BE/APh161: Physical Biology of the Cell
Homework 2
Due Date: Wednesday, January 20, 2021

"Whatever you can do, or dream you can do, begin it. Boldness has genius, power and magic in it." - Goethe

Extra Credit. Provide comments on chap. 4, "Thinking Big About Data" of the upcoming third edition of *Physical Biology of the Cell*. Note that this is an unfinished draft of the chapter. Figure placements are not necessarily correct and there are still a number of internal discussions amongst the author team about how to finish things off. We are especially interested in mistakes, flaws in logic, confusing figures, unclear discussions, etc., but are happy to entertain comments at all scales. This extra credit will constitute an additional 15% on your score on the homework.

1. A concentration rule of thumb.

(a) If there is one molecule of a given species within a bacterium, what is that concentration?

(b) As an application of this idea, how many H^+ ions are there in a bacterial cell if the pH is 7.0?

2. Post-Translational Modifications and "nature's escape from genetic imprisonment"

In a very interesting article ("Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding"), Prof. Jeremy Gunawardena discusses how we should think about post-translational modifications as a way of expanding the natural repertoire of the 20-letter amino acid alphabet. Similarly, Prof. Christopher Walsh (also at Harvard) wrote a whole book entitled "Posttranslational Modifications of Proteins: Expanding Nature's Inventory", again making the point that by adding chemical groups to proteins we can significantly change their properties.

(a) Provide at least one mechanistic idea about how adding a chemical group to a protein can alter its structure or function. Your answer should be offered in less than a paragraph, but should be concrete in its assertions about how these modifications change the protein. Why does Gunawardena refer to this process of post-translational modification as "escape from genetic imprisonment"?

(b) As a toy model of the combinatorial complexity offered by post-translational modifications, let's imagine that a protein has N residues that are able to be phosphorylated (NOTE: please comment on which residues these are - the answer is different for bacteria and eukaryotes). How many distinct states of the protein are there as a result of these different phosphorylated states? Make an approximate estimate of the mass associated with a phosphate group and what fraction of the total mass this group represents. Similarly, give some indication of the charge associated with a phosphate group. What ideas do you have about how we can go about measuring these different states of phosphorylation?

(c) In this part of the problem, we make a very crude estimate of the number of sites on a protein that are subject to phosphorylation. To do so, imagine that the protein is a sphere with N residues. How does the radius of that sphere depend upon the number of residues in the protein? Given that estimate, what is the number of residues that are on the surface? Given that number, what fraction of those are phosphorylatable? Remember, these are crude estimates. Work out these results for a concrete case of a typical protein with roughly 400 amino acids.

(d) Let's close out these estimates by thinking about a bacterial cell. If all 3×10^6 proteins in such a cell can be phosphorylated with the number of different phosphorylation states that you estimated above, how many distinct cells could we make with all of these different states of phosphorylation.

3. Phosphorus, Sulfur and the Lives of Cells

In addition to the big ticket chemical elements in cells (carbon, nitrogen, oxygen, hydrogen), other elements come in at lower concentrations, but still with enormous functional importance. Two such elements are phosphorus and sulfur and in this problem, we will try to figure out how much of the

cell's dry weight is taken up by these elements and what this implies about the transport of these elements into the cellular interior. A useful vignette to watch in order to do this problem is "Rate-Limiting Hypothesis - Carbon Transport."

(a) Let's begin by trying to estimate the number of phosphorus atoms in a cell. Where do we find phosphorus? There is 10 mM of ATP in a typical bacterium. We all know that in both RNA and DNA, every base carries its own phosphate. Many lipids are phospholipids, with polar heads containing phosphate atoms as well. Proteins are phosphorylated (see problem 1!!). Don't forget ribosomes. They too are full of phosphorus atoms because they are 2/3 by mass RNA. Given these various facts, estimate the total number of phosphorus atoms in a bacterium. Given a division time of $f \times 10^3$ s, how many phosphate transporters (PitA) are needed to bring all those phosphorus atoms into the cell during that time?

(b) Next, we consider sulfur. Where do we find sulfur atoms in cells? Clearly one of the main amino acids, cysteine, has its known covalent binding properties precisely because of its sulfur atom. The metabolite glutathione has a concentration of 17 mM. Like in the previous part of the problem, in light of these facts, make an estimate of the total number of sulfur atoms in a bacterial cell. Given a division time of $f \times 10^3$ s, how many sulfur transporters (CysUWA) are needed to bring all those sulfur atoms into the cell during that time?

