 10^{-12} \text{ M}$
- b. $1.5 \times 10^{-9} \text{ M}$
- c. $2 \times 10^{12} \text{ M}^{-1}$
- d. $2 \times 10^9 \text{ M}^{-1}$

c is correct

For the next two questions, consider a stage after infection of *E. coli* with λ in which the λ phage has not yet replicated, so there is one copy of the λ genome per cell. For simplicity, let's imagine a λ phage with a single repressor binding site (also referred to as the operator, or O). The *E. coli* genome has about 4.2×10^6 bp, and let's assume there is only one genome per cell. The volume

of the cell is 1×10^{-15} L, giving a concentration of nonspecific binding sites of 7×10^{-3} M, a concentration of specific binding sites of 1.7×10^{-9} M. The following equations apply, and we will ignore the effects of Cro for these problems.



$$K_{s,r} = \frac{[RO]}{[R][O]} = 10^{11} \text{ M}^{-1} \quad (\text{eqn 2})$$

$$K_{ns,r} = \frac{[RD_{ns}]}{[R][D_{ns}]} = 10^5 \text{ M}^{-1} \quad (\text{eqn 3})$$

where

R = repressor dimer

O = operator site

D_{ns} = a nonspecific binding site in the genomic DNA

A10=B8. (5 pts) When the repressor is present in excess over the concentrations of operator sites, what fraction of the repressor is not bound to either specific or nonspecific sites?

- a. 1×10^{-11}
- b. 0.00001
- c. 0.9986
- d. 0.0014
- e. Cannot be calculated from the information given.

d is correct

A11=B9. (10 pts) How much repressor must be present in the cell to occupy 90% of the operator sites? This is equivalent to $\frac{[RO]}{[O]} = 9$. First calculate the [R] needed to achieve this, and then express this in molecules of repressor per cell. Avagadro measured about 6.02×10^{23} molecules per mole.

- a. 4 molecules per cell
- b. 37 molecules per cell
- c. 100 molecules per cell
- d. 4000 molecules per cell
- e. Cannot be calculated from the information given.

b is correct

For the next three questions, consider the following information about a protein called Gcn5p. [This problem is based on Brownell et al. (1996) Cell 84: 843-851.]

[1] Gcn5p is needed for activation of some, but not all, genes in yeast.

[2] Gcn5p does not bind with high affinity to any particular site on DNA.

[3] Gcn5p will interact with acidic transcriptional activators.

[4] When incubated with histones and the following substrates, Gcn5p will have the designated effects. A + in the column under "Effect" means that the histones move slower than unmodified histones on a polyacrylamide gel that separates on the basis of charge, with the histones moving toward the negatively charged electrode. A - means that the treatment has no effect on the histones. S-adenosylmethionine is a substrate for some methyl transfer reactions, and NADH is the substrate for ADPribosyl-transferases.

<u>Mixture</u>	<u>Effect</u>
Gcn5p + histones	-
Gcn5p + histones + ATP	-
Gcn5p + histones + S-adenosylmethionine	-
Gcn5p + histones + acetyl-coenzyme A	+
Gcn5p + histones + NADH	-

A12=B12. (5 pts) Which of the following conclusions is consistent with these observations?

- Gcn5p is a transcriptional activator that works by binding DNA and stimulating the level of transcription of target genes.
- Gcn5p aids in the activation of target genes by interacting with other transcriptional activators which bind to specific DNA sequences.
- Gcn5p is a repressor of most genes in yeast.
- Gcn5p acts to turn off expression of genes by increasing the level of methylation of the DNA.

b is correct

A13=B13. (5 pts) What enzymatic activity is associated with Gcn5p?

- histone kinase
- histone methyltransferase
- histone acetyltransferase
- DNA ligase
- DNA methyltransferase

c is correct

A14=B14. (5 pts) Which step in the gene expression pathway is likely to be regulated by Gcn5p?

- chromatin remodeling/ activation
- rate of initiation of transcription
- rate of transcription elongation
- rate of initiation of translation

a is correct

Match the following phrases with the RNA splicing event listed in the next two questions. Pick the choice with the most correct matches.

- [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 3' OH of a guanine nucleotide or nucleoside.
- [5] Is self-splicing in vitro.
- [6] Initiation of the splicing reaction uses the 2' OH of an internal adenine nucleotide.

A15=B30. (5 pts) Splicing of group I introns (which include the precursor to rRNA in *Tetrahymena*.)

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

c is correct: statements 2, 4 and 5 are true for this process.

A16=B31. (5 pts) Splicing of introns from nuclear pre-mRNA.

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

b is correct: statements 1, 2, 3 and 6 are true for this process.

