APh161: Physical Biology of the Cell
Homework 3
Due Date: Thursday, January 26, 2006

"Thinking, analyzing, inventing are not anomalous acts; they are the normal respiration of the intelligence." Jorge Luis Borges in *Pierre Menard, Author of the Quixote*

1. Poster Session Title and Abstract

Please email Eric Peterson, Jen Barnet and me your title and abstract for the poster session project. Your abstract should be about half a page long and should state what the phenomenon of interest is, how you will model it and what will be learned. The style of the project is that you begin by identifying your problem (it needs to be a problem about a biological system). Next, you need to give a "feeling for the numbers" for your problem. Finally, construct a model of the phenomenon which provides intuition and, if possible, makes predictions. Some possible choices for your project are posted with this homework, but feel free to choose your own idea.

2. Mind your media: carbon content of cells

Cells such as *E. coli* can be grown in a liquid solution (medium) that contains various salts and sources of energy. In this problem we will think about "minimal media" which have a single carbon source such as glucose. Start by estimating the number of carbon atoms in an *E. coli* cell. Make sure you carefully explain all of your assumptions. Next, estimate how many sugar molecules your growth medium must deliver for each new cell. Over the period of a typical cell cycle, how many sugar molecules does this mean that the cell is taking on every second? If possible, make some argument about the membrane proteins that bring sugar on board - how many of them might there be and how many sugars do they transport each second?

3. Statistical Mechanics of Gene Regulation.

(a) In class, I derived an expression for the probability that RNA polymerase will be on the promoter of interest in the absence of any transcription

factors. Reproduce the entirety of that argument including the missing algebraic steps that were glossed over in class and show that p_{bound} may be written as

$$p_{bound} = \frac{1}{1 + \frac{N_{NS}}{P} e^{\beta \Delta \epsilon_{pd}}}. \tag{1}$$

Make a log-log plot of the probability that polymerase is bound as a function of the number of polymerase molecules using $N_{ns} = 5 \times 10^6$ and $\Delta\epsilon_{pd} = -5k_BT$. Also, make sure to discuss the implications of this result for a weak promoter, in particular, comment on the basal transcription rate.

(b) Now generalize the problem you did above to the case in which there is a second promoter competing with the promoter of interest. Assume that the binding energy for that site is identical to that of the promoter of interest and derive an expression for p_{bound} for this promoter. Comment on the relative importance of the nonspecific sites and the competing promoter in inhibiting the binding to the promoter of interest.

(c) Müller-Hill, in Oehler et al. (1994), performed a series of impressive measurements of repression in the lac operon for the case in which only a single repressor binding site (the primary operator) was present. In this part of the problem you will reproduce the derivation given in class for the problem of repression, culminating in an expression for p_{bound}. Once you have that expression, use your algebraic expressions to fit the repression measured by Müller-Hill - their results are in the paper posted with this homework. Note that for the purposes of this analysis, we define repression as

$$\text{repression}(R) = \frac{p_{bound}(R = 0)}{p_{bound}(R \neq 0)}. \tag{2}$$

To effect the fit, you will use the measured value of repression and the number of repressors (remember that Oehler *et al.* report the number of repressor monomers - divide by 4 to find the number of active repressors) - this leaves as the only unknown the value $\Delta\epsilon_{rd}$ since you already know the value of $\Delta\epsilon_{pd}$ from our earlier treatment of the problem in the absence of repression and its role when calculating fold-activities. Once you have all of these numbers in hand, make a plot of the "fold-activity" as a function of the number of repressors. In fact, "fold-activity" is the inverse of what we mean by

repression and is defined as

$$\text{fold-activity}(R) = \frac{p_{bound}(R \neq 0)}{p_{bound}(R = 0)}. \tag{3}$$

Make a log-log plot of the fold-activity in the case of pure repression. What features of the curve change by varying parameters such as the binding energy? Examine your expression in the limit when the promoter binding is weak and show that in this case the fold change is given approximately by $F_{reg}(R)$ itself. Work the numbers for the case of interest in this problem and show that this is the appropriate limit.

d) Calculate p_{bound} in the case of an activator "recruiting" RNAP. As we saw in class this process can be described using a binding energy of the activator to DNA ($\Delta\epsilon_{ad}$) and an interaction energy between the activator and RNAP (ϵ_{ap}). Demonstrate that the regulation factor is greater than one and show what happens in the limit that the interaction energy between activator and the polymerase goes to zero. Make a log-log plot of the fold activity in this pure activation case. What features of the curve change by varying the parameters?