BE/APh161: Physical Biology of the Cell
Homework 5
Due Date: Wednesday, Feb. 8, 2017

"One of the principal objects of theoretical research in any department of knowledge is to find the point of view from which the subject appears in its greatest simplicity." - Josiah Willard Gibbs

1. MWC Ion Channel: One Equation that Rules Them All

In class, we introduced the idea of allosteric proteins as those that have a regulatory binding site that cause the protein to switch between inactive and active states. We worked out the theory of such allosteric proteins in the specific case of transcription factors such as the Lac repressor. However, the allostery idea has much broader reach. In this problem, we will take the same ideas developed in class and apply them to the so-called ligand-gated ion channels. These channels are relevant in contexts ranging from our neuromuscular junctions to the photoreceptors in our eyes to olfactory neurons. Figure 1 shows two classic examples of these channels.

(a) Write a paragraph that summarizes the function of the two ion channels shown in Figure 1. The point here is just to make sure you have a little understanding of their physiological function before we start working out their statistical mechanical properties.

(b) Make a diagram with your version of the statistical mechanics protocol showing the states and weights for the nAChR ion channel. Make sure you explain all of your notation for the parameters that appear here.

(c) Write an equation for the probability that the channel is open $p_{open}(c)$, where c is the concentration of acetylcholine. How does this equation compare to the equation we wrote for the activity of the Lac repressor as a function of inducer concentration?

(c) Work out the leakiness, dynamic range and the EC50. Leakiness refers to the probability that the channel is open in the absence of ligand and can be thought of as p_{min}, the minimum probability the channel is open. Dynamic range refers to the difference between p_{max} and p_{min}, where p_{max} is

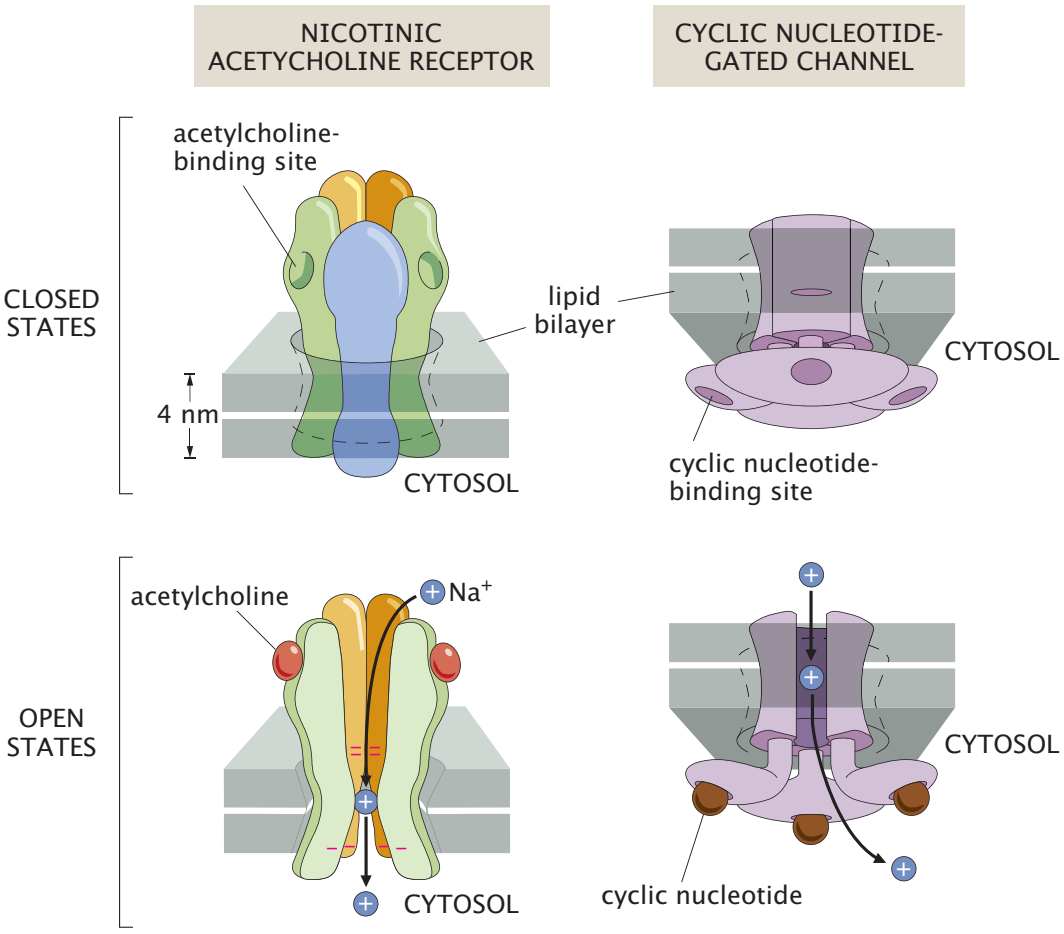

Figure 1: Key examples of ligand-gated ion channels. (A) Nicotinic acetylcholine receptor, revealing it's heteropentameric structure with two binding sites for acetylcholine. (B) cGMP-gated ion channel. These channels have four cGMP binding sites.

the probability of being open at saturating concentrations of ligand. Find explicit expressions for both p_{min} and p_{max} and then use their difference to obtain the dynamic range. EC50 is the concentration of ligand at which the channel is halfway between p_{min} and p_{max}.