For the next two questions, choose the polymerase that will be best for the designated task.

A17=B32. (5 pts) Labeling by nick translation, i.e. introducing radioactive nucleotides into nicked, circular duplex DNA, starting with [³²P] dNTPs.

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

b. is correct

A18=B33. (5 pts) Amplification of a defined segment of DNA by the polymerase chain reaction.

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

a is correct.

A19=B34. (5 pts) How does DNA polymerase III achieve high processivity? Choose the correct statement.

- a. The core () is inherently highly processive.
- b. The complex is a "clamp" that holds the polymerase core onto the template.
- c. The complex loads the β "sliding clamp" onto the primer-template junction, and after ATP hydrolysis the β clamp will exchange from the complex to the core.

d. The β "sliding clamp" is irreversibly linked to the polymerase core.
c. is correct.

A20=B26. (5 pts) Much important confirmatory evidence on the genetic code has come from the nature of single-residue changes in the amino acid sequence of mutant proteins. Using the "universal" genetic code, which of the following single-residue amino acid replacements *cannot* be the result of single-base mutations?

- a. Phe Leu
- b. Lys Ala
- c. Ala Thr
- d. Ile Leu

b. is correct

A21=B27. (5 pts) Which of the following statements about protein synthesis in *E. coli* is true?

- a. The peptidyl transferase step requires hydrolysis of high energy phosphate bonds in GTP.
- b. Binding of f-Met-tRNA to the mRNA on the small ribosomal subunit utilizes ATP and IF3.
- c. Translocation of the peptidyl-tRNA from the A site to the P site of the ribosome is catalyzed exclusively by an RNA component of the ribosome.
- d. The 16 rRNA on the small subunit holds the initiator AUG in register for formation of the initiation complex (via base pairing).

d. is correct

A22=B28. (5 pts) Which RNA polymerase(s) need(s) additional general transcription factors for accurate initiation at promoter sequences?

- [1] eukaryotic RNA polymerase I
- [2] eukaryotic RNA polymerase II
- [3] eukaryotic RNA polymerase III
- [4] *E. coli* RNA polymerase holoenzyme

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

b is correct (5 pts), and 2 pts for c and 1 pt for d.

A23=B29. (5 pts) Which RNA polymerases are used to transcribe genes encoding proteins?

- [1] eukaryotic RNA polymerase I
- [2] eukaryotic RNA polymerase II
- [3] eukaryotic RNA polymerase III
- [4] *E. coli* RNA polymerase

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

c is correct

A24=B15. (5 pts) Which of the following statements are true about duplex DNA? For this question, it does not matter whether the DNA is in A, B, or Z form.

- [1] A purine nucleotide is paired with a pyrimidine nucleotide.
- [2] A keto-base nucleotide is paired with an amino-base nucleotide.
- [3] The base-pairing specificity is dependent on the tautomerization (keto versus enol) of the nucleotides.
- [4] The two strands are parallel with respect to their 5' to 3' orientation.

- a. 1, 2, 4 b. 3, 4 c. 1, 2, 3 d. 1, 3
c. is correct

A25=B16. (5 pts) Which of the following will **decrease** the melting temperature of duplex DNA?

- a. Increasing the G+C content of the duplex.
- b. Increasing the number of mismatches in the duplex.
- c. Increasing the NaCl concentration of the DNA solution.
- d. Increasing the DNA concentration in the solution.

b is correct.

For the next three questions, choose the enzyme(s) that carry out the designated process in *E. coli*.

A26=B17. (5 pts) Mismatch repair

- a. UvrA, UvrB, UvrC, UvrD
- b. MutH, MutL, MutS
- c. UmuC, UmuD
- d. Uracil-N-glycosylase

b is correct

A27=B18. (5 pts) Base excision repair

- a. Uracil-N-glycosylase and AP endonuclease
- b. MutH, MutL, MutS
- c. UmuC, UmuD
- d. UvrD

