

BE/APh161: Physical Biology of the Cell

Homework 2

Due Date: Wednesday, January 18, 2017

“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. Diffusion

We have been interested in making estimates of the time scale associated with a number of biological processes. One of the most ubiquitous processes is diffusion, which is our “go to” null hypothesis for how molecules get around in cells.

In class I noted that the time scale for diffusing a distance L is given by $t = L^2/D$, where D is the diffusion constant. In this problem, we will formally derive this result. Note that parts (a) and (b) are effectively problem 13.2 of PBoC. Also, reading much of chap. 13 of PBoC will be very helpful

for doing this problem.

(a) Our goal is to find the diffusive profile for some molecular species as a function of time. If we are given an initial concentration, we can use the diffusion equation to determine the concentration distribution at a later time. To that end, consider the one-dimensional diffusion equation in free space given by

$$\frac{\partial c(x, t)}{\partial t} = D \frac{\partial^2 c(x, t)}{\partial x^2}. \quad (1)$$

In particular, consider that the initial concentration distribution is given by $c(x, 0) = \delta(x)$, where $\delta(x)$ is the Dirac delta function and basically means that there is a spike at the origin. In particular, you will show that

$$G(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}, \quad (2)$$

where we introduce the Green function $G(x, t)$ to signify that this is the concentration profile for the special case in which the initial concentration is the spike at the origin as represented by the delta function.

To obtain the solution, we will Fourier transform the diffusion equation in the spatial variable x according to the Fourier transform convention

$$\tilde{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx, \quad (3)$$

and

$$f(x) = \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} dk. \quad (4)$$

Using these definitions, Fourier transform both sides of the diffusion equation to arrive at the ordinary differential equation

$$\frac{d\tilde{c}(k, t)}{dt} = -Dk^2 \tilde{c}(k, t). \quad (5)$$

Note that to do this Fourier transform of the right side, you will need to use integration by parts twice. Solve the differential equation that emerges from your Fourier transform to obtain $\tilde{c}(k, t)$ and make sure to use the initial condition $c(x, 0) = \delta(x)$ to find $\tilde{c}(k, 0)$. Then invert the Fourier transform on $\tilde{c}(k, t)$ to find $c(x, t)$. NOTE: You will need to use completion of the square to carry out the inversion. Make sure you explain all of your steps. We are

big on having you not only do the analysis correctly, but also to explain what you are doing and why you are doing it. Also, explain why I said this is the solution for “free space”. Why would this solution fail to describe diffusion in a finite box?

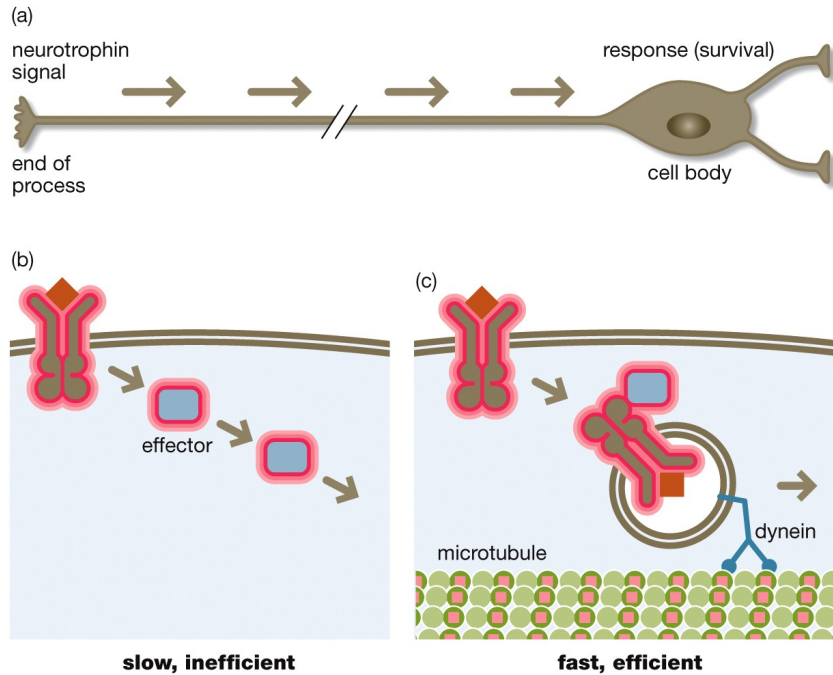


Figure 5.17 Cell Signaling (© Garland Science 2015)

