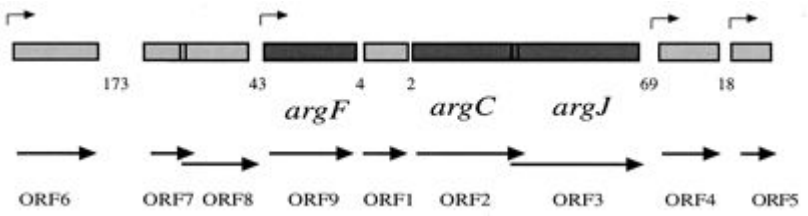
Molecular Biology Midterm 4

- - -

1. (1 pt.) How is it that transcription can induce positive supercoils (overwinding) in part of the template? Which part? (Ignore other factors, such as topoisomerases.)

2. (6 pts.) (continues on next page) The following Fig is from a recent paper. It is a map of several contiguous genes in a bacterium. The genes are simply named ORF with a number; ORF means "open reading frame", a name for a region that appears to be a gene, but has not yet been specifically identified. Some of the ORFs were later shown to be part of the pathway for synthesizing the required amino acid nutrient, arginine; these genes are also labeled with the name *arg*. The bent arrows at the top identify transcriptional start sites.



The numbers below the top map (173, etc) are the number of nucleotides between genes. You can mostly ignore these, though one will be somewhat useful in part a.

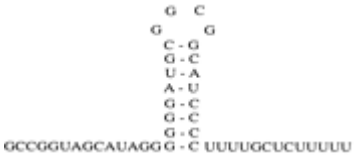
All parts of this question are intended to be about transcription and its regulation, from chapters officially covered by this test. They are not intended to be about translational issues. (The article itself is about such issues, in part, and you may think they are relevant. But it would be unfair for the test to grade on any translational issues here.)

⇒ The parts of this question (next page) are mostly independent.

a. (2 pts.) They show that a regulatory protein binds a site overlapping the argF promoter. Part of their evidence was a footprint experiment, which shows protection of a 29 base region, which extends from the middle of the -10 region to somewhat upstream of the -35 region. Sketch the result of the footprint experiment that would lead to this conclusion. Be sure to include a control. Label key features of your sketch well enough to make the main point. Relate the sketch to the map shown above. (That is, show what region of the top map is involved in this part.)

b. (1 pt.; no credit without explanation) Judging from the information given in part a, is it most likely that this protein is a positive control protein or a negative control protein? Explain; that is, how can you tell from the given information? Be sure your explanation is clear enough to distinguish what the two terms mean.

c. (2 pts.) The Figure at the right shows part of an RNA from a region near the beginning of ORF5. What is the significance of this structure? Which transcript is it from? Point to two specific features that help you identify its function.



d. (1 pt.) Adding arginine to the bacteria increases transcription of ORF4. What does this suggest to you about the function of the ORF4 protein? In particular, is ORF4 likely to be involved in making arginine? Explain.

3. (2 pts.) Consider a DNA sequence upstream of a gene. For simplicity, let's just look at one strand, and say that the "base sequence" is abcdefghijklmnopqrstuvwxyz. You want to explore the importance of this region in gene expression, so you make many mutants in this region. In one set of experiments, you make small deletions (say, 2-3 bases at a time). These experiments show that deletions of any bases from d to w reduce gene expression. In another set of experiments, you do extensive linker scanning tests over this region. The linker scanning results show that only changes in bases d-i and r-w reduce gene expression. Suggest a reasonable model for what this region does, based on these two sets of experiments. In particular, explain why the two experiments seem to point to different parts of the DNA as being important.

4. (6 pts.; 2 pts. per part) (continues on next page) The following is an abstract of a recent paper.

The Escherichia coli *proP* P2 promoter, which directs the expression of an integral membrane transporter of proline, glycine betaine, and other osmoprotecting compounds, is induced upon entry into stationary phase to protect cells from osmotic shock. Transcription from the P2 promoter is completely dependent on RpoS (σ^{38}) and Fis. Fis activates transcription by binding to a site centered at -41, which overlaps the promoter, where it makes a specific contact with the C-terminal domain of the α subunit of RNA polymerase (α-CTD). We show here that Fis and cyclic AMP (cAMP) receptor protein (CRP)-cAMP collaborate to activate transcription synergistically in vitro. Coactivation both in vivo and in vitro is dependent on CRP binding to a site centered at -121.5, but CRP without Fis provides little activation. The contribution by CRP requires the correct helical phasing of the CRP site and a functional activation region 1 on CRP. We provide evidence that coactivation is achieved by Fis and CRP independently contacting each of the two α-CTDs. Efficient transcription in vitro requires that both activators must be preincubated with the DNA prior to addition of RNA polymerase. (end of abstract)

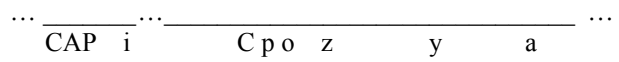
Answer the following parts (next page) based on the information in this abstract. The parts are substantially independent. (The particular σ factor discussed is of no concern for the questions below.)

a. Sketch the genome over the region discussed here. Your sketch should show the transcriptional start site, the gene whose transcription is of interest, and all the features discussed in or clearly implied by the abstract. (You need not show the genes that code for regulatory proteins; their locations are not indicated or known here.) (You can use the *lac* operon sketch on p 6 as a guide to style and level of detail needed.)

b. Using the features of this system, give an example of a genetic test that would show cis dominance. That is, describe a diploid strain for this region and describe the result that would show cis dominance. Be sure to explain what is meant by cis dominance. (You can add a reporter gene, as needed.)

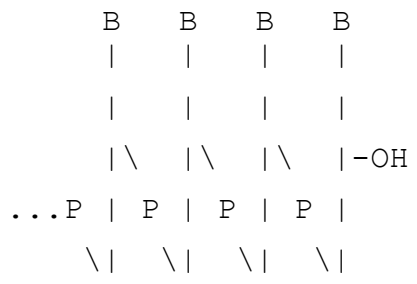
c. Sketch the "pre-initiation complex" -- the DNA with appropriate proteins bound, ready to initiate transcription. Try to incorporate as many features described in the abstract as possible. (The core RNA Pol subunit composition is $\alpha_2\beta\beta'$.)

The sketch below is for your general guidance; it shows the general (simplified) organization of the *lac* operon:



The CAP (= CRP) and i proteins are the regulatory proteins, which act at the sites C or o, respectively. p is the promoter; the remaining letters denote genes for the *lac* enzymes.

5. (3 pts.) The diagram below shows part of a growing RNA chain. Each | represents one position of the ribose sugar. (P = one phosphate group; B = any of the four normal bases. I have omitted the 2′-OH, which is not relevant to the question.)



a. (1 pt.) I showed horizontal lines above and below the main chain. Which (one) of those horizontal lines (best) represents where the template strand would be? Mark it, and label the 5′ and 3′ ends of this template strand.

b. (2 pts.) The immediate precursor for RNA synthesis is an activated nucleotide. Using the same diagram system for a nucleotide that I used in the main drawing, show the next nucleotide approaching the chain, about ready to form a bond to extend the chain. Show the activated structure of the incoming nucleotide and where it is about to bond. Use an arrow to show where/how the activated nucleotide precursor will connect to the chain.