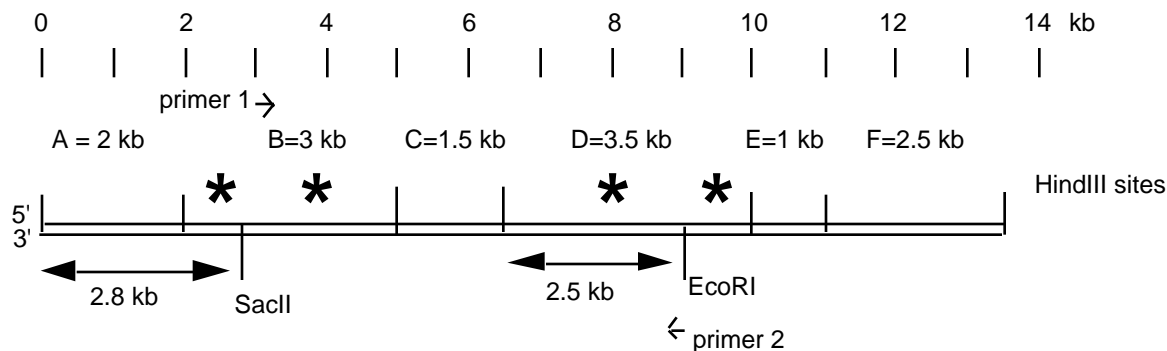
a is correct

A28=B19. (5 pts) Generation of single stranded ends for strand invasion during recombination

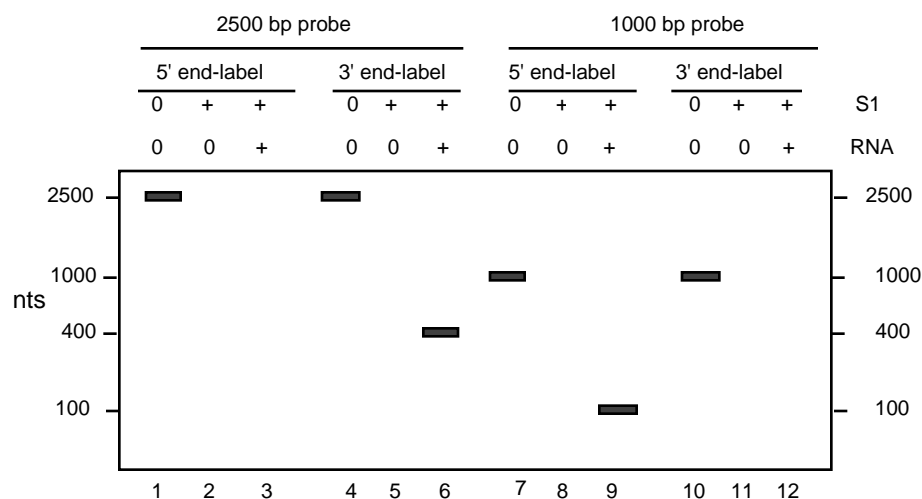
- a. RecBCD
- b. RecA
- c. The 5' to 3' exonuclease activity of DNA polymerase I
- d. RuvC

a is correct.

Use the following information for the next 4 problems. The *GCN5* gene encodes a protein of 439 amino acids, called Gcn5p. A hypothetical restriction map of the region containing the entire *GCN5* gene is shown below. The *Hind*III fragments are marked along the top of the map, and the large asterisks indicate restriction fragments that hybridize to *GCN5* mRNA.



The positions of the exons of the *GCN5* gene were mapped more precisely using an nuclease S1-protection assay. The clone of genomic DNA was digested with *Eco*RI (which is located at position 9000 on the restriction map), radiolabeled on the 5' end using [³²P]ATP and polynucleotide kinase, and digested with *Hind*III. An 2500 bp fragment extending from the C/D junction to the labeled *Eco*RI site was isolated, as was a 1000 bp fragment extending from the labeled *Eco*RI site to the D/E junction. In a parallel experiment, the same DNA fragments were labeled on the 3' end at the *Eco*RI site. The labeled DNA fragments were denatured and mixed with *GCN5* mRNA. RNA-DNA heteroduplexes were formed by annealing the labeled DNA and the RNA, after which 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.



A29=B22. (5 pts) Relative to the restriction map above, what direction is the *GCN5* gene transcribed?,

- It is transcribed from left to right.
- It is transcribed from right to left.

- c. Two genes are present, and they are divergently transcribed (i.e. the 5' ends of the genes are close together).
- d. Two genes are present, and they are convergently transcribed (i.e. the 3' ends of the genes are close together).

b is correct

A30=B23. (5 pts) If the only exon in the HindIII fragment D is the one mapped by the S1 nuclease protection assay, what can you conclude about the *GCN5* gene?

- [1] The 5' end of the *GCN5* gene is located at position 9100 on the restriction map.
- [2] The 3' end of the *GCN5* gene is located at position 8600 on the restriction map.
- [3] The 5' end of the *GCN5* gene is located at position 10,000 on the restriction map.
- [4] The 3' end of the exon of the *GCN5* gene in fragment D is located at position 8600 on the restriction map.
- [5] The 3' end of the exon of the *GCN5* gene in fragment D is located at position 6500 on the restriction map.

Which statement(s) is(are) correct?

