APh161: The Physics of Biological Structure and Function
Homework 3
Due Date: Thursday, February 3, 2005

"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*

Reading: Chap. 15 and 17 of ECB. Chap. 2 of **Genes and Signals**, chap. 11 of Schleif.

1. 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 $\Delta \epsilon_{pd} = -5 k_B T$. Make sure that your analysis includes a complete description of how we got the numbers we got - that is, invoke the arguments about equilibrium constants that I gave in class and use the measured values for the specific and nonspecific binding equilibrium constants to check the claims that I made. 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. 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?

(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}). Make a log-log plot of the fold activity in this pure activation case. What features of the curve change by varying the parameters?

(e) Work out p_{bound} for a system in which a repressor competes for promoter binding with RNAP and for which there is an activator which can

assist the binding of RNAP. In addition, write out the regulation factor explicitly for this case. This is an idealization of the case for the lac operon, or alternatively, can be thought of as the appropriate model in which the secondary operators in the lac operon have been deleted. For a reasonable choice of energies make a 3D plot of fold-activity as a function of R and A. Notice that if you just care about the concentration extremes ($A = 0$ and $R = 0$, $A \to \infty$ and $R = 0$, etc.) you can think of this as a logic function. Think of the design possibilities; in particular design your own logic function (AND, OR, etc.) and plot it for a reasonable choice of parameters. You might find useful to see Terry Hwa's work in Buchler et al. (2003).

(f) In this part of the problem, first rederive all of the stuff I did in class having to do with repression in the presence of looping. Look at the table in fig. 4 of Oehler et al. and convince yourself of the fact that there is a cooperative effect. By carefully considering the possible states work out the change in fold-activity in this case. Notice that now you only have one unknown parameter, since the binding energies have already been calculated, and that it should only depend on the separation of the sites, not on the particular sites that have been chosen. Do a fit of your model to Müller-Hill's data and check our claim on its dependence exclusively on the geometry. Check out Vilar and Leibler's (2003) work as well as the papers by Bintu *et al.* to see how this whole story works out.

(g) Consider the case of two activators which bind independently to sites upstream of the promoter and which both have favorable interactions with polymerase without directly interacting with each other. This example is shown as entry 10 in Table 1 of the paper by Bintu et al. Show that for two independent activators that the regulation factor is given by the product of the separate regulation factors. Make a log-log plot of fold-activity as a function of the concentration of A_1 for different values of A_2 and identify the contribution of each activator to the graph's figures.