APh161: The Physics of Biological Structure and Function
Homework 1
Due Date: Tuesday, January 18, 2005

"A traveler who refuses to pass over a bridge until he has personally tested the soundness of every part of it is not likely to go far; something must be risked, even in mathematics." – Horace Lamb

Reading: Read chap. 1 of Schleif.

1. A feeling for the numbers: microbes as the unseen majority

One of the key arguments that I will make throughout the course is that sometimes just having a feel for magnitudes is a useful guide to intuition.

(a) Justify the assumption that a typical (i.e. *E. coli*) bacterial cell has a volume of 1 μm^3. Also, express this volume in femtoLiters. The claim is made (see the paper on prokaryotes as the unseen majority on the course website) that in the top 200 m of the world's oceans, there are roughly 10^{28} prokaryotes. Work out the total volume taken up by these cells in m^3 and km^3.

(b) In his famed book *The Story of Mankind*, Hendrik Willem van Loon makes an amazing estimate of the size of a box that all of the humans from all of history would have fit into. I read this when I was about 15 years old and found it really odd, but cool! Your task is to work out the size of a box that would hold (close packing) all of the current human inhabitants of the Earth. Compare this number to the volume of the box that will hold all of the bacteria in the top 200 m of the oceans.

(c) Also, recall that roughly 2-3 kg of bacteria are to be found in the waste factory of your large intestine. Make an estimate of the total number of bacteria inhabiting your intestine and then all of the intestines of all of the humans currently on the Earth.

(d) On the course website, there is a fascinating paper by Zimmerman and Trach in which they attempt to measure the crowding in the cellular interior. In table 3 they tell us their estimated macromolecular concentrations in the cellular interior. Use these numbers to make an estimate of the mean protein spacing. Also, estimate the number of lipid molecules in the cell membrane of *E. coli* by computing the approximate area of the cell membrane and dividing by the area taken up by each lipid, the total number of protein molecules and the number of water molecules.

(e) If a particular protein in an *E. coli* cell is found there at nM concentrations, how many molecules are there per cell? Are you happy with the notion of a "concentration" in this case? Explain your reasoning. Make a plot of the number of copies of a molecule in an *E. coli* cell as a function of the concentration - make the plot for concentrations from nM to mM.

2. A Feeling for the Numbers: The Scales of Things

In class, I tried to convey a sense of the meaning of some of the key length scales that will be on constant duty during this course. The goal of this problem is to try and get a sense of the *relative* sizes of the various length scale units we have introduced.

For the purposes of this problem, scale up a typical protein so that it is the size of an apple. In such a world, use familiar objects or distances from everyday life (i.e. the size of a house, the distance between New York and LA, etc.) to characterize the sizes of:

(a) the ribosome

(b) a bacteriophage

(c) an *E. coli* bacterium

(d) a typical eukaryotic nucleus

(e) a eukaryotic cell

(g) a fruit fly

(h) a human

3. Manipulating Atomic Coordinates

Visualization of the various structures populating the cell is a key part of fulfilling the objective of structural biology to connect structure and function. In this problem, you will learn how to manipulate pdb files from the Protein Databank and to view them using one of the various plotting programs.

(a) Obtain coordinates for ATP, phosphatidylcholine, B-DNA, G-actin, the lambda repressor/DNA complex or lac repressor/DNA complex, myoglobin, green fluorescent protein (GFP) and RNA polymerase. You can do this by visiting sites such as "http://chemistry.gsu.edu/glactone/PDB/pdb.html" and the Protein Databank itself. You may have to search around a bit. Give a brief description of each one of these molecules and its role in cellular life.

(b) Download a structural viewing code such as VMD (University of Illinois), Rasmol (University of Massachusetts) or DeepView (http://www.expasy.ch/spdbv/) and create a plot of each of the molecules you downloaded above. Experiment with the orientation of the molecule and make sure you print out pictures of each and every molecule.

(c) Later we will see that phosphatidylcholine is one of the molecules that can self assemble to form a lipid bilayer (see chap. 11 of Alberts et al.). Part of our analysis of such structures will be to consider their geometry. As a first step down that path, estimate the cross sectional area of the polar head of phosphatidylcholine. Make sure you get the coordinates for this molecule and plot it as well. Note that you will find this result useful for the estimate of the number of lipids in an *E. coli* cell in problem 1.

(d) ATP is the energy currency for many processes in biochemistry. The action of ATP is mediated by ATP binding onto other molecules which then exploit the energy associated with hydrolysis of ATP. Use your coordinates for ATP to estimate the size of the regions in which ATP might bind when

it encounters other molecules.

4. A Feeling for the Numbers: Kilodaltons and Beyond

The description of proteins is partially advanced by referring to the number of kilodaltons (i.e. number of atomic mass units) making up that protein. For example, G-actin is a 42kD protein, while the motor molecule myosin is an enormous 520kD protein (made up of several different domains). The goal of this problem is to develop some rules of thumb for thinking about these numbers and to use these rules of thumb to generate a feel for protein size (both mass and geometry).

(a) Generate an estimate for the size of a "typical" amino acid in daltons. Justify your estimate by explaining how many of each type of atom you chose. Also, to quantify the goodness of your estimate, consider the actual masses of several key amino acids such as glycine, proline, arginine and tryptophan. Report these masses in daltons also.

(b) On the basis of your result for part (a), deduce a rule of thumb for converting mass of a protein (reported in kD) into a corresponding number of residues. Apply this rule of thumb to myosin, G-actin, hemoglobin and hexokinase. How well does the rule of thumb compare to the exact values for the number of residues in each of these proteins?

(c) Generate an estimate for the spatial size of a typical amino acid. You may choose to work either in $\mathring{A}^3$ or nm^3. Similarly, work out a rule of thumb that allows for a conversion between a statement of the mass of a protein in kD and its volume.

(d) Apply the rule of thumb from part (c) to the same set of proteins you considered in part (b).

(e) As a point of amusement, provide brief explanations concerning the function of each of the proteins you have considered in this problem.

5. Read and Red Ink

Read chap. 2 of the book that I am writing with J. Kondev and comment on the draft using Adobe Acrobat and email your comments to me . Please comment on anything and everything - logic, clarity, figures, the use of English, etc. The draft for the chapter will be put on the course website and *not* handed out in class.