4. DNA replication rates.

Do problem 3.3 of PBoC2. However, as you do this problem, please come at it a few different ways. First, when estimating how much of the full fly genome is shown in the figure, account for the fact that the DNA is compacted by nucleosomes. Second, given that the entire fly genome has been claimed to have ≈ 6000 origins of replication, figure out the mean spacing between such origins and use that estimate as the basis of your own independent estimate of the replication time for the *Drosophila* genome.

5. Migration of the bar-tailed godwit

Animal migrations are one of the greatest of interdisciplinary subjects, bring-

ing together diverse topics ranging from animal behavior to the physics of navigation to the metabolism required for sustained long-distance travel. The bar-tailed godwit is a small bird that each year travels between Alaska and New Zealand on the same kind of incredible nonstop voyage taken by happy tourists in modern long-distance jetliners. During a visit to New Zealand's South Island, one of us had the chance to see these amazing birds in Okarito Lagoon with a naturalist guide who claimed that over the course of their ten-day, ten-thousand kilometer trip, these migratory birds lose 1/3 of their body mass. In this problem, we make a series of simple divide-and-conquer estimates to see whether this claim might be true. A useful vignette to watch to do this problem is "Semantics On Scale."

(a) Using dimensional-analysis arguments, work out how the drag force experienced by flying godwits depends upon the density of air, the speed of the birds and the size of the birds. Specifically, work out the coefficients α, β and γ in the expression

$$F_{drag} = \text{const.}\rho^{\alpha} v^{\beta} L^{\gamma}. \tag{1}$$

(b) Work out the power expended by the bar-tailed godwit to overcome the drag force. Then, work out the total energy expended during the ten-day migration in overcoming this drag force.

(c) Given that burning fat yields 9 kcal/g, work out the number of grams of fat that would need to be burned to sustain the ten day flight of the bar-tailed godwit.

6. Protein density on the membrane vs in the cytoplasm

Which is larger, the protein concentration in membranes or in the cytoplasm? Report both of them in units of number per μm^3.

7. Laws of Cellular Growth Dynamics.

Much of our understanding of the natural world is couched in the language of the subject now known as "dynamical systems." In a nutshell, the idea is

to write down equations that tell us how some variable(s) of interest change in time. Often, this ends up being written in the form of coupled differential equations. Perhaps the most important and simplest of such dynamical systems is the law of exponential growth (or decay), relevant to thinking about the early stages of growth of a culture of cells, for example. In my recorded vignette "Laws of Cellular Growth," I give a brief introduction to the way we can write dynamical equations that describe the evolution of the size of a population of cells ($N(t)$) as a function of time. In this problem, you are going to revisit the discussion I give there by solving for the dynamics of a population of bacterial cells both analytically and numerically.

(a) In this first part of the problem, write down and justify the differential equation for a population of dividing bacteria that is not limited by nutrient availability. Write an analytic solution. In addition, write a code in Python that integrates the equation over time. The basic idea for solving a differential equation of the form

$$\frac{dx}{dt} = f(x,t), \tag{2}$$

is time stepping. Here we propose you use the most naive method which instructs us to write the solution at time t as

$$x(t) = x(t - \Delta t) + f(x, t - \Delta t)\Delta t. \tag{3}$$

The structure of for loops permits us to solve this numerically for any "well behaved" $f(x,t)$. Use this algorithm to solve your growth equation numerically and plot your numerical solution on the same graph as the analytic solution.

(b) The logistic growth equation is a phenomenological way of curbing unchecked growth and is written as

$$\frac{dN}{dt} = kN\left(1 - \frac{N}{K}\right), \tag{4}$$

where K is the so-called carrying capacity of the population. As in part (a), find an analytic solution for this equation and then find a numerical solution as well. Estimate the parameters relevant to your solution by thinking about a saturated bacterial culture. Comment on the meaning of the carrying capacity.

8. The pandemic elephant in the room.

(a) What is the information density of the SARS-CoV-2 virus? What I mean is that there are a certain number of bits of information contained in the viral genome, so you can report a density of bits/nm^3.

(b) What is the information density of a typical hard drive for backing up our laptops?

(c) Given your answer to part (a), how many SARS-CoV-2 viruses would it take to capture all the information in the Library of Congress? How much volume would such a "library" take up? Could the whole Library of Congress fit into one 5 mL tube, a one L flask, one shelf of a -80 freezer?