- a. 1, 2 b. 3, 5 c. 2, 3 d. 1, 4 e. none are correct

d is correct

A31=B24. (5 pts) If the only exon in the HindIII fragment D is the one mapped by the S1 nuclease protection assay, what can you conclude about the hybridization to HindIII fragment B?

- a. Fragment B contains an exon encoding the C terminus of the Gcn5p protein.
- b. Fragment B contains an exon encoding the N terminus of the Gcn5p protein.
- c. Fragment B contains an gene related to but distinct from *GCN5* (i.e. it will hybridize to *GCN5* but has a somewhat different DNA sequence).
- d. Fragment B contains a gene unrelated to *GCN5*.

a. is correct

A32=B25. (5 pts) Let's say that you wanted to devise a sensitive assay to detect *GCN5* mRNA. You synthesize two oligonucleotides to amplify DNA by the polymerase chain reaction (PCR). Primer 1 contains positions 3000 to 3050 of the *GCN5* locus shown above, reading 5' to 3' left to right along the top strand of the DNA. Primer 2 contains positions 9000 to 8950, reading 5' to 3' right to left along the bottom strand of the DNA. You isolate nucleic acids from cells; this preparation is highly enriched for RNA but still contains a small amount of DNA. The RNA is converted to cDNA by reverse transcriptase and then amplified by PCR. You find two major products of the PCR reaction, one that is 6000 bp long and one that is 700 bp long. What do you conclude?

- a. Two different RNAs are generated from the *GCN5* gene by alternative splicing.
- b. Two different RNAs are generated from the *GCN5* gene, one with the introns removed by splicing and the other unspliced.
- c. Two different homologs of the *GCN5* gene encode different sizes of RNA.
- d. The 6000 bp PCR product is the result of contaminating genomic DNA, and the 700 bp PCR product is generated from the spliced mRNA.

d is correct, 3 points for b.

For the next two questions, consider the following. The kinetics of reassociation of a genomic DNA shows two components. Component I occupies 0.2 of the genome and renatures with a measured $C_{0t_{1/2}}$ of 0.001, and component II occupies 0.8 of the genome and renatures with a measured $C_{0t_{1/2}}$ of 1000 (the measured $C_{0t_{1/2}}$ values are for the mixture of components I and II in genomic DNA). Under identical conditions, a phage genome with a complexity of 10^5 renatures with a $C_{0t_{1/2}}$ of 1. Assume that the slow renaturing component is single-copy.

A33=B21. (5 pts) What is the complexity of the two components?

	<u>Component I</u>	<u>Component II</u>
a.	0.001	1000
b.	0.2	0.8
c.	20	8×10^7
d.	200	8×10^9

c is correct

A34=B35. (5 pts) What is the repetition frequency of the fast-renaturing component I?

- a. 20
- b. 800
- c. 1×10^6
- d. 4×10^8

c. is correct.

35. (25 points) The bacterium *E. coli* can synthesize guanylic acid (or GMP) from inosinic acid (or IMP) in two enzymatic steps. The first is oxidation of IMP, catalyzed by IMP dehydrogenase, to make xanthylic acid (or XMP). The second step replaces a carbonyl on the purine ring by an amino group, and is catalyzed by GMP synthetase. The genes for these enzymes are encoded in an operon. Growth in the presence of guanine (which leads to increased intracellular levels of guanine nucleotides) will repress the operon. Consider the following hypothetical data for the operon. Four genes or loci, *guaA*, *guaB*, *guaC*, and *guaD* that affect the activity or regulation of the biosynthetic 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 *gua* gene or locus and mutant alleles are indicated by a - under the letter. The activities of the two enzymes, IMP dehydrogenase and GMP synthetase, were assayed after growth in the presence or absence of guanine . The units of enzyme activity are 100 = non-repressed activity of the wild-type enzyme, 1 = repressed activity of the wild-type enzyme (in the presence of guanine), and 0 = no measurable activity. In the diploid analysis, one copy of each operon is present in each cell.

Strain number	<i>gua</i>				IMP dehydrogenase		GMP synthetase		
	A	B	C	D	- guanine	+ guanine	-guanine	+guanine	
Haploid 1	+	+	+	+	100	1	100	1	
2	-	+	+	+	100	1	0	0	
3	+	-	+	+	0	0	100	1	
4	+	+	-	+	100	100	100	100	
5	+	+	+	-	100	100	100	100	
Diploid 6	A	B	C	D/A	B	C	D		
6	-	+	+	+/+	-	+	+	100	1
7	+	+	-	+/+	+	+	+	200	2
8	+	-	+	-/-	+	+	+	100	100
9	-	+	+	-/+	-	+	+	100	1
10	+	-	-	+/-	+	+	+	100	1

Use the data in the table to answer the following questions.

a) (10 pts) Describe the phenotype of the following the strains with respect to IMP dehydrogenase and GMP synthetase activities. A single word will suffice for each phenotype.

	IMP dehydrogenase	GMP synthetase
Strain 2	_____	_____
Strain 3	_____	_____
Strain 4	_____	_____
Strain 5	_____	_____
Strain 6	_____	_____

- b)** (2 pts) What does strain 6 tell you about the relationship (dominant or recessive) between wild-type and mutant alleles of the genes *guaA* and *guaB*?

guaA⁺ is _____ to *guaA*⁻ and

guaB⁺ is _____ to *guaB*⁻.

- c)** (4 pts) What are the roles of *guaA* and *guaB* in synthesis or regulation of GMP?

guaA :

guaB:

- d)** (1 pt) What do strains 7 and 10 tell you about the relationship (dominant or recessive) between wild-type and mutant alleles of the gene *guaC* ?

guaC⁺ is _____ to *guaC*⁻.

- e)** (4 pts) What do strains 8 and 9 tell you about the relationship (dominant or recessive) between wild-type and mutant alleles of the genes *guaD* and *guaD*? Explain your answer.

- f)** (4 pts) What are the roles of *guaC* and *guaD* in synthesis or regulation of GMP?

guaC

guaD

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	ATTTGCAT

S = G or C W = A or T

Transcription factor	DNA binding site
GAL4	CGGASGACWGTCTCCG
Glucocorticoid Receptor	TGGTACAAATGTTCT
MyoD/E47	CAGCTG

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):

$$N_n = C_0 t_{1/2}^{\text{pure},n} \times \frac{N^{\text{std}}}{C_0 t_{1/2}^{\text{std}}}$$

$$R_n = \frac{f_n G}{N_n} = \frac{C_0 t_{1/2}^{\text{mix single copy}}}{C_0 t_{1/2}^{\text{mix } n}}$$

Answer to number 35:

- a) (10 pts) Describe the phenotype of the following the strains with respect to IMP dehydrogenase and GMP synthetase activities. A single word will suffice for each phenotype.

	IMP dehydrogenase	GMP synthetase
Strain 2	_____ repressible _____	_____ inactive _____
Strain 3	_____ inactive _____	_____ repressible _____
Strain 4	_____ constitutive _____	_____ constitutive _____
Strain 5	_____ constitutive _____	_____ constitutive _____
Strain 6	_____ repressible _____	_____ repressible _____

- b) (2 pts) What does strain 6 tell you about the relationship (dominant or recessive) between wild-type and mutant alleles of the genes *guaA* and *guaB*?

guaA⁺ is _____ **dominant** _____ to *guaA*⁻ and

guaB⁺ is _____ **dominant** _____ to *guaB*⁻.

- c) (4 pts) What are the roles of *guaA* and *guaB* in synthesis or regulation of GMP?

guaA : **encodes GMP synthetase**

guaB : **encodes IMP dehydrogenase**

- d) (1 pt) What do strains 7 and 10 tell you about the relationship (dominant or recessive) between wild-type and mutant alleles of the gene *guaC* ?

guaC⁺ is _____ **dominant** _____ to *guaC*⁻.

Note that both these strains are heterozygous for *guaC*⁺/*guaC*⁻, and the wild-type alleles of *guaA* and *guaB* are repressible. However, in the haploid *guaC*⁻, the *guaA* and *guaB* genes are constitutive.

- e) (4 pts) What do strains 8 and 9 tell you about the relationship (dominant or recessive) between wild-type and mutant alleles of the genes *guaD* and *guaD*? Explain your answer.

The allele of *guaD* that is in *cis* to the active allele of *guaA* or *guaB* (consider these the reporter genes) is the dominant allele. For example in strain 8, GMP synthetase is constitutive because its wild-type gene, *guaA* is in *cis* to the mutant allele of *guaD*. In contrast, IMP dehydrogenase is repressible because its wild-

type gene, *guaB*, is in *cis* to the wild-type allele of *guaD*. Strain 9 shows the same effects, except that now the mutant *guaD* leads to constitutive expression of *guaB* in *cis*. Thus *guaD* shows *cis*-dominance, and thus one surmises that *guaD* is a site on the DNA, probably an operator to which a repressor binds.

f) (4 pts) What are the roles of *guaC* and *guaD* in synthesis or regulation of GMP?

guaC: encodes an apo-repressor to which guanine nucleotides bind to make an active repressor. Full credit (2 pts) for just saying "repressor."

guaD: operator (full credit for saying "operator") to which the active repressor binds.