BE/APh161: Physical Biology of the Cell
Homework 1
Due Date: Wednesday, January 14, 2015

"Satisfaction of one's curiosity is one of the greatest sources of happiness in life." - Linus Pauling

Comments from RP to Class

In my view, homeworks are one of the primary tools at a teacher's disposal to push an educational agenda. I am big on sports analogies, and the simple fact is this: if you want to shoot free throws like Kobe Bryant you have to practice your craft. It isn't always amusing, but it pays off later. As a result, I put lots of time into both thinking up problems and writing text that goes along with those homeworks that I think will give you the opportunity to practice things that will help you do science better later when it is "for real". The reading that goes along with my homeworks is a key part of the course material, so please read my commentaries (and argue with them if you have an alternative perspective).

This first homework has as its main objective the development of a feel for the numbers associated with various biological problems. This particular homework will probably involve more searching around on the web than others. Please make sure to report your sources.

1. Who Are You?

It is said that you have more foreign cells in your body than those containing your own DNA. Make a simple estimate of the number of human cells you are made up of, the number of bacterial cells you harbor in your gut (assume there are 2 kg of bacteria in your gut), the number of human genes you carry and the number of genes associated with the more than 200 different species of bacteria you are carrying around. Also, by appealing to the so-called "Complete Blood Count" (CBC - the blood test that nearly all of us have had) and characterized in Table 4.1 of PBOC2 (our textbook), estimate what fraction of our cells are red blood cells and what fraction of our cells

are white blood cells.

2. RNA vs Protein

Using the kind of estimates we have talked about in class, give a simple characterization of the relative sizes of mRNAs and the proteins they code for. Specifically, first comment on the mean mass of amino acids and nucleotides as well as their typical physical sizes. Use both of these metrics as a way to provide a rough sense of the relative sizes both in mass and physical dimensions of proteins and the mRNAs that code for them.

3. 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 protein with roughly 400 amino acids.

4. To Build a Cell.

The idea for this problem is to imagine the costs of building up a cell from pure thought. That is, by knowing just a few key facts, what can we say about the chemical requirements to build a cell? Given that cells grow in media that has a carbon source such as glucose and a nitrogen source such as ammonium chloride, we will try to figure out from the recipe of one of these media how many cells can grow there before the resources are exhausted.

Do Problem 2.5 of PBOC2. This problem will walk you through the analysis of the nutrients provided in growth media and how they support the synthesis of proteins, nucleic acids and lipids as cells grow.

5. 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.

6. RNA Polymerase and Rate of Transcription.

Do problem 3.4 of PBoC2.