BE/APh161: Physical Biology of the Cell
Homework 2
Due Date: Wednesday, January 22, 2020

"Doubt is the father of creation." - Galileo Galilei

1. Number of mRNA in a cell.

In this problem, we are going to work our way through an estimate of the number of mRNA molecules found in a bacterium and in a yeast cell. The idea of the estimate is to try to figure out over the entire set of genes in the organism, how many *total* copies of mRNA will be found in the cell. To do the estimate, we will first consider the case of a bacterium and then for yeast, we will make the assumption that things play out the same way and simply scale up our bacterial estimate. Our starting point is the number of proteins in a cell, which for a bacterium we take to be 3×10^6. This means that in order to make a new cell, this many proteins have to be synthesized in the 1000-3000 s of the cell cycle (depending upon growth conditions). If the ribosome translates at a rate of 20 aa/s, figure out a range of values for how many proteins can be synthesized from each mRNA per minute. The range comes from how tightly packed the ribosomes are. What is the highest rate at which translation could occur (hint: think about the size of the ribosome and how tightly packed they can be)? Now use this to estimate the total number of mRNAs that are needed to supply the protein needed during a cell cycle. Provide estimates for both bacteria and budding yeast.

2. A feeling for the numbers: the chemical composition of a cell by pure thought

Make an estimate of the composition of carbon, hydrogen, oxygen, and nitrogen in the dry mass of a bacterium. Using knowledge of the size and mass of a bacterium, the fraction of that mass that is "dry mass" (that is, $\approx 30\%$) and the chemical constituents of a cell, figure out the approximate small integers (<10) for the composition $C_mH_nO_pN_q$, that is, find m, n, p, and q.

3. The Benjamin Franklin Problem

In his travels between America and Europe, Benjamin Franklin was subjected to the vicissitudes of the sea which led him to reflect on his reading of Pliny the Elder and claims of how oil was known to smooth the waves. Upon arriving in England, Franklin took the concept to the test. He tells us of his experience thus: "At length at Clapham where there is, on the common, a large pond, which I observed to be one day very rough with the wind, I fetched out a cruet of oil, and dropped a little of it on the water. I saw it spread itself with surprising swiftness upon the surface... the oil, though not more than a teaspoonful, produced an instant calm over a space several yards square, which spread amazingly and extended itself gradually until it reached the leeside, making all that quarter of the pond, perhaps half an acre, as smooth as a looking glass."

(a) Though Franklin himself never made the estimate (that was to await Lord Rayleigh), use Franklin's description of the experiment to work out the thickness of the oil film (the height of a lipid!) that covered the surface of Clapham common pond.

(b) Using a typical molecular mass for a lipid (say, 1000 g/mol), work out the number of lipid molecules that covered that surface of the pond and use that number to compute the area per lipid. How do your results compare to the modern values for the size of lipids?

4. Cell cycle in *E. coli*.

During our first week of class, we spent a lot of time talking about time scales of various processes in biology. In fact, I like to think of the cell cycle time as the standard stopwatch of biology. That is, cells are the individual "quanta" of biology and the time scale for one cell to make another reflects perhaps the most important process undertaken by cells. A particularly beautiful and fun class of experiments on cellular dynamics are those in which time-lapse imaging is carried out on a microscope resulting in series of images like those shown in Figure 1. Your task is to analyze one of these movies in order to measure the cell division time for *E. coli*. There will be a help session offered by the TAs that will explain how to load the files and segment them to find the number of cells as a function of time. Make sure you show how to connect

to the growth equation

$$\frac{dN}{dt} = kN. \tag{1}$$

Specifically, how is k related to the division time? In your estimation, use the fact that the time between consecutive frames is 5 min.

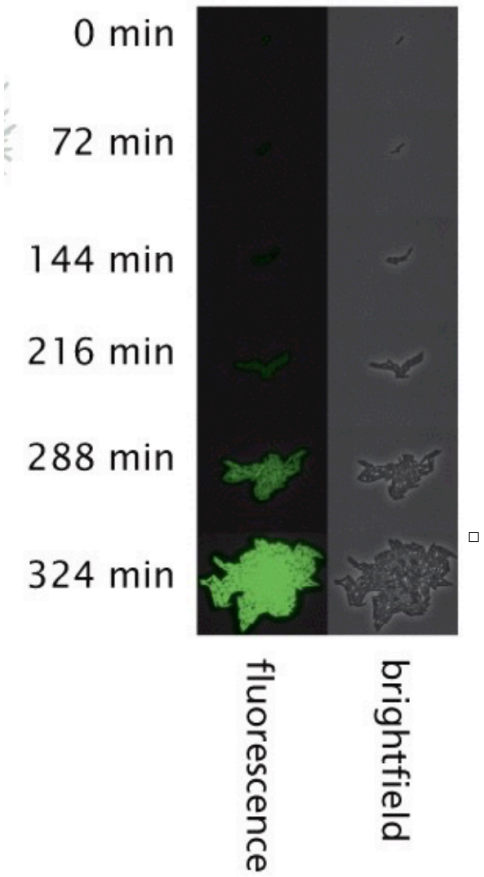

Figure 1: Cell cycle of *E. coli*. Phase microscopy and fluorescence microscopy images of growing *E. coli* cells.