Figure 1: Comparison between passive diffusion and active transport in neurons. (a) Schematic of a neuron. (b) An effector molecule is activated and then diffuses along the axon to the cell body. (c) Receptor is incorporated into a vesicle and then actively transported by a dynein molecule along a microtubule.

(b) Using the solution we obtained above, find $\langle x \rangle$ and $\langle x^2 \rangle$. In general, we have that

$$\langle x^n \rangle = \frac{\int_{-\infty}^{\infty} x^n c(x, t) dx}{\int_{-\infty}^{\infty} c(x, t) dx}. \quad (6)$$

Explain what you find for both the first and second moments of the distribution as a function of time and explain how it relates to the estimated diffusion

time $t = L^2/D$ which we use to find the time scale for diffusion over a length L . Using the Einstein-Stokes relation given by

$$D = \frac{k_B T}{6\pi\eta a}, \quad (7)$$

where η is the viscosity which for water is $\eta_{water} = 10^{-3} Pa s$ and a is the radius of the diffusing particle, estimate the diffusion constant for a protein in water and make a log-log plot of diffusion time vs distance (with distances ranging from 1 nm to 1 μ m) and comment on its biological significance. Also, make a plot of the solution for the point source as a function of time by showing $c(x, t)$ at various times t using the same diffusion constant.

(c) In their book “Cell Signaling”, Lim, Mayer and Pawson give the classic story about diffusion in neurons and how diffusion will take prohibitively long times. See Figure 1 for their depiction of the comparison between passive diffusion and active transport. Using what we have learned about diffusion, work out the time for diffusion of a protein over the 10 cm length of a neuron. Compare this to the time for a molecule to be transported actively by a motor. Do you agree with their assessment that active transport is efficient?

3. 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 2. 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

$$\frac{dN}{dt} = kN. \quad (8)$$

Specifically, how is k related to the division time?

4. 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.

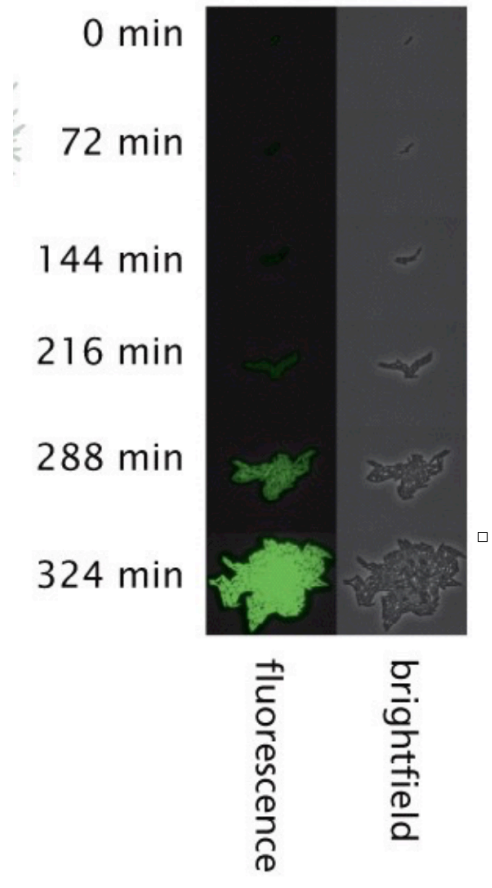


Figure 2: Cell cycle of *E. coli*. Phase microscopy and fluorescence microscopy images of growing *E. coli* cells.