(d) Figure 2 shows data for the wild-type nAChR ion channel from the laboratory of our own Prof. Henry Lester. With your TAs, use digitizeit to extract the data and then make a fit using the MWC model you worked out earlier in the problem. This is Figure 1B of the paper by Labarca *et al.* included with the homework. Note that unfortunately, they chose to plot "normalized current" rather than $p_{open}(c)$. As a result, your fit will have to be to the normalized current given as

$$\text{normalized current} = \frac{p_{open}(c) - p_{min}}{p_{max} - p_{min}}. \tag{1}$$

I am excited for you to learn how to use Digitizeit because it is liberating: with it, you can take figures from anyone's papers and grab their experimental data and export it into a spreadsheet so that you can unleash your theoretical analysis on it.

2. The Standard Candle: Counting Proteins with Partitioning Statistics.

(a) Begin by reading the paper by Rosenfeld *et al.* entitled "Gene Regulation at the Single-Cell Level" (posted on the website with the homework) and write a one paragraph commentary on the paper with special reference to how they used the binomial partitioning as a way to count repressor proteins. What is the experiment they did and what were they trying to learn?

In the rest of the problem we work out for ourselves the ideas about binomial partitioning introduced in the Rosenfeld *et al.* paper in order to consider the concentration of mRNA or proteins as a function of time in dividing cells. In particular, the point of this problem is to work out the concentration of mRNA or protein given that we start with a single parental cell that has N copies of this mRNA or protein (in the experiments of Rosenfeld *et al.* this is a fluorescently-labeled transcription factor). In the Rosenfeld experiment, 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

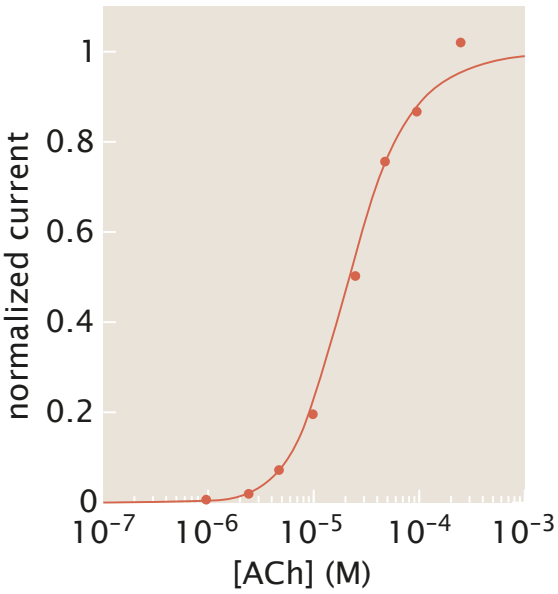

Figure 2: Ion channel currents as a function of ligand concentration. (Adapted from Labarca *et al.*, Nature, 1995).

copies per cell is reduced as a result of dilution as the cells divide.
Interestingly, this problem 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.

(b) If we think of the N copies of the protein as being divided between the two daughters with N_1 going to daughter 1 and $N - N_1$ going to daughter 2, write the probability distribution $p(N_1, N)$. Next, work out the expected fluctuations in the partitioning process 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 $\sqrt{< (N_1 - N_2)^2 >} = \sqrt{N}$. When I do this calculation, I find it convenient to write $N_2 = N - N_1$. Basically, this reduces the problem to having to calculate $\langle N_1 \rangle$ and $\langle N_1^2 \rangle$ since once you have those two quantities you can

evaluate $\sqrt{< (N_1 - N_2)^2 >}$.

(c) 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. Specifically, assume that the fluorescence intensity in each cell can be written as $I = \alpha N$, where α is an as-yet unknown calibration factor and N the number of proteins in the cell. Explain what this equation means and why you think it is justified. Derive an expression relating I_1, I_2 and I_{tot} using the result of part (b). 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.

(d) Now we are going to repeat the Rosenfeld experiment numerically in order to *fit* the calibration factor. Consider a fluorescent protein such that the calibration factor between the intensity and the number of fluorophores is 50, that is $I = 50N$. Generate intensity data by choosing $N_1 + N_2 = 10, 50, 100, 1000$ and 5000 and for each case, "partition" the proteins from the mother cell to the two daughters 100 times (i.e. as if you are looking at 100 mother cells divide for each choice of the protein copy number). Then, make a plot of the resulting $\sqrt{< (I_1 - I_2)^2 >}$ vs I_{tot} just as we did analytically in the previous problem. What I mean is that you need to make a plot of all of your simulation results. Then, do a fit to your "data" and see how well you recover the calibration factor that you actually put in by hand. Plot the fit on the same graph as all of the "data".