5. Post-Translational Modifications and "natures escape from genetic imprisonment"

In a very interesting article ("Post-translational modification: natures 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 pro-

tein 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.

6. Estimating the diffusion constant

In this problem, we are going to use the observed trajectories of diffusing GFP molecules to estimate the diffusion coefficient.

(a) Conventional microscopy to observe individual fluorescent proteins won't work. In this part of the problem, we are going to work out why. During a traditional experiment, the microscope shutter is open during some time interval of order 10s to 100s of milliseconds. By assuming a diffusion constant of 10 μm^2/s, work out how far the fluorescent protein will move during the time that the shutter is open and compare that distance to the size of the cell itself and comment on how this limits our ability to measure the diffusion constant. Perform the estimate a second time, this time using the 0.3 ms exposure time shown in Figure 2(A).

(b) Use the trajectories shown in Figure 3 and our simple rule of thumb that $t_{diffusion} = L^2/D$ to estimate the diffusion constant for GFP. Explain your reasoning carefully.

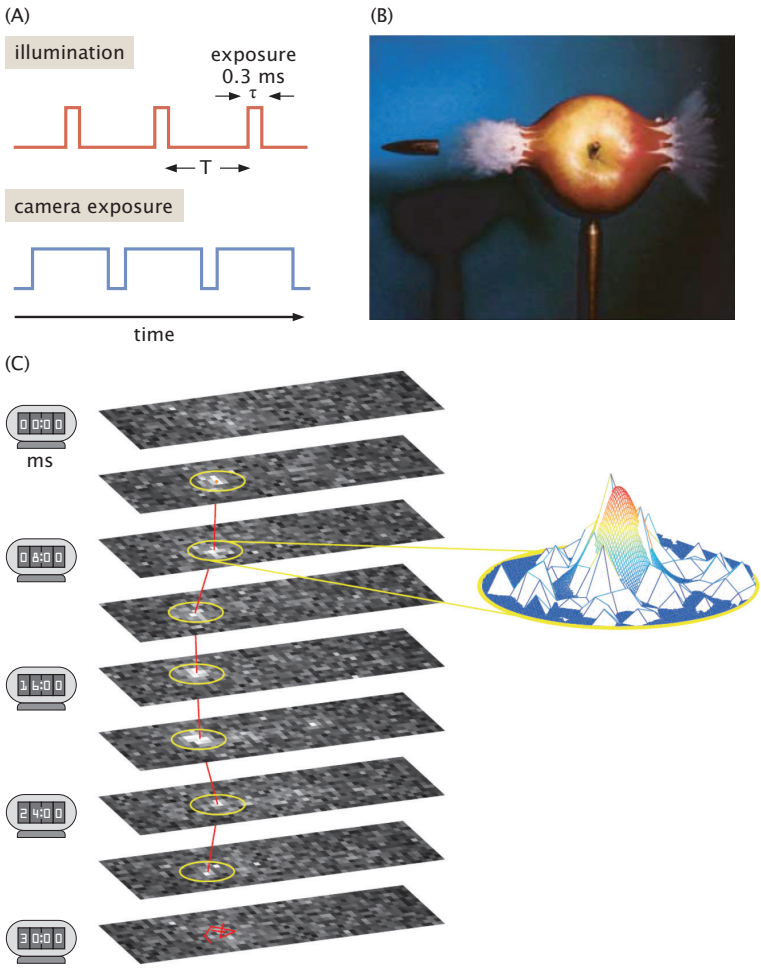

Figure 2: Stroboscopic illumination for high speed photography. (A) Comparison of the pulse of light and the camera exposure times. Brief illumination periods guarantee that the diffusing molecule doesn't move very far. (B) A classic photo from MIT legend Harold Edgerton who pioneered stroboscopic photography for science and fun. (C)

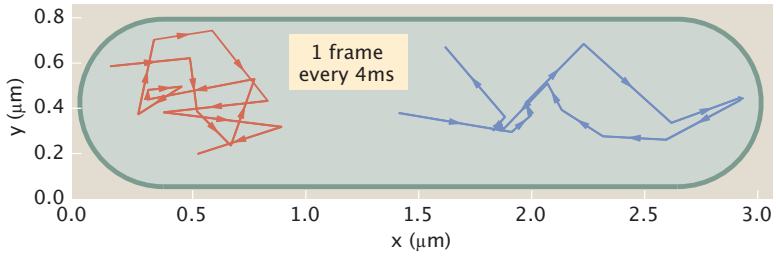

Figure 3: Time series showing positions of diffusing GFP molecules at different times. The red and blue traces correspond to different molecules. The lines are a guide to the eye.