APh161: Physical Biology of the Cell
Homework 4
Due Date: Tuesday, February 17, 2009

"We must travel in the direction of our fear." – John Berryman

Reading:

Read the first part of chap. 19 of PBoC, pgs. 721-746.

1. Chromosome Geography in Prokaryotes.

(a) Read the paper by Fiebig *et al.* in order to understand the chromosome geography experiments they have performed and write a one paragraph "Scientific American" style description of the experiment.

(b) Repeat in detail with careful explanations all of the steps leading up to eqn. 8.50. This means you need to show your work for determining the Fourier coefficients, explain the normalization in the denominator, etc.

(c) In this part of the problem I would like you to fit the model from part (b) to experimental data for Vibrio cells. First, observe the geometry of the Vibrio cells (though treat them as straight cylinders rather than curved cells) and determine their length and their width. Now, make two plots using the Vibrio data provided with this homework as an Excel spreadsheet. One plot should be of the distribution of the fluorescence marker along the long axis and the other of the distribution along the short axis. Each of the figures should include both the provided data points with error bars and a curve showing the best fit of the model from (b) to the data. To do this fit you will need to use Matlab or Mathematica or some other computer program of your choice.

Your fit should depend upon the parameters x_0 and N (assume the naked DNA value of the Kuhn length). Comment on the best-fit values of those parameters (are they reasonable?), and give a description of how the plots change as the position of the marker gets farther from the tether point. Does

confinement matter along either of these directions? Note that the way I am having you model the data is an approximation. In particular, you are using the one-dimensional solution for the distribution twice, once for each of the two perpendicular directions, but both representing the same polymer. In reality, you would want to solve the full 3D problem (which is much tougher).

2. Statistical Mechanics of Gene Regulation.

- from chap. 19 of PBoC, work out probs. 19.1 and 19.2 (a).

3. Protein Concentration by Dilution.

In this problem we consider the concentration of mRNA or proteins as a function of time in dividing cells. This exercise provides some of the conceptual tools we will need to write down rate equations describing gene expression. In particular, the point of this problem is to work out the concentration of a protein given that we start with a single parental cell that has N copies of this protein (in the experiments of Rosenfeld *et al.* this is a fluorescently-labeled transcription factor). At some point while the culture is growing, the production of the protein is stopped by providing a chemical in the medium and then the number of copies per cell is reduced as a result of dilution as the cells divide.

(a) Work out a differential equation for the change in protein concentration as a function of the time that has elapsed since production of the protein was stopped. Solve the equation and relate the decay constant to the cell cycle time. Note that here we are only interested in the dilution that results from the original N copies of the protein being partitioned into an ever-larger number of daughter cells, not in the dilution that occurs as each individual cell lengthens in preparation for the next round of division. Note also that in this part we're interested in a continuous model—you'll look at the discrete version in part (b). HINT: there are two ways to approach this problem. You can consider the change in the concentration as a function of the change in the number of cells into which the original N proteins are partitioned. Or you can note that for a bacterium like *E. coli*, it is a reasonable assumption to imagine that the cell diameter is unchanged and that the size is controlled by the cell length, such that the change in

volume with time is simply the change in length with time times a constant prefactor; then consider the change in protein concentration as a function of the change in the total volume into which the original N proteins are diluted.

(b) We can repeat a calculation like that given above using a discrete language in which the number of proteins per cell is a discrete integer. Imagine that before cell division, the number of copies of a given transcription factor in the cell is N. In particular, for every cell doubling, the number of proteins is reduced by a factor of 2. Using such a picture, write a formula for the average number of proteins per cell as a function of the number of cell divisions and relate this result to that obtained in part (a). Furthermore, by using the fact that $2 = \exp(\ln 2)$, reconcile the discrete and continuous pictures precisely.

(c) Interestingly, the model used in part (b) opens the door to one of the most important themes in physics, namely, that of fluctuations. In particular, as the cells divide from one generation to the next, each daughter does not really get $N/2$ copies of the protein since the dilution effect is a stochastic process. Rather the partitioning of the N proteins into daughter cells during division follows the binomial distribution. Analyzing these fluctuations can actually lead to a quantification of the number of copies of a protein in a cell. In this part of the problem, work out the expected fluctuations after each division by noting that the fluctuations can be written as $\sqrt{<(N_1 - N_2)^2>}$, where N_1 and N_2 are the number of proteins that end up in daughter cells 1 and 2 respectively. Show that $\sqrt{<(N_1 - N_2)^2>} = \sqrt{N}$ (hint: you'll need to use the binomial theorem.)

Next, look at the Rosenfeld paper and explain how measuring fluorescence variations can be used to calibrate the exact number of copies of the fluorescent protein in a cell. Assume that the fluorescence intensity in each cell can be written as $I = \alpha N$, where α is some calibration factor and N the number of proteins. Make a plot of $\sqrt{<(I_1 - I_2)^2>}$ versus I_{tot} and explain how to get the calibration factor α from this plot.