Reading: Read chap. 2 of ECB, especially the parts later in the chapter on the noncovalent interactions central to macromolecular structure and function. Look especially at Panel 2-7.

1. Scaling of Protein Size.

The scaling of a polymer’s size as a function of the number of monomers is one of the central results to emerge from simple lattice models of polymers. The goal of this problem is to investigate the extent to which such arguments are in fact appropriate for biological polymers, and in particular, proteins. To that end, use the Protein Databank in order to download the coordinates for a variety of globular proteins, including myoglobin, hemoglobin, bovine pancreatic trypsin inhibitor (BPTI), lysozyme, cytochrome c, G-actin and tubulin. (see pg. 150 of Alberts et al. to get a preliminary feel for these proteins).

(a) Using whatever graphical package you prefer, make plots of all of these structures and comment on their geometric makeup, with special reference to the elements of their secondary structure.

(b) Give a few lines of description of the role of each of these proteins in living systems.

(c) In each case, compute the radius of gyration and then make a single plot which shows the radius of gyration for each of these proteins as a function of the number of residues in the protein.

(d) On the same plot given above, show the scaling of the radius of gyration for models of the polymer which are a) space filling and b) scale in the way suggested by the simplest random walk model, and comment on the
comparison between these simple scaling estimates and the real structures.

2. The Frances Arnold Estimate Problem

In a 2001 Bioengineering seminar, Professor Frances Arnold made a startling remark that it is the aim of the present problem to examine. The basic point is to try and generate some intuition for the **HUGE, ASTRONOMICAL** number of ways of choosing amino acid sequences. To drive home the point, she noted that if we consider a protein with 300 amino acids, there will be a huge number of different possible sequences.

(a) How many different sequences are there for a 300 amino acid protein?

But that wasn’t the provocative remark. The provocative remark was that if we took only one molecule of each of these different possible proteins, it would take a volume equal to five of our universes to contain all of these different distinct molecules.

(b) Estimate the size of a protein with 300 amino acids. Justify your result, but remember it is an estimate. Next, find an estimate of the size of the universe and figure out whether Frances was guilty of hyperbole or if her statement was on the money.

3. A Feeling for the Numbers: Statistical Configurations of Macromolecules

A shocking feature of macromolecules when viewed from the statistical perspective is the number of different configurations available to such systems. We have already examined the astronomical variety of primary structures available to proteins in the Frances Arnold problem. Now we turn to an examination of the number of different spatial configurations of macromolecules. As a simplifying assumption, we assume that our polymer is described via a lattice model in which each monomer is situated at a particular site on the lattice and different sites are connected by bond vectors which can point along any of the directions \( \{ \mathbf{e}_1, -\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_2, \mathbf{e}_3, -\mathbf{e}_3 \} \).

(a) For a polymer with \( N \) monomers, make an estimate for how many
different possible configurations there are?

(b) Comment on any suspect assumptions we may have made in the estimate given in part (a). In particular, think about whether all of the configurations you counted are really legitimate configurations.

4. Entropic Spring in Three-Dimensions Revisited

In class, we have already described entropic springs from several different perspectives. In this problem, you will work out the force-displacement relation for a simple random walk model of DNA in two- and three-dimensions.

(a) In your two-dimensional model, assume that the segments can point in the plus and minus x and y directions only. Use the partition function to compute the force-extension curve for both this model and its three-dimensional analog. Plot these curves on the same graph as well as with the results for the one-dimensional model discussed in class and the fully-three-dimensional freely jointed chain discussed in class. In addition, linearize all four of these expressions and find the effective spring constant in each case.

(b) In class, I handed out several papers from Julio Fernandez in which they have measured the force-extension properties of titin. The idea espoused by Fernandez and coworkers is that after each domain of the protein ruptures, one now has an entropic spring with a new contour length. For the paper that I have, the contour lengths of the successive states of the extended protein are 52.8, 80.3, 108.1, 136.2, 164.4, 192.4, 220.7, 248.7, all measured in nanometers. Using the freely-jointed chain result in three-dimensions, make plots of the force-extension curves for all of these different contour lengths and put them all on the same plot. Comment on how well your results correspond to those reported by Fernandez et al.. NOTE: This part will require some thought and is open-ended.

5. The Benjamin Franklin Problem

Benjamin Franklin is known as a man of great curiosity. One of the scientific spirits that I most admire is those who are able to carry out extremely simple experiments in order to quantify (approximately) important physical
phenomena. A classic example is G. I. Taylor who was able to deduce the yield of the first atomic explosion (which was classified) by examining the radius of the fireball as a function of time as revealed on the cover of Time Magazine. A more recent brilliant experimenter of this ilk is Harry Swinney at the University of Texas (whose website I encourage you to visit). A figure of mythical proportions in this regard is Benjamin Franklin who carried out a simple experiment that allowed for the determination of molecular dimensions and Avogadro’s number and it is the examination of this experiment that is the charter of the present problem.

(a) What Franklin found is that a given quantity of oil always covered the same overall area. If he attempted to spread the oil over a larger area, the result was that the oil slick would split up into different parts. The conclusion that we reach is that the oil spread out to the point that it was a single monolayer thick. Given that Franklin found that $5 \text{ cm}^3$ of oil covers $2500 \text{ m}^2$ of pond surface, estimate the size of the relevant molecules. To do so, recall earlier problems in which you examined the structure of oily molecules that have polar heads and hydrophobic tails. Indeed, if you like, you can imagine that Franklin used phosphatidylcholine and work from there to see what thickness the monolayer film he found had.

(b) Continuing with the same reasoning as used in the previous problem, figure out roughly how many molecules that Franklin spread over the pond.

(c) Finally, let’s see how well we can do from these experiments at deducing the value of Avogadro’s number. For the case of oleic acid, the molar mass is $282 \text{ gm/mol}$. If the density of the oil used by Franklin is $0.9 \text{ gm/cm}^3$, how many grams did he have? Given the molar mass, how many moles did he have? Given what you found in part (b), you are now ready to deduce an approximate value of Avogadro’s number, please do so.

Consider a Cubic Bug (courtesy of Adrian Parsegian).

In class we have discussed the importance of noncovalent interactions for thinking about biological systems. In this problem, we scale up the estimates made in class to consider the ability of a bug to hold on to a